Secretion and engineering of unspecific peroxygenases for diverse terpene oxyfunctionalization

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften
(Dr. rer. nat.)

der
Naturwissenschaftlichen Fakultät II
Chemie, Physik und Mathematik
der Martin-Luther-Universität
Halle-Wittenberg

vorgelegt von

Judith Doan-Münch

The experimental work presented within this thesis has been conducted at the Leibniz-Institute of Plant Biochemistry (IPB) in Halle (Saale) and the Martin-Luther-University Halle-Wittenberg (MLU) within the period of November 2019 to December 2023.

First Reviewer and Supervisor: Prof. Dr. Martin-Weissenborn (Martin-Luther-University Halle-Wittenberg)

External Reviewer: Prof. Dr. Sabine Flitsch (University of Manchester)

Date of Submission: 16 September 2025

Date of Defense: 16 December 2025

We share deep admiration for evolution, a force of Nature that has led to the finest chemistry of all time, and to all living things on this planet. […] Yet Nature has explored only a tiny fraction of the life and life's molecules that are possible. With evolution in our hands, with the ability to set genetic diversity and to tailor the forces of selection, we can now explore paths that Nature has left unexplored. We can also explore paths that Nature will never explore: we can select living organisms and their chemistries for our benefit—to create new sources of energy, to fix the carbon in our atmosphere, to cure disease or to make us younger. Or we can make new weapons of terror or of state control … which shall it be?

— Prof. Fances H. Arnold

Speech at the Nobel Banquet, 10 December 2018

Contents

Acknowledgement

The completion of this dissertation has been possible thanks to the support, guidance, and encouragement of many people. I would like to take this opportunity to express my sincere gratitude to those who have contributed in different ways throughout this journey.

First and foremost, I thank Prof. Dr. Martin Weissenborn for welcoming me into his group and mentoring me from my master's studies to the very end of my PhD. I am grateful for his abundance of ideas, his instinct for suggesting the right next step in moments of uncertainty, and his support when experiments produced frustrating results. I especially valued the trust he placed in me and the freedom he gave me to pursue my academic ideas. Beyond science, I enjoyed our discussions about politics and philosophy, which helped me grow in more than one dimension. I am also thankful for his support in financing the final stage of my work, which ensured that I could complete both my research and my publications under comfortable conditions.

I owe special thanks to Prof. Sabine Flitsch (University of Manchester) for agreeing to serve as the external reviewer of this dissertation and for welcoming me into her group for a one-month internship. That stay not only gave me insights into new topics in biocatalysis by exploring enzyme cascades, but also allowed me to experience another culture—both in a different country and in a different lab environment—broadening my perspective on how science can be done. I also want to thank Christian Schnepel, who mentored me during my time in Manchester and generously continued to provide advice afterwards. Thanks as well to the PhD students in Manchester who welcomed me into such a great community, making my stay motivating and inspiring for the final year of my PhD.

I am especially grateful to my colleagues in the Weissenborn lab. The biggest and warmest thanks go to Niklas Dietz, the most outstanding master's student one could wish for. At a time when I felt stuck and unsuccessful, working with Niklas reminded me of how much I had learned and how far I had come. Sharing my knowledge with him often gave me new inspiration for my own work. Since he joined, I felt steady progress in my projects, and I am grateful for the two publications we completed together. It has been a great joy to move from supervising his master's work to collaborating as colleagues, and I am glad he decided to stay in the group for his PhD.

Thanks also go to Dominik Homann, who accompanied me through most of my PhD, often rescuing me when I struggled with the GC-MS or chemical synthesis. I am equally thankful to Linda Kulka, Bonnie Winter, Li Wan, Yannick Wolf, and Vico Adjedje, as well as to Lewis, Eugen, Paul Benedikt, and Anton, for the ideas shared, the teamwork, and the good times spent together in the lab.

I also thank Nicole Hünecke, who was instrumental in helping me settle into the lab and the group, as well as Anja and Pascal, from whom I learned much about UPOs, GC-MS, and cloning, and on whose work I could build.

I am deeply grateful to my collaboration partners Prof. Sarel Fleishman, Prof. Marc Garcia-

Borràs, Jordi Soler, and Shiran Barber-Zucker for the wonderful partnerships. I am very proud of what we achieved together, combining my experimental work with your expertise in computational chemistry. Special thanks to Sarel Fleishman for in-depth discussions of strategies and results, and for the continuous refinement of our shared publications.

Both the Leibniz Institute of Plant Biochemistry (IPB) and the Biocenter of the University Halle provided excellent working environments. I was very fortunate to be surrounded by people who offered support and shared their expertise, including Dr. Franziska Seifert, Dr. Susanne Matschi, Prof. Bernhard Westermann, and Dr. Martin Dippe.

I am grateful to the Friedrich-Naumann-Stiftung für die Freiheit for generously supporting my research with a scholarship during the first three years, and to Prof. von Renesse for serving as my Vertrauensdozent, supporting my dissertation, and giving valuable feedback on my time management.

Of course, I could not have finished this dissertation without the encouragement of my friends and family. You reminded me that there was more to life than the next experiment and helped me regain perspective when things seemed overwhelming. Thank you for listening to both my excitement and my complaints, and for cheering me on throughout these years. Special thanks go to my parents, to Max, Lukas, Kristin, Daniel, and all my siblings—for their patience, belief in me, and countless ways of supporting me along the way.

Finally, the biggest thanks go to Anh Vu—for constant emotional support, for motivating me when I was frustrated, for giving sharp feedback when I needed a second opinion, and for patiently helping with everything from R scripts to LaTeX formatting. Without you, this dissertation would not have reached the finish line in the way it did.

Summary

Unspecific peroxygenases (UPOs) are a unique class of extracellular fungal enzymes capable of catalyzing a wide range of oxyfunctionalization reactions. Their portfolio includes hydroxylations of aliphatic and aromatic C-H bonds, epoxidations, heteroatom oxidations, and halogenations. Such transformations are of high value for chemical synthesis, providing access to building blocks for pharmaceuticals and fine chemicals, as well as enabling late-stage functionalization of complex molecules under mild and sustainable conditions. However, despite this versatility, UPOs face limitations that hinder their broader application. Many enzymes are not readily accessible through heterologous expression, others display insufficient stability or low activity, and their selectivity often falls short of the stringent requirements of industrial biocatalysis. As with many enzymes, systematic engineering strategies are therefore required to unlock their potential for preparative-scale use.

To place these challenges in context, the thesis first surveys the wider field of hydroxylating enzymes through a review of C-sp^3 hydroxylases. The subsequent experimental studies explore how established computational tools, combined with extensive laboratory work, can help overcome key bottlenecks in UPO engineering. Three complementary directions are pursued: (i) stability and secretion engineering to enable reliable expression of UPOs, including those initially inaccessible, (ii) activity and selectivity optimization, achieved either through structure-guided directed evolution supported by molecular dynamics (MD) simulations or through algorithm-based enzyme design, and (iii) algorithm-assisted multi-substrate engineering to improve catalytic performance across broader substrate panels.

Chapters I and II establish methodological foundations. Chapter I describes strategies for heterologous expression and secretion of UPOs in *S. cerevisiae*, emphasizing high-throughput analytics, in particular MISER-GC-MS. Chapter II applies a secretion-based expression strategy that broadens the available UPO repertoire. The hydroxylation of a pharmaceutically relevant substrate highlights the catalytic potential of the newly obtained enzymes, including identification of *Mth*UPO as a promising starting point for engineering.

Chapter III expands the accessible UPO repertoire by combining computational stability design with secretion engineering. Ab initio structural models were analyzed with the PROSS algorithm, generating enzyme design with stabilizing mutation and a signal-peptide shuffling was applied together enabling the functional secretion of nine out of ten enzymes. This approach allowed the secretion of several UPOs previously inaccessible as wild type enzymes, including the first experimentally characterized non-fungal UPOs. The determiniation their activity towards selected substrates and their thermostability further benchmarked these UPOs. Overall, these study shows that computational stability design, together with secretion engineering, can substantially broaden the UPO toolbox.

Chapter IV targeted the improvement of enantioselectivity in terpene hydroxylation using

*Mth*UPO as a model. Starting from poor enantioselectivity, a structure-guided directed evolution campaign based on a homology model was performed, utilizing a high-throughput screening technique. Two consecutive rounds of computationally guided mutagenesis and screening yielded variants with enhanced activity, regioselectivity, and enantioselectivity, producing both enantiodivergent products. This study illustrates how density functional calculations and MD simulations can guide engineering campaigns combined with experimental screening fine-tune enzyme properties. MD simulations also provided supporting insights into the observed functional changes of the improved enzyme variants.

Chapter V adopts a different strategy, using the FuncLib algorithm to design a compact library of *Mth*UPO designs for a panel of terpene substrates. All algorithm-designed variants were secreted successfully, and screening revealed 2.2-7.1-fold increases in activity relative to the wild type, as well as altered regio-, chemo-, and enantioselectivity. In contrast to the stepwise nature of directed evolution, this approach allowed parallel optimization across multiple substrates from a single design round.

These studies show how computational tools can support enzyme engineering—by enabling secretion of otherwise inaccessible UPOs, guiding mutational strategies for activity and selectivity, or generating designed variant libraries even for improving multiple substrates in parallel. Experimental validation and screening were essential to realize these advances, selected reactions were demonstrated at preparative scale, yielding milligram quantities of products underscoring the applicability of engineered UPOs beyond analytical assays. By comparing different computationally tools and engineering strategies, the thesis highlights their respective advantages and limitations, illustrating how the choice of approach can be tailored to the specific engineering goal. In this way, the work provides guidance for selecting suitable combinations of computational and experimental methods for engineering UPOs and, more broadly, for monooxygenases.

Overall, this thesis illustrates how combining computational and experimental approaches can systematically advance UPOs toward more versatile and robust biocatalysts. The studies demonstrate that key challenges—such as limited accessibility, insufficient stability, and suboptimal activity or selectivity—can be addressed, providing practical strategies to improve enzyme performance. By establishing methodological foundations and generating well-characterized variants, the work takes important steps toward bringing these versatile catalysts closer to practical use and lays a platform for future exploration of their potential.

Zusammenfassung

Unspezifische Peroxygenasen (UPOs) stellen eine einzigartige Klasse extrazellulärer Pilzenzyme dar, die in der Lage sind, ein breites Spektrum an Oxyfunktionsalisierungsreaktionen zu katalysieren. Ihr Reaktionsspektrum umfasst Hydroxylierungen von aliphatischen und aromatischen C-H-Bindungen, Epoxidierungen, Heteroatom-Oxidationen sowie Halogenierungen. Solche Transformationen sind für die chemische Synthese von hohem Wert, da sie Zugang zu Bausteinen für Arzneimittel und Feinchemikalien bieten und die späte funktionelle Modifikation komplexer Moleküle unter milden und nachhaltigen Bedingungen ermöglichen. Trotz dieser Vielseitigkeit stoßen UPOs jedoch auf Einschränkungen, die ihre breitere Anwendung behindern. Viele Enzyme sind nicht einfach heterolog exprimierbar, andere zeigen unzureichende Stabilität oder geringe Aktivität, und ihre Selektivität entspricht häufig nicht den strengen Anforderungen der industriellen Biokatalyse. Wie bei vielen Enzymen sind daher systematische *Engineering*-Strategien erforderlich, um ihr Potenzial für den präparativen Einsatz zu erschließen.

Um diese Herausforderungen einzuordnen, untersucht die Arbeit zunächst das Feld der C-sp^3 hydroxylierenden Enzyme anhand eines Reviews. In den anschließenden experimentellen Studien wird untersucht, wie etablierte computergestützte Werkzeuge in Kombination mit experimenteller Laborarbeit helfen können, zentrale Engpässe im UPO-*Engineering* zu überwinden. Drei komplementäre Ansätze werden verfolgt: (i) Stabilitäts- und Sekretions-*Engineering* zur zuverlässigen Expression und Sekretion von UPOs, einschließlich solcher, die zunächst nicht zugänglich waren, (ii) Optimierung von Aktivität und Selektivität, entweder durch strukturgeleitete *directed evolution* unterstützt von Molekulardynamik-(MD)-Simulationen oder durch algorithmusbasiertes Enzymdesign, und (iii) algorithmusgestütztes Multi-Substrat-*Engineering* zur Verbesserung der katalytischen Leistung über ein breiteres Substratpanel.

Die Kapitel I und II legen methodische Grundlagen. Kapitel I beschreibt Strategien zur heterologen Expression und Sekretion von UPOs in *S. cerevisiae*, mit Schwerpunkt auf Hochdurchsatzanalytik, insbesondere MISER-GC-MS. Kapitel II wendet eine sekretionsbasierte Expressionsstrategie an, die das verfügbare UPO-Repertoire erweitert. Die Hydroxylierung eines pharmakologisch relevanten Substrats hebt das katalytische Potenzial der neu gewonnenen Enzyme hervor, einschließlich der Identifizierung von *Mth*UPO als vielversprechendem Ausgangspunkt für das *Engineering*.

Kapitel III erweitert das zugängliche UPO-Repertoire durch die Kombination von computergestütztem Stabilitätsdesign mit Sekretions-*Engineering*. Ab-initio-Strukturmodelle wurden mit dem PROSS-Algorithmus analysiert, um Enzymdesigns mit stabilisierenden Mutationen zu generieren, zusätzlich wurde und ein Signalpeptid-*Shuffling* wurde angewendet. Neun von zehn Enzyme konnten funktional sekretiert werden unter anderem mehrerer als Wildtyp nicht zugängliche, einschließlich der ersten experimentell charakterisierten nicht-pilzlichen UPOs. Die Bestimmung ihrer Aktivität gegenüber ausgewählten Substraten und ihrer Thermostabilität lieferte weitere

Vergleichsdaten. Insgesamt zeigt die Studie, dass computergestütztes Stabilitätsdesign zusammen mit Sekretions-*Engineering* das UPO-Werkzeug erheblich erweitern kann.

Kapitel IV fokussiert sich auf die Verbesserung der Enantioselektivität bei der Hydroxylierung von Terpenen unter Verwendung von *Mth*UPO als Biokatalysator. Ausgehend von geringer Enantioselektivität wurde eine strukturgeleitete gerichtete Evolutions-Kampagne auf Basis eines Homologiemodells durchgeführt, unter Verwendung einer Hochdurchsatz-*Screening*-Technik. Zwei aufeinanderfolgende Runden computergestützter Mutagenese und *Screening* führten zu Varianten mit verbesserter Aktivität, Regioselektivität und Enantioselektivität, die enantiodivergente Produkte erzeugten. Diese Studie zeigt, wie Dichtefunktionaltheorie-Berechnungen und MD-Simulationen *engineering*-Kampagnen leiten und in Kombination mit experimentellem *screening* Enzymeigenschaften feinjustieren können. MD-Simulationen lieferten zudem weitere Einblicke in die beobachteten funktionellen Veränderungen der verbesserten Enzymvarianten.

Kapitel V verfolgt eine alternative Strategie und nutzt den FuncLib-Algorithmus, um eine kompakte Bibliothek von *Mth*UPO-Designs für ein Panel von Terpensubstraten zu erstellen. Alle algorithmusgestalteten Varianten wurden erfolgreich sekretiert, und *screening* zeigte 2,2- bis 7,1-fache Aktivitätssteigerungen im Vergleich zum Wildtyp sowie veränderte Regio-, Chemo- und Enantioselektivität. Im Gegensatz zur schrittweisen Natur der gerichtete Evolution ermöglichte dieser Ansatz die parallele Optimierung von mehreren Substraten in einer einzigen Designrunde.

Diese Studien zeigen, wie computergestützte Werkzeuge das Enzym-*engineering* unterstützen können—sei es durch die Ermöglichung der Sekretion sonst nicht zugänglicher UPOs, durch das Leiten von Mutationsstrategien zur Verbesserung von Aktivität und Selektivität oder durch die Generierung von Enzym-Design-Bibliotheken zur gleichzeitigen Verbesserung mehrerer Substrate. Experimentelle Validierung und *screening* waren entscheidend, um diese Fortschritte zu realisieren; ausgewählte Reaktionen wurden auf präparativer Skala demonstriert und lieferten Milligramm-Mengen an Produkten, was die Anwendbarkeit der UPOs über analytische Assays hinaus unterstreicht. Durch den Vergleich verschiedener computergestützter Werkzeuge und *engineering*-Strategien werden deren Vor- und Nachteile aufgezeigt und illustriert, wie die Wahl der Strategie auf das spezifische *engineering*-Ziel zugeschnitten werden kann. Auf diese Weise bietet die Arbeit Orientierung bei der Auswahl geeigneter Kombinationen von computatorischen und experimentellen Methoden für das UPO-*engineering* und, weiter gefasst, für Monooxygenasen.

Insgesamt zeigt diese Arbeit, wie mit diese Methoden UPOs systematisch zu vielseitigeren und robusteren Biokatalysatoren entwickelt werden können. Die Studien demonstrieren, dass zentrale Herausforderungen—wie begrenzte Zugänglichkeit, unzureichende Stabilität und suboptimale Aktivität oder Selektivität—adressiert werden können, und liefern praktische Strategien zur Verbesserung der Enzyme. Durch die Etablierung methodischer Grundlagen und die Generierung gut charakterisierter Varianten unternimmt die Arbeit wichtige Schritte, um diese vielseitigen Katalysatoren der praktischen Anwendung näherzubringen und eine Plattform für die zukünftige Erforschung ihres Potenzials zu schaffen.

List of Figures

List of Schemes

[2]The asterisk (*) indicates an engineered variant.

List of Tables

List of Acronyms

***Aae*UPO** UPO from *Agrocybe aegerita*
ABTS 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid)
AF2 AlphaFold2
AI artificial intelligence
ASR ancestral sequence reconstruction

***Bci*UPO** UPO from *Botrytis cinerea*
***Bla*UPO** UPO from *Bremia lactucae*
BM3 cytochrome P450 from *Bacillus megaterium*

CAST combinatorial active-site saturation test
***Cci*UPO** UPO from *Coprinopsis cinerea*
CD consensus design
***Cgr*UPO** UPO from *Colletotrichum graminicola*
***Chi*UPO** UPO from *Colletotrichum higginsianum*
CO carbon monoxide
CYP cytochrome P450

DFT density functional theory
DMP 2,6-dimethoxyphenol
DNA desoxyribonucelic acid

ee enantiomeric excess
EI electron-impact ionization

***Fgr*UPO** UPO from *Fusarium graminearum*
***Fox*UPO** UPO from *Fusarium oxysporum*
FRISM focused rational iterative site-specific mutagenesis

GC gas chromatography
GFP green fluorescent protein
***Gma*UPO** UPO from *Galerina marginata*

ISM iterative saturation mutagenesis

MD molecular dynamic
MISER multiple injection in a single experimental run
ML machine learning
MM molecular mechanics
***Mro*UPO** UPO from *Marasmius rotula*

MS mass spectrometry
***Mth*UPO** UPO from *Myceliophthora thermophila*

NAC near-attack conformation
NADH nicotinamide adenine dinucleotide
NADPH nicotinamide adenine dinucleotide phosphate
nanoDSF nano differential scanning fluorimetry
NBD 5-nitro-1,3-benzodioxole

***Pca*UPO** UPO from *Phytophthora infestans*
***Pin*UPO** UPO from *Phytophthora cactorum*
P. pastoris *Pichia pastoris*
PROSS Protein Repair One-Stop Shop

QM quantum mechanics
QM/MM quantum mechanics / molecular mechanics

S. cerevisiae *Saccharomyces cerevisiae*
SUMS substrate multiplexed screening

TON turnover number
TS transition state

UPO unspecific peroxygenase

1 Introduction

1.1 Enzymatic C-H oxyfunctionaliszation

This section is taken from my review paper published in ACS Catalysis in 2021.[1]

C-H bonds are the most common structural units in organic compounds. The selective activation and transformation of C (sp^3)-H bonds into more active functional groups is expected to substantially broaden the synthetic potential inorganic chemistry.[2] However, the general inertness of C-H bonds renders their selective transformation one of the most challenging reactions. Aliphatic C-H bond oxidations, for example, for late-stage functionalizationin synthesis were achieved using acid, H_2O_2 and an electrophilic nonheme Fe(PDP) catalyst.[3–5]

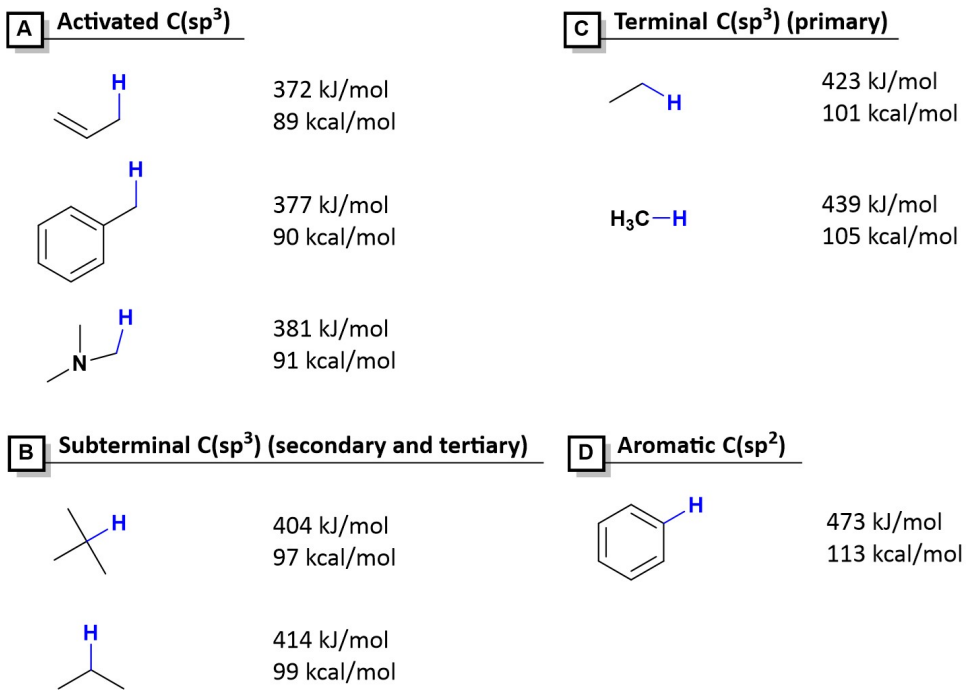
Enzymes are evolved by nature and often combine excellent activity and selectivity, rendering them powerful catalysts for C(sp^3)-H bond functionalization. Hydroxylating enzymes catalyze the oxyfunctionalization of versatile substrates unlocking a broad range of compounds in the primary and secondary metabolism in all kingdoms of life. They directly utilize dioxygen or hydrogen peroxide and are essential catalysts for the natural functions of all aerobic life forms. In anaerobic organisms, hydroxylating enzymes utilizing water as the oxygen source are present.[6] However, it is not yet fully understood, why such a diversity of enzymes with similar functions exist.

When molecular oxygen is supplied as the donor, two electron-equivalents need to be provided as both oxygen atoms are reduced from oxidation state 0 to -II, whereas the carbon is oxidized only by two oxidation states. Monooxygenases, which transfer one oxygen into the substrate and form water from the second oxygen atom, require electrons provided by electron transport chains often commencing from NAD(P)H.[7] Dioxygenases obtain the overall four electrons required for reducing molecular oxygen from oxidizing two carbon atoms, therefore always performing dihydroxylation reactions.[8,9] For hydrogen peroxide as the oxygen source, no external electron transfer is required since both oxygens are only reduced from -I to -II and hence the electron demand is balanced by the carbon oxidation. Although only a few examples exist, the molybdoenzymes utilizing water as electron source even release two electrons, as the oxygen is already in the oxidation state -II.[10,11]

The hydroxylation of sp^3-hybridized carbons mostly takes place through homolytic C-H bond cleavage. For Mo-dependent enzymes, there is still a debate about this. The mechanism of most of the discussed flavin-dependent enzymes does not take place via a radical mechanism. The stability of a carbon radical determines the bond dissociation energy (Scheme 1).[12] The higher the dissociation energy, the generally fewer enzymes are available for its hydroxylation. This observation allows judging the hydroxylation potential of an enzyme from the converted carbon substrates. An enzyme that can hydroxylate terminal carbons can also hydroxylate all carbons with lower bond dissociation energies like benzylic or allylic carbons. However, hydroxylating biocatalysts that hydroxylate benzylic carbons, are not necessarily able to functionalize nonactivated carbons. Methane is a particular case as only specific methane monooxygenases can hydroxylate this compound having the highest bond dissociation energy (105 kcal/mol) of all C(sp^3)-H bonds. As one would predict from this high hydroxylating potential, methane monooxygenases can oxidize a broad range of substrates.[13,14]

The bond dissociation energy of aromatic C(sp^2)-H bonds (113 kcal/mol) is higher than all C(sp^3)-H bonds. This high bond dissociation energy is caused by the formed radical upon C-H dissociation not being able to stabilize in the aromatic ring. Nonetheless, aromatic hydroxylation reactions are substantially easier to perform than C(sp^3)-H modifications due to an altered reaction mechanism. The reaction does not proceed via a homolytic cleavage but rather via an initial epoxide intermediate, leading to the loss of aromaticity followed by a heterolytic H-abstraction and rearomatization, known as NIH shift mechanism (Scheme 2). These differences between

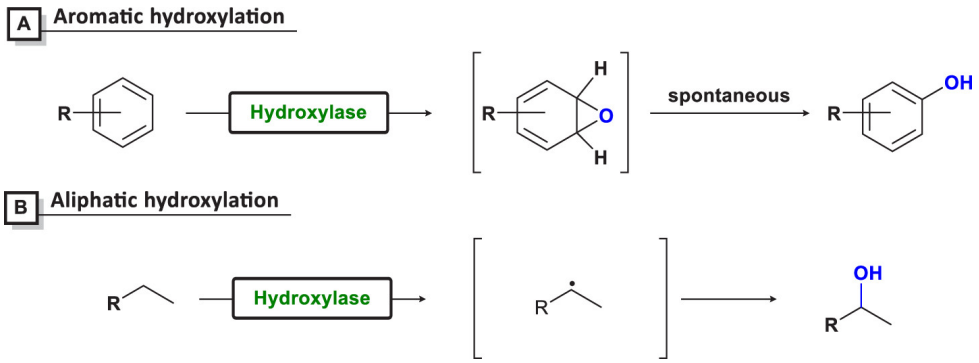
Scheme 1 C-H Bond Dissociation Energy of Selected Organic Molecules



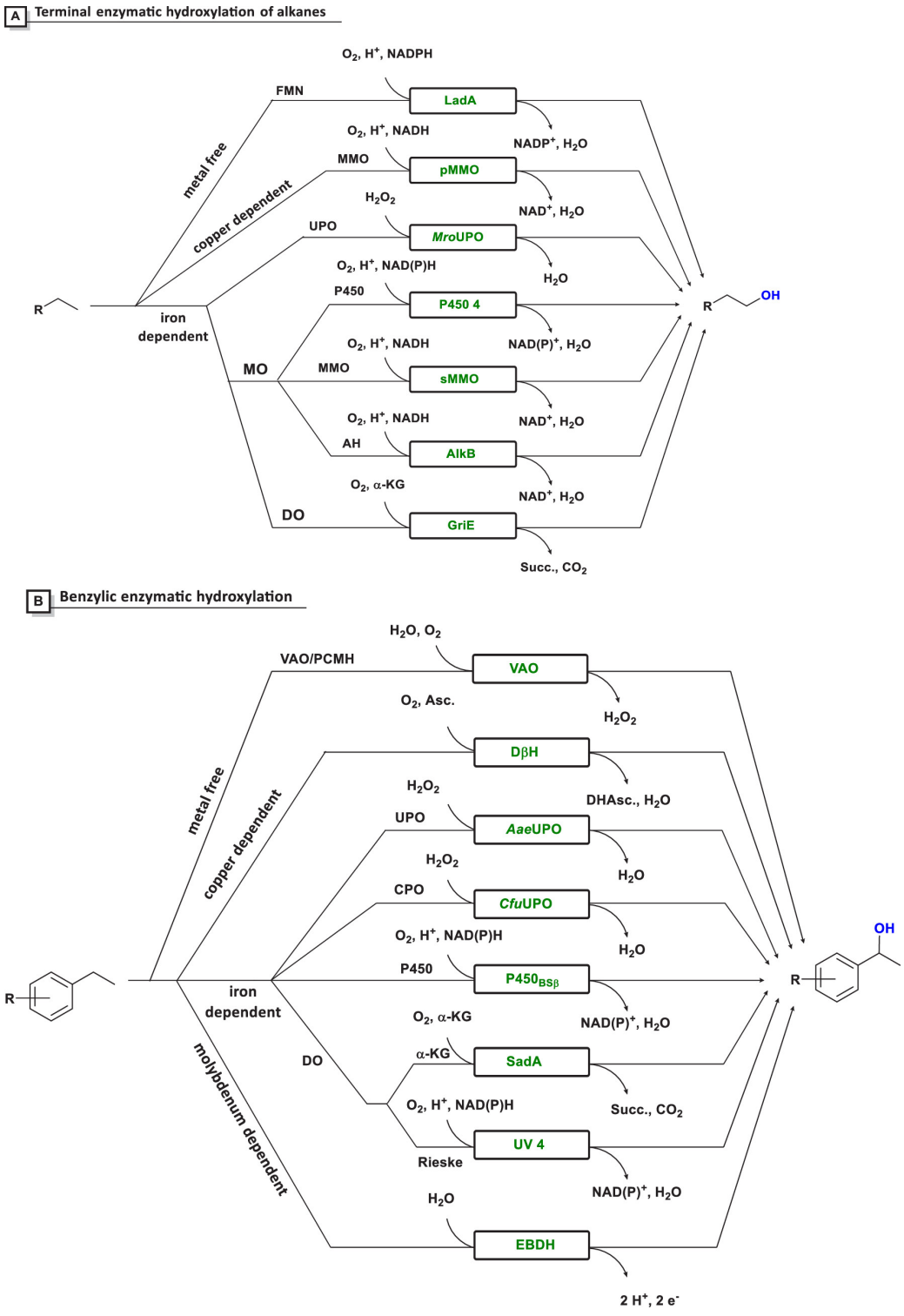
the reaction mechanism of C(sp^2) and C(sp^3) hydroxylation highlight why C(sp^3) hydroxylation reactions are more difficult to perform. Ullrich and Hofrichter gave an overview of the enzymatic hydroxylation of aromatic compounds.[15]

Our review focuses on enzymes performing C(sp^3) hydroxylation reactions. We structured it by nature of the prosthetic group: metal-free, Fe-, Cu- or Mo-depending proteins. Scheme 3 gives an overview of the different enzyme classes that can hydroxylate primary and benzylic carbons. Table 1 summarizes the enzymes discussed in this Review and lists their prosthetic group, associated proteins, cofactors, and positions that can be hydroxylated.

Scheme 2 Different Mechanisms of Enzymatic Aromatic and Aliphatic Hydroxylation



Scheme 3 Overview of Hydroxylating Enzymes at Terminal and Benzylic Positions Discussed in This Review.



1.2 Heme-Containing Enzymes

Heme, a Fe-containing metalloporphyrin, serves as a prosthetic group in numerous proteins that mediate essential redox processes across all domains of life. It facilitates electron transfer between proteins, enables oxygen transport and storage, and plays a central role in various enzymatic oxidations. While many heme enzymes utilize molecular oxygen (O_2) as both oxidant and oxygen source, others employ hydrogen peroxide (H_2O_2) as a prereduced alternative. The following sections introduce two heme-thiolate enzyme families relevant to this context: unspecific peroxygenases (UPOs), which form the core focus of this dissertation and are applied throughout the experimental chapters, and cytochrome P450 monooxygenases (P450s), which catalyze mechanistically and functionally related oxidations. The comparative discussion of P450s provides a reference framework for evaluating the biocatalytic potential and distinct features of UPOs.

1.2.1 Unspecific peroxygenases

This section is taken from my review paper published in ACS Catalysis in 2021.[1]

> Unspecific peroxygenases (UPOs) are heme-thiolate proteins forming a unique, solely fungal (exceptions are oomycetes) family of secreted, extracellular enzymes. Other enzymes sharing the term peroxygenases are plant peroxygenases, which are integral membrane proteins and oxidize unsaturated fatty acids thereby forming epoxide products. Another enzyme group called peroxygenases are P450 peroxygenases, which are part of the cytochrome P450 family and can hydroxylate fatty acids.[16] This subgroup will be discussed in the cytochrome P450 section below. Further oxidoreductases also named peroxygenases are 3-methyl-L-tyrosine peroxygenase[17] and L-tyrosine peroxygenase.[18]
>
> UPOs can catalyze a large variety of reaction types comparable to P450 monooxygenases. These include the hydroxylation of aliphatic and aromatic compounds, epoxidation, heteroatom oxidation, such as sulfoxidation/*N*-hydroxylation and chemoenzymatic halogenation reactions, the latter not being reported for P450s.[19] The first UPO was described in 2004 by Hofrichter and co-workers occurring in the culture supernatant of the basidiomycete *Agrocybe aegerita* (*Aae*UPO) commonly known as black poplar mushroom. It was initially classified as a haloperoxidase because of its ability to oxidize halides and aryl alcohols similar to chloroperoxidases (CPOs).[20] Further UPOs were discovered in *Marasmius rotula* (*Mro*UPO) and *Coprinellus radians* (*Cra*UPO), both also belonging to the order of Agaricales known as "gilled mushrooms" of the Basidiomycota phylum. On the basis of further investigations regarding the catalyzed reactivities, the class descriptor changed from haloperoxidase-peroxygenase to aromatic peroxygenase and finally to the currently used name unspecific peroxygenase.[21–23] It would not be surprising if a more eligible name will arise for this enzyme class in the future as it continuously exhibits excellent selectivity for C–H oxyfunctionalizations[24] and "unspecific" has a rather negative implication for an enzyme. UPOs are generally classified into two groups: The Group-I short-type UPOs with an average mass of 29 kDa and the Group-II long-type UPOs with an average mass of 44 kDa. Group-I UPOs are widespread within the whole fungal phyla, whereas Group-II UPOs have so far been only found in Ascomycota and Basidiomycota, while also not being omnipresent within them. UPOs are not found in Taphrinomycotina and Sacharomycotina, for example, in *Saccharomyces cerevisiae* (true yeast) and *Schizosaccharomyces* (fission yeast).[25] The available large set and diversity of long and short-type UPOs are emphasized by identifying more than 4300 putative UPO-encoding sequences.[26,27] When compared to P450s, UPOs generally exhibit higher stability and water solubility correlated to the naturally occurring substantial degree of glycosylation.[28] Because of their high stability, the

independence of other colocalized associated proteins such as reductases and reduction agents, sufficient protein secretion and the usage of H_2O_2 instead of molecular oxygen, UPOs are on the brink of becoming a viable catalyst for industrial applications. So far, impressive TONs of up to 300,000 for the stereoselective benzylic hydroxylation of ethylbenzene were achieved.[29] This high activity was further harnessed in initial upscaling reactions to reach g/L titers of the enantiopure product.[30] Recent application focussed work includes the utilization of UPOs for the epoxidation of various styrene derivates under neat conditions, resulting in high product concentrations of up to 360 mM.[31]

The natural function of UPOs still remains elusive. However, experiments could prove that they can convert 35 out of 40 compounds listed as priority pollutants by the Environmental Protection Agency of the United States (EPA), thus showing a pronounced detoxification and biodegradation ability.[32] Combined with their extracellular occurrence, it is assumed that UPOs might be involved in the detoxification of their surrounding microenvironment.[32] They are assumed to complement P450s, which are key players of intracellular detoxification. Initial productions of UPOs were performed homologously as (fed)-batch cultivation of the natural host organism in shake flasks or bioreactors using optimized media compositions that allowed for a maximal volumetric UPO titer of 41,000 U veratryl alcohol/L, which correlates roughly with 445 mg/L peroxygenase enzyme in the case of *Mro*UPO.[20,33,34] Initial heterologous production of a UPO derived from *Coprinopsis cinerea* (*Cci*UPO) could be achieved by utilizing an industrial strain of the mold *Aspergillus oryzae*, which is an attractive host for industrial protein production because of the extremely high natural protein secretion capacity.[35]

However, the lack of a heterologous expression platform for UPOs using a fast-growing, easy to manipulate and highly transformable host has been a strongly limiting factor ever since. Alcalde and co-workers performed an intensive directed evolution campaign using the yeast *S. cerevisiae* that led to the successful expression of the *Aae*UPO variant PaDa-I. This secretion variant could be successfully produced in *S. cerevisiae*, bearing nine amino acid exchanges compared to the fungal wild type enzyme reaching titers of 6500 U ABTS/L and 1300 U NBD/L, respectively, corresponding to approximately 8 mg/L.[36] Especially four mutations within the signal peptide led to a boosted secretion by the heterologous yeast host. Subsequently, this variant was also produced on a larger scale with titers of 5826 U ABTS/L and 728 U NBD/L, respectively, using the methylotrophic yeast *Pichia pastoris* and a standard genomically integrated methanol-inducible (AOX1 promoter) system.[37] These comparable activity values indicate that the shake flask format's secretion levels were maintained irrespective of the yeast used. Fed-batch fermentation led to 27-fold enhanced production of the UPO in *P. pastoris* (217 mg/L).[37]

Recent work in our group has aimed to construct a modular yeast UPO expression system allowing the shuffling of a panel of versatile signal peptides for optimal secretion and thereby aiding the quick verification and production of already described as well as novel wildtype UPOs in *S. cerevisiae*. The system was developed as a Golden Gate platform allowing the rapid assembly of 17 different signal peptides, eight genes and a range of C-terminal protein tags for purification and split-GFP-based protein quantification (Yeast Secrete and Detect Kit; Addgene). By harnessing this system, we could expand the panel of recombinant UPO enzymes by the wild type enzymes *Mwe*UPO (*Marasmius wettsteinii*), *Mro*UPO, and *Cgl*UPO. Two putative peroxygenases derived from thermophilic fungal organisms—*Myceliophthora thermophila* (*Mth*UPO) and *Thielavia terrestris* (*Tte*UPO)—were produced and characterized for the first time (Table 1). The modular episomal signal peptide shuffling expression platform could furthermore be adapted to *P. pastoris*. A compatible integration plasmid offers access to genomically stable integrated UPO producer strains. The high obtained *Mth*UPO expression titers of up to 24 mg/L in shake flasks furthermore enabled enantioselective C(sp^3) hydroxylation reactions on a preparative scale.[38] Subsequent work on episomal UPO production in *P. pastoris* led to elevated production titers of known UPOs as well as the first heterologous production of *Mhi*UPO (*Myceliophthora hinnulea*) and *Mfe*UPO (*Myceliophthora fergusii*).[39]

Table 1 Successful Production of Different Unspecific Peroxygenases in Diverse Organisms

			hetereologous expression in				homologous
organism	enzyme	group	*S. cerevisiae*	*E. coli*	*A. oryzae*	*P. pastoris*	expression
Agrocybe aegerita	*Aae*Upo	group-II	yes[36a]			yes[37a]	yes[20]
Marasmius rotula	*Mro*Upo	group-I	yes[38]	yes[40]		yes[38]	yes[33]
Coprinopsis cinerea	*Cci*Upo	group-II			yes[35]		
Daldinia caldariorum	*Dca*Upo	group-I		yes[41]		yes	
Collariella virescen	*Cvi*Upo	group-I		yes[41]			
Coprinus radians	*Cr*Upo	group-II					yes[34]
Marasmius wettsteinii	*Mwe*Upo	group-I	yes[38]				yes[42]
Chaetomium globosum	*Cgl*Upo	group-I	yes[38]			yes[38]	yes[43]
Humicola insolens	*Hin*Upo	group-I			yes[43]		
Psathyrella aberdarensis	*Pab*Upo	group-II					yes[44]
Myceliophthora thermophila	*Mth*Upo	group-I	yes[38]			yes[38]	
Thielavia terrestris	*Tte*Upo	group-I	yes[38]			yes[38]	
Myceliophthora fergusii	*Mfe*Upo	group-I				yes[39]	
Myceliophthora hinnulea	*Mhi*Upo	group-I				yes[39]	

[a] No expression of wildtype UPO but evolved mutant

Recently, the expression of UPOs utilizing a prokaryotic host were achieved leading to the production of *Cvi*UPO[45] (*Collariella virescens*), *Mro*UPO,[38] and *Dca*UPO[41] (*Daldinia caldariorum*) in *E. coli* (Table 1). However, the described system needs to be further studied to compare homologously, yeast and *E. coli* expressed UPOs as the post-translational modifications were previously shown to be highly important.[38]

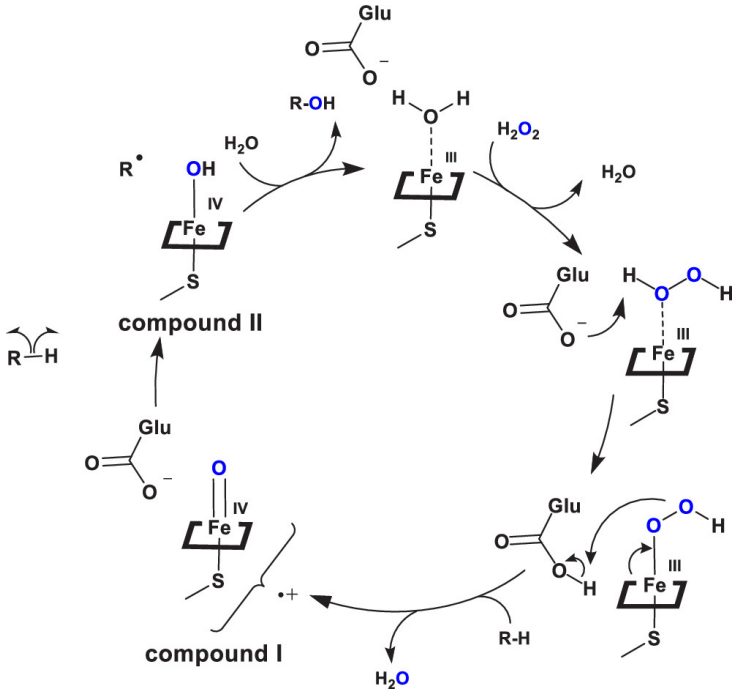
UPOs carry a heme cofactor with a highly conserved cysteine proximal axial ligand flanked by two proline residues (PCP motif). Another highly conserved residue is the catalytic acid glutamate (Glu196 in *Aae*UPO, Glu158 in *Mro*UPO), which acts together with arginine (long-type) or histidine (short-type) as acid-base catalyst pair. The PCP motif and the glutamate residue are also present in the closely related CPO.

The catalytic cycle of UPOs is similar to P450s and other monooxygenase enzymes. The resting-state heme contains an Fe^{3+} ion that reacts with H_2O_2 forming a Fe-peroxo-complex. This complex is heterolytically cleaved and water is released under the aid of the catalytic acid glutamate. After electron rearrangement the key intermediate compound I is formed.[46,47] The kinetics of this oxo-ferryl cation radical complex and its formation and decomposition have been studied intensively.[48,49] In a subsequent reaction, UPO compound I abstracts a hydrogen atom from the substrate (R-H) yielding protonated compound II, a ferryl hydroxide complex and a substrate radical (R•). The radical quickly recombines with the hydroxide leading to the alcohol product (Scheme 4).

UPOs are capable of hydroxylating linear, branched, and cyclic aliphatic, as well as aromatic, compounds. They can catalyze hydroxylations of activated, benzylic, subterminal, and terminal C(sp^3). Depending on the enzyme and reaction conditions, overoxidation reactions can occur, leading to aldehyde, ketone, or carboxylic acid products. Short alkanes can be hydroxylated, starting at n-propane (C3). n-Propane and n-butane are hydroxylated by *Aae*UPO giving rise to 2- propanol and 2-butanol, longer alkanes predominantly yield a mixture of 2-alcohol and 3-alcohol.[50]

*Mro*UPO catalyzes the hydroxylation and epoxidation of unsaturated fatty acids such as oleic acid (**29**) (Scheme 5A). Hydroxylation mainly takes place subterminal (ω-1) **31** and at position ω-7 **30**. Targeted amino acid exchanges within the entrance tunnel drastically modulated the epoxida-

Scheme 4 Catalytic Cycle of Unspecific Peroxygenases(UPOs)



tion/ hydroxylation ratio of the enzyme. It was possible to increase the epoxidation ratio but also to abolish it completely so that only terminal and subterminal hydroxylation at C(sp^3) takes place. k_{cat}/K_m values of 2.1 × 10^4 M^{-1} s^{-1} (ω-1 hydroxylation of oleic acid) were obtained with variant *Mro*UPO I153F/S156F which abolishes epoxidation.[40] In contrast to *Aae*UPO and *Cci*UPO, which exclusively hydroxylate tetradecane at the subterminal position, *Mro*UPO can catalyze the terminal hydroxylation of dodecane (408 TON) and tetradecane (**33**) (240 TON), followed by overoxidation to a dicarboxylic acid (Scheme 5B). Some of the terminal-oxygenation products **34** also displayed additional oxygenation at subterminal (ω-1 and ω-2) positions. Concluding, this peroxygenase has high potential to be used as catalyst for late-stage functionalization of organic molecules due to its ability to hydroxylate the least activated terminal position.[51]

Utilizing the above-described modular secretion system enabled the protein engineering of the newly discovered *Mth*UPO. By screening more than 5000 transformants, the catalytic efficiency on the conversion of the model substrate 5- nitro-1,3-benzodioxole (NBD) was improved 16.5-fold. Improved variants led to the chemodivergent formation of selective oxyfunctionalizations at the aromatic **37** and benzylic positions **36**, respectively, using 2-methylnaphthalene (**35**) as a substrate (Scheme 5C).[52]

*Cgl*UPO from *Chaetomium globosum* can catalyze the epoxidation and hydroxylation of testosterone (**38**) yielding testosterone 4,5-β-epoxide (**40**) and 16α-hydroxy testosterone (**39**) (Scheme 5D). Those steroids exhibit crucial roles in physiological processes. 4,5-β-epoxide (**40**) regulates cell proliferation and cholesterol homeostasis, whereas 16α-hydroxy testosterone (**39**) is a metabolite in the oxidative metabolism in liver cells and an essential intermediate in the formation of estriol in late pregnancy.[53–55] In contrast, other known UPOs, such as *Aae*UPO, *rCci*UPO, r*Hin*UPO, and *Mro*UPO failed to hydroxylate the testosterone's gonane ring in comparative experiments.[43] Epoxidation is strongly favored over the hydroxylation with a ratio of 9:1. Both products **39** and **40** could be purified with high purity (>96%) on a preparative scale. Overall turnover numbers of up to 7000, including the epoxidation, and a k_{cat}/K_m value of 1.60 × 10^3 M^{-1} s^{-1} were determined, which is moderate in contrast to other peroxygenase activities and substrates.[43] Nevertheless, it could be demonstrated that *Cgl*UPO might be a promising catalyst for the oxyfunctionalization of bulky steroid core struc-

tures. The successful heterologous production of *Cgl*UPO in *S. cerevisiae* and *P. pastoris*[38] can enable subsequent protein engineering.

The enantioselective hydroxylation of phenethyl amine derivatives at the benzylic position leads to many pharmaceutically important substances such as beta-blockers and sympathomimetics.[56] As shown by our group, *Mth*UPO and *Cgl*UPO can catalyze the benzylic hydroxylation of naphthaloyl-phenylethyl amine (**41**) with high enantioselectivities of >98% ee. We were able to perform this reaction in a proof-of-concept experiment using *Mth*UPO on a preparative scale with 57% yield of the hydroxylated product **42** (Scheme 5E).[38]

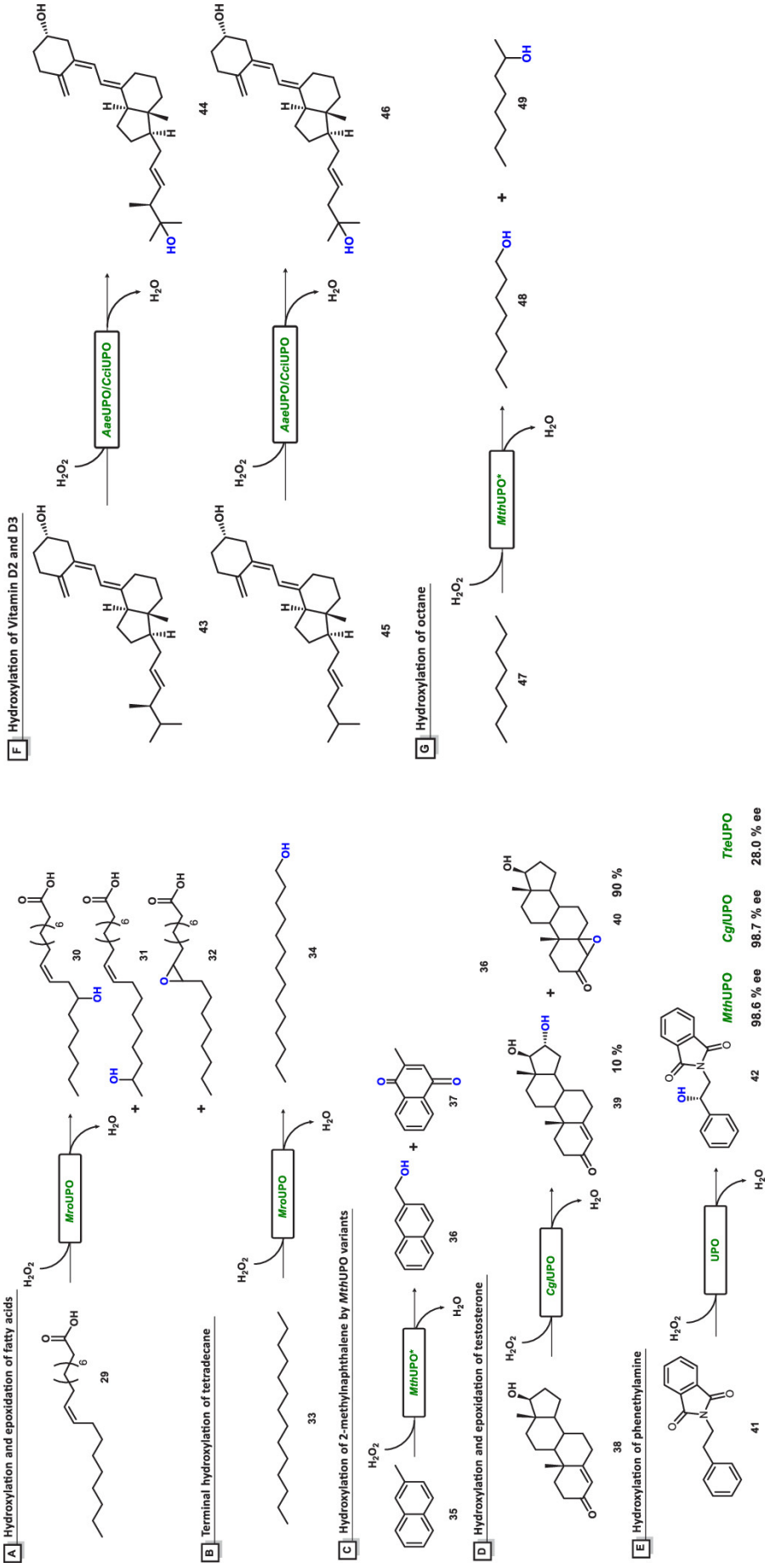
The hydroxylation of vitamin D2 (ergocalciferol) (**43**) and D3 (cholecalciferol) (**45**) to 25-monohydroxylated vitamin D2 (**44**) and 25-monohydroxylated vitamin D3 (**46**) is of high interest as these hydroxylated compounds show positive effects on different human diseases and raise interest for feeding broiler chickens and other farm animals to reduce skeleton problems.[57–59] The conversion of ergocalciferol (**43**) and cholecalciferol (**45**) can be catalyzed by *Aae*UPO, as well as *Cci*UPO (Scheme 5F). The cholecalciferol conversion yielded 100% product **46** formation after 60 min of reaction catalyzed by *Cci*UPO and only a slightly lower yield when catalyzed by *Aae*UPO (90%). Ergocalciferol conversion was also observed to be similar for both enzymes (81–86%). *Cci*UPO converted both substrates **43** and **45** regioselectively. *Aae*UPO only catalyzed ergocalciferol's regioselective conversion, whereas several byproducts were formed in the course of vitamin D3 hydroxylation. These differences could be rationalized computationally and obtained differences in the energy profile confirmed the experimental data.[60,61]

Working with enzymes in chemical reactions implies some challenges, which can be addressed through protein and reaction engineering to improve product formation and reduce reaction costs. We will briefly discuss some solutions addressing the challenges of peroxygenases.

To increase thermostability and the tolerance toward organic solvents and H_2O_2, directed evolution and neutral genetic drift approaches were performed by Alcalde and coworkers leading to variants with increased stabilities (increased T_m up to 4.9 fold) and stability toward acetonitrile (up to 2.9-fold, from C_{50}(%) 7.0 to 20.3). However, increases in stability partially lead to decreases in activity.[36,62] H_2O_2 has a Janus role in peroxygenases: on the one hand, it is indispensable for the peroxygenase activity as it omits cosubstrates for the reduction. On the other hand, the heme inactivation through oxidative damage is a major challenge when using H_2O_2 as cosubstrate. Supplying the reaction over time with H_2O_2 instead of adding it stoichiometrically has proved to be a beneficial strategy to achieve maximal product formation. Several smart procedures to maintain low H_2O_2 concentrations by continuous in situ H_2O_2 generation systems have been developed, for example, the utilization of biocatalysts such as formate oxidase (Fox) were employed.[63–66] Further concepts are based on the usage of light-driven reactions,[67] gold-palladium nanoparticles,[63] or radiolytic water splitting to produce H_2O_2.[68,69]

Thus, far, only one long-type UPO could be heterologously expressed in yeast after an intensive directed evolution campaign.[36] Our group, therefore, investigated the construction of long-type chimeric enzymes. Three long-type UPO genes were selected for a simple gene, predefined shuffling approach and divided into 5 subdomains (243 possible combinations). In this setup, subsequent screening using a multi-injection GC-MS technique with tetralin as a substrate led to the identification of six functional UPO chimeras. Some of these enzymes exhibited elevated secretion levels when compared to the parental UPOs. This initial work suggests that gene shuffling of long-type UPOs can offer a viable option to obtain a library of active and structurally diverse long-type UPOs expanding the rather small panel of available recombinant peroxygenases.[70] Combining the multi-injection GC-MS technique with the simultaneous analysis of three substrates and six products enabled the identification of significantly modified chemo- and regioselectivities. *Mth*UPO variants with selective alkene epoxidation and shifted regioselectivities of the octane (**47**) hydroxylation toward 2-(**49**) and even 1-octanol (**48**) formation were identified (Scheme 5G).[71]

Scheme 5 Example C(sp3) Hydroxylations Catalyzed by Unspecific Peroxygenases (UPOs)



Since the publication of the 2021 review, research on unspecific peroxygenases (UPOs) has continued to expand rapidly, with the enzymes often lauded as "the pot of gold at the end of the oxyfunctionalization rainbow"[72] or "the jewel in the crown of C–H oxyfunctionalization biocatalysts".[73] Numerous recent studies have further demonstrated their efficiency and versatility in C(sp^3)–H hydroxylation. Examples include the selective hydroxylation of terpenes such as menthol at the C8 position to yield *p*-menthane-3,8-diol,[74] the hydroxylation of fatty acids at C4 and C5 followed by spontaneous lactonization to form valuable γ- and δ-lactones,[75] the oxyfunctionalization of cyclic ethers to produce a wide range of chiral cyclic hemiacetals with >99% enantiomeric excess and turnover numbers up to 95,170 under neat reaction conditions,[76] and the selective remote-site functionalization of halogenated and unsaturated hydroxyarbons.[77]

But UPOs are not restricted to C(sp^3) hydroxylations, they can catalyse a diverse portfolio of oxyfunctionalisation reactions. These include epoxidation of C=C double bonds, aromatic C(sp^2)–H hydroxylations, chemoenzymatic halogenations, and heteroatom oxidations (*N*- or *S*-oxygenations).[19] In one example, a UPO from *Aspergillus brasiliensis* (*Abr*UPO) catalysed the hydroxylation of substituted benzenes, the preference for aromatic versus benzylic hydroxylation proofed to be determined by the number, chemical properties, and chain length of the substituents. Oxidation of ethylbenzene resulted in an approximately 1:1 ratio of ring and side-chain hydroxylation, increasing the alkyl chain length shifted the selectivity toward the benzylic oxidation. Substituted phenols, in contrast, were oxidised exclusively at the aromatic ring. The resulting phenolic products are valuable intermediates for the synthesis of dyes, pharmaceuticals, and agrochemicals.[78] Another study showed, that the ratio of aromatic vs. benzylic hydroxylation can be shifted through mutagenesis.[52]

The heteroatom oxidation capabilities of UPOs have been harnessed for the synthesis of azoxy compounds, which exhibit diverse biological activities including antimicrobial and cytotoxic effects.[79] *Aae*UPO was shown to catalyse the transformation of aniline derivatives into azoxy products with turnover numbers up to 48,450, turnover frequencies of 6.7 s^{-1}, and up to 99% conversion at 98% chemoselectivity.[80]

Another heteroatom oxyfunctionalization catalyzed by UPOs is the asymmetric sulfoxidation, which was achieved with immobilized *Aae*UPO in non-aqueous media. A large variety of sulfides was converted to chiral sulfoxides with up to 99% ee, gram-scale synthesis of the products has been demonstrated.[81]

The examples discussed here and previously in the review paper illustrate that unspecific peroxygenases constitute a uniquely versatile class of enzymes, capable of catalyzing an exceptionally broad range of oxyfunctionalization reactions. Relying solely on hydrogen peroxide as oxidant, UPOs operate without the need for additional cofactors or redox partners, offering operational simplicity and environmental benefits. The breadth of available UPO sequences, combined with recent advances in heterologous expression in *E. coli* and yeast, and the first successful enzyme engineering campaigns, highlights their potential as adaptable and efficient biocatalysts. Taken

together, these attributes position UPOs as promising tools for both fundamental biocatalysis research and the sustainable synthesis of valuable chemical entities.

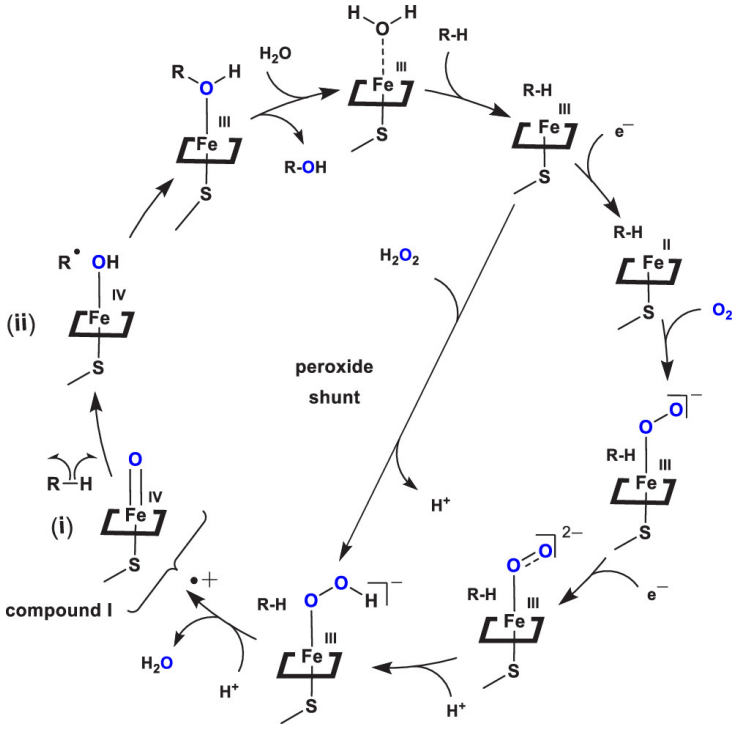
1.2.2 Cytochrome P450 Monooxygenases

To provide context and to highlight both the similarities and distinctions between these related enzyme classes, the following section introduces cytochrome P450 monooxygenases, drawing on the description presented in the same review article.

The following section is taken from my review paper published in ACS Catalysis in 2021.[1]

> Cytochromes P450 (CYPs or P450s) were first described in 1958. Because of their unknown enzymatic function, they have been named after the unusual Soret peak of the carbon monoxide complex at 450nm.[82] They catalyze a large variety of oxyfunctionalization, which all rely on molecular oxygen transfer.[7] P450s are distributed in all kingdoms of life. In mammals, they are primarily known to oxidize steroids, fatty acids, xenobiotics and are a crucial part of hormone synthesis and breakdown.[83] In plants, they represent approximately 1% of the plant genome and exhibit leading roles in the biosynthesis of defensive compounds and fatty acids, hormones, lignins, and secondary metabolites.[84] Archaea and eubacteria often do not possess P450 genes; neither does the prokaryotic modelorganism *E. coli*. In fungi, they seem to play a manifold and crucial role in the metabolism—especially regarding the adaptation of fungi to specific ecological niches.[85] More than 300,000 P450-coding sequences are known today considering all kingdoms.[86]
>
> P450s are heme-thiolate enzymes containing a heme prosthetic group linked to the protein via an axial conserved cysteine residue.[87] They utilize molecular dioxygen and use the universal cellular cofactors NADH or NADPH as reducing agents.[88,89] Respective P450 enzymes utilize different electron delivering systems providing the necessary electrons for oxygen cleavage and substrate hydroxylation. In the case of some bacterial and mitochondrial P450s, electrons are transported from NAD(P)H via flavin-dependent ferredoxin reductases and ferredoxin similars to the class of Rieske-type oxygenases.[90] Other bacterial and microsomal P450s only utilize cytochrome P450 reductases (CPR) or CPR combined with Cytochrome b_5—an electron transport hemeprotein.[91]
>
> The reaction mechanism requires two hydrogens (2 H+ , 2e-) for the formation of the side-product water. The two electrons are derived from the cofactor NAD(P)H. They are transferred in two consecutive one-electron transfer steps from NAD(P)H.[88,89] The Fe resting state of P450s is Fe(III). In the absence of a substrate, a water molecule serves as the distal ligand.[92] After substrate binding, the ligand exchange from water to substrate leads to a spin shift of Fe(III) from a low-spin to high-spin state.[93] This causes a lowered redox potential. This lowered redox potential allows the transfer of the first electron resulting in the Fe(II) formation. Fe(II) binds O_2 on the distal site, and after the second electron transfer, a nucleophilic double negatively charged Fe(III)-peroxo intermediate is formed.[94] After protonation, the ferric hydroperoxo intermediate (also called compound 0) is generated,[95,96] which is also an intermediate in the catalytic mechanism of UPOs[7,97] and peroxidases.[95,96,98] Further steps of the mechanism occur as described for UPOs. Water release leads to compound I formation that abstracts a hydrogen atom through homolytic bond cleavage from the substrate yielding compound II.[99,100] The substrate-based radical generated from C-H abstraction recombines, and the alcohol product is formed and released (Scheme 6).[7,101] P450s can also utilize H_2O_2 instead of molecular oxygen via the peroxide shunt; however, in most cases with substantially reduced activities relative to the NAD(P)H-driven reactions.[102]
>
> Because of their omnipresence and immense catalytic potential, P450s can function as alternative catalysts leading to new synthetic strategies for assessing a large variety of organic molecules of outstanding pharmaceutical and agricultural interest.

Scheme 6 Catalytic Cycle of Cytochrome P450 Monooxygenases



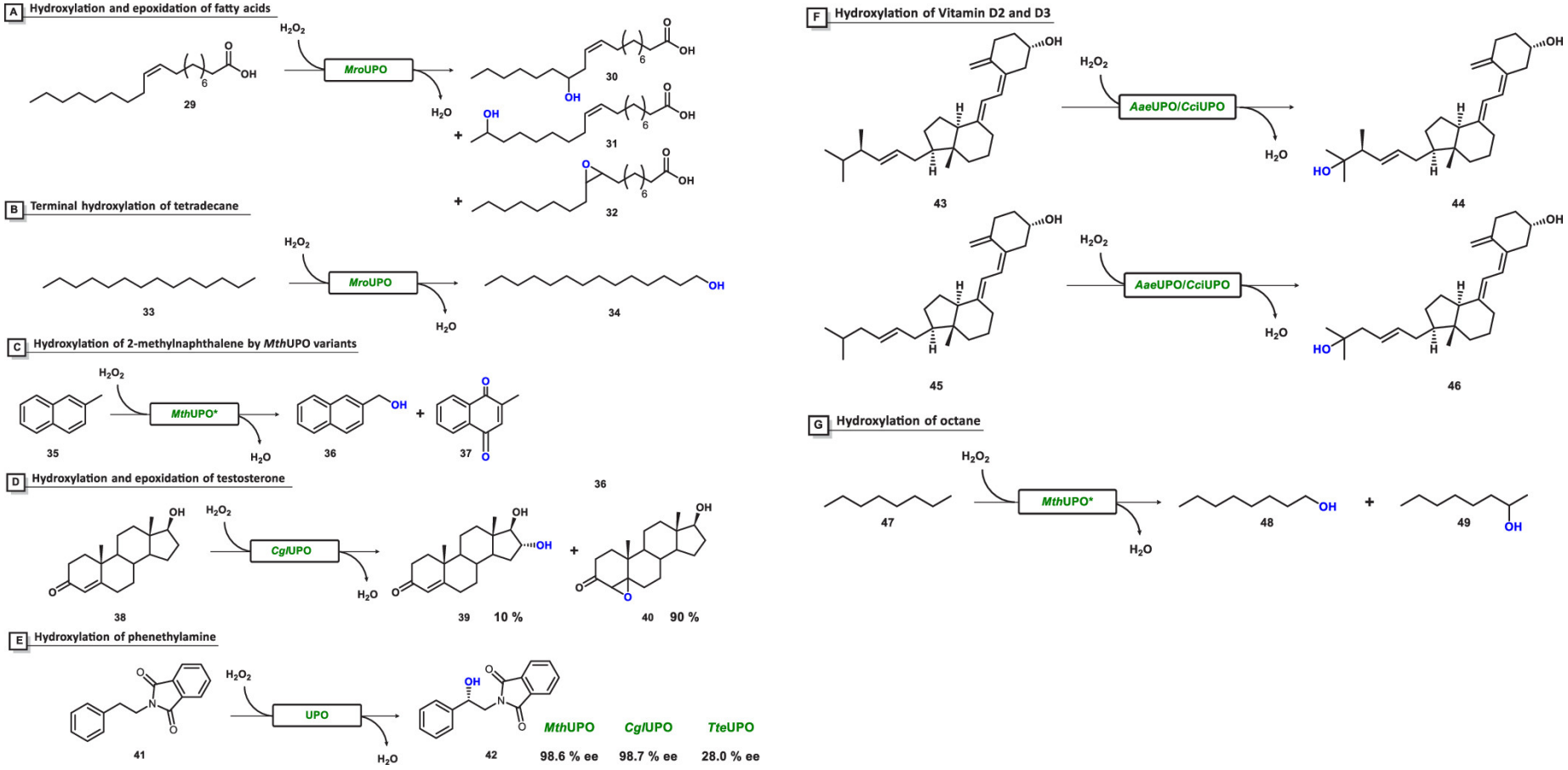
They have been extensively studied resulting in several thousands of publications. We can, therefore, only cover few examples focusing on protein engineering and industrial applications perspectives. For more detailed accounts, the readers are referred to comprehensive P450 reviews.[86,103–106] For overviews on the 2012 discovered P450 catalyzed carbene- and nitrene-transfer reactions readersare referred to theses overviews.[107–109]

P450s catalyze all types of C(sp^3) hydroxylation reactions ranging from activated to primary carbons. Even marginal activities were observed for methane hydroxylation.[110] Unfortunately, many P450s—especially if plant-derived—suffer from shortfalls regarding heterologous expression rates, activity and stability.[111] To address stability issues, P450s have been extensively engineered toward temperature resistance and solvent stability.[112,113] Besides limited stability, low activities can be a severe problem impeding the industrial application of P450s.[111,114] The main limiting parameter when employing cell-free P450 enzyme biotransformations is the availability of cofactors.[115] A multitude of solutions have been investigated to address this issue. One possibility is the targeted engineering of P450s for the utilization of H_2O_2 instead of dioxygen thereby avoiding an additional reducing agent. Hydroxylation of short-chain alkanes using H_2O_2 was shown to be possible. It achieved turnover numbers of up to 1800.[116] Another approach is the coexpression of glucose dehydrogenase within a wholecell setup to facilitate NADPH-regeneration, resulting in an approximately 22-fold increase in product concentration.[117] Switching the cofactor from NADPH to NADH or other cofactors, such as *N*-benzyl-1,4-dihydronicotinamide, might be beneficial, if applicable, as they are more stable and can reduce process costs.[118–120] The immobilization of P450s on electrodes was performed to investigate and further improve the electrochemical performance by fine-tuning their catalytic parameters and substrate recognition leading to possible applications as amperometric sensors.[121,122] Another possibility, which has been investigated in detail, is the replacement of NADPH by a light-generated source of electrons. This substitution is possible as the ferredoxin-mediated electrontransfer in mitochondrial and some bacterial P450 systems is similar to the electron carriers in the photosynthetic chain. In vivo and in vitro studies demonstrated the possibility of light-driven biosynthesis.[123,124] Further activity issues were addressed by using tailored fusion-proteins based on the CYP102A1 model en-

zyme from *Bacillus megaterium* ($P450_{BM3}$). Fusion-proteins can decrease both inefficient electron transfer between reductase domains and the heme domain, as well as uncoupling events.[125,126]

The monooxygenase families CYP4,[127] CYP152,[128] and CYP15351,[129] hydroxylate the terminal C-H bonds of alkanes and fatty acids **50**. These alcohols can further be oxidized by dehydrogenases leading to long-chain dicarboxylic acids (DCAs), a valuable chemical intermediate for polymers, glues, esters, and surfactants (Scheme 7A). DCAs rangingfrom C_9 to C_{12} are chemically produced in large quantities.[130,131] The demand of dodecanedioic acid DDDA is estimated to exceed 90.4 kilotons per month in 2023.[130,132] Diacid concentrations higher than 100 g L^{-1} at a > 100 m^3 fermentation scale could be enzymatically produced after enhancement of P450 expression with volumetric productivity higher than 1 g L^{-1} h^{-1}.[133]

Scheme 7 C(sp^3) Hydroxylation Reactions Catalyzed by P450s Discussed in This Review[†]



[†]The asterisk (*) indicates an engineered variant.

Another notable example for the industrial application of P450 enzymes is the hydroxylation of vitamin D3 (**45**) to 1α,25-dihydroxyvitamin D_3 (**53**) (Scheme 7B). Solely the double-hydroxylated form of vitamin D_3 exhibits physiological activity in humans and can treat hyperthyroidism, osteoporosis, and renal failure. This reaction is catalyzed by P450s from the CYP105 and CYP107 family (derived from *Pseudonocadria autotrophica*).[134] Enzyme engineering through rational mutagenesis design led to an enzyme variant with a 400-fold increase in catalytic efficiency (Table 2).[135,136] The biocatalytic production of hydroxylated vitamin D_3 by P450 enzymes is a more sustainable alternative for a previous inefficient 20-stage process.[137]

Table 2 Enzyme Engineering of CYP105A1 Leading to Improved Hydroxylation of Vitamin D3

CYP105A1	product	K_m (M)	k_{cat} (s^{-1})	K_m/k_{cat}(M^{-1} s^{-1})	relative activity
WT	1α(OH)	1.0×10^{-5}	1.3×10^{-4}	1.3×10^{1}	1
	25α(OH)	4.4×10^{-6}	4.3×10^{-5}	9.8	1
R73V	1α(OH)	1.8×10^{-5}	6.2×10^{-3}	3.5×10^{2}	28
	25α(OH)	2.7×10^{-6}	4.2×10^{-4}	1.3×10^{2}	15
R84A	1α(OH)	8.7×10^{-8}	4.2×10^{-3}	4.8×10^{2}	39
	25α(OH)	2.4×10^{-6}	4.2×10^{-4}	1.7×10^{2}	17
R73V/R84A	1α(OH)	6.5×10^{-6}	$3.5 \times 10^{-}$	5.4×10^{3}	432
	25α(OH)	2.2×10^{-6}	2.3×10^{-3}	1.0×10^{3}	105

Adapted with permission from ref.[135] Copyright 2008, American Chemical Society

The engineering possibilities of CYP101A1 ($P450_{cam}$) from *Pseudomonas putida* and CYP102A1 ($P450_{BM3}$) have been thoroughly investigated for the hydroxylation of alkanes.[138,139] As natural function, $P450_{cam}$ catalyzes the hydroxylation of the terpenoid (+)-camphor (**54**) regio- and enantioselectivity leading to (+)-5-hydroxy-camphor (**55**) (Scheme 7C). One engineering approache aimed to reduce the active site volume through the introduction of bulky amino acid residues. These variants showed fast and efficient n-butane hydroxylation, yet n-propane oxidation was modest, and shorter alkanes could not be hydroxylated.[140] Other approaches focused on directed evolution utilizing both enzymes CYP101A1 and CYP102A1 and led to ethane and propane hydroxylating variants (Scheme 7C).[138,139,141,142] The activity and product formation rate of the propane hydroxylation even exceeded those obtained with naturally occurring alkane monooxygenases on their native substrate.[143]

Several protein engineering studies of P450 enzymes catalyzing steroid hydroxylation were reported.[144–148] Reetz and co-workers engineered CYP102A1 ($P450_{BM3}$) toward higher regio- and stereoselectivities using directed evolution (Scheme 7D). They started with CYP102A1 (F87A), which hydroxylates testosterone (**38**) resulting in a 1:1 mixture of 2β-(**57**) and 15β-alcohol (**58**). Using iterative saturation mutagenesis (ISM) as engineering technique, they obtained mutants that exhibited selectivities of up to 97% for either of the two regioisomers without any loss of diastereoselectivity (β-selectivity). Moreover, mutants with more than 100-fold increased k_{cat}/K_m were identified.[149] Further engineering studies based on mutability landscapes and molecular dynamics simulations were focused on testosterone's C16 hydroxylation. Three combinatorial libraries were created, each with 1024 possible variants. As only less than 1,000 samples per library were screened the library coverage was only around 53%. Mutants WIFI-WC and WWV-QRS displayed k_{cat}/K_m values of 4.1×10^7 and 2.4×10^7 M^{-1} s^{-1}, respectively, while having high selectivities (95% for 16α-hydroxy testosterone (**39**) and 92% for 16β-hydroxy testosterone (**59**)) (Table 3).[150] Those activities are close to CYP102A1 natural substrate conversion (fatty acids, 5-6 $\times$ 10^7 M^{-1} s^{-1})[151] and suggest their potential industrial application in steroid hydroxylations.[138,139]

Another important late-stage functionalization of a larger substrate catalyzed by P450s is the hydroxylation of amorphadiene to artemisinic alcohol. This key step for the efficient synthetic production of the antimalarial drug artemisinin (**60**) is catalyzed by CYP71AV1 (coexpressed with CPR1

Table 3 Enzyme Engineering of CYP102A1 Leading to Improved Refined Regio- and Stereoselectivity of Testosterone[a]

mutant	selectivity	K_m (M)	k_{cat} (s^{-1})	K_m/k_{cat}(M^{-1} s^{-1})	regioselectivity
F87A	2β,15β	9.8×10^{-5}	3.1×10^{-1}	3.2×10^{3}	52% / 45%
KSA-8	2β	6.5×10^{-5}	1.4	2.1×10^{4}	84%
KSA-14	15β	1.6×10^{-5}	6.6	4.2×10^{5}	96%

[a] Kinetik parameters were otained measuring NADPH consumption.
Data derived from Kille et al.[149]

and CYB5).[152] Furthermore, engineered P450s were also able to hydroxylate remote, unactivated C(sp^3) of the final product artemisinin with refined regio- and stereo-selectivity (Scheme 7E). This was one of the first examples for fine-tuning the regio- and stereoselectivity of P450s for late-stage functionalization of complex molecules through enzyme engineering. As a mutagenesis method, a three-tier strategy involving first-sphere active site mutagenesis, high-throughput P450 fingerprinting and fingerprint-driven P450 reactivity predictions were used. Those enabled a rapid evolution of three efficient CYP102A1 enzymes hydroxylating a primary C(sp^3) leading to **63** and a secondary C(sp^3) with both *S* (**61**)and *R* (**62**) stereoselectivity (Table 4).[153]

Table 4 Enzyme Engineering of CYP102A1 for Regio- and Stereoselective Hydroxylation of Remote and Unactivated C(sp^3) in Artemisinin

	product distribution (%)		
CYP102A1 variant	**61**	**62**	**63**
parental (FL#62)	83	10	7
IV-H4	100	0	0
II-H10	0	100	0
X-E12	4	2	94

Data derived from Zhang et al.[153]

In 2018, the group of Bornscheuer initially reported on P450s of marine bacteria catalyzing the oxidative demethylation of 6-*O*-methyl-D-galactose (**64**), which demonstrates the ability of P450s to act on carbohydrate substrates (Scheme 7F).[154] Sugar-*O*-methylations shield polysaccharides of algae toward microbial hydrolytic enzymes,[155] removing the *O*-methylation modification gives access to glycoside binding hydrolases and thus the biotechnological usage of those polysaccharides for microbial fermentation, e.g., to produce biofuels.[156] Next to LPMOs these novel P450s are the second group of monooxygenases involved in carbohydrate degradation. Successful demethylation of further small substrates has already been demonstrated in 2006 with CYP116B2 (P450 *Rhf*).[157] Moreover, demethylation of monosaccharides using CYP102A1 was already demonstrated by the group of Arnold in 2009.[158]

Through the astonishing abundance of P450s in the kingdom of life and their vast engineering potential, investigating this enzyme family thrived for many decades and will continue in the future. Occurring bottlenecks such a slow production rates, low activities, dependence on expensive cofactors/redox protein complexes and enzyme stability issues will be the future challenges that need to be addressed to take this enzyme family beyond its current applicability.

1.3 Enzyme Engineering

With their high specificity, enzymes are increasingly utilized as catalysts within industry context. Limitations in activity, stability or selectivity of native enzymes, however, often impede their usage at industrial level. In response, enzyme engineering emerges as a pivotal methodology for tailoring enzyme properties to align with the requirements of specific applications. Directed evolution, semi-rational design and rational design are the most commonly used techniques and strategies to evolve an enzyme properties. Advancements in computational engineering introduce novel tools integrating machine learning and artificial intelligence (AI), set to revolutionize the landscape of enzyme engineering methodologies.

1.3.1 Discovery and Design as Starting Points for Enzyme Engineering

A fundamental prerequisite for any enzyme engineering endeavor is the availability of an expressible enzyme scaffold that exhibits at least minimal catalytic activity toward the substrate of interest. This initial activity does not have to be the enzyme's native function; often, a promiscuous or side activity provides a viable starting point for engineering efforts. While the discovery of such candidate enzymes—whether through metagenomic mining, in-silico bioprospecting, activity-based screening, or structure-guided annotation—is a critical step, an in-depth discussion of enzyme discovery and expression methodologies is beyond the scope of this work and not critical for the understanding of the following studies. Interested readers are referred to recent reviews.[159–163]

An alternative approach to identifying starting scaffolds involves the de novo design of enzymes, wherein entirely new protein structures are constructed to catalyze specific reactions. Recent advancements in computational tools, including deep learning algorithms and improved understanding of sequence-structure-function relationships, have significantly enhanced the feasibility of de novo enzyme design. Given that the focus of this work is on engineering of existing enzymes guided by computational tools, a comprehensive exploration of de novo enzyme design is not included here. For readers interested in this field, the following reviews and studies provide detailed insights into current strategies and applications.[164–169]

1.3.2 Laboratory Techniques and Strategies

There are three main strategies to evolve an enzyme: directed evolution, rational design and semi-rational-design, which is a combination of the first two methods. The choice of method depends on the engineering goal and knowledge about the enzyme, its structure and properties as well as the viability of the screening system. The better the understanding about the sequence-function relationship within the enzyme, the more rational decision making can be employed within the engineering strategy.

Directed Evolution and Semi-Rational Design

Darwinian Evolution, the concept of evolving through variation, selection and inheritance is the insipration for directed evolution. Directed evolutions mimics this natural process and is conducted across multiple iterations wherein the enzyme undergoes incremental refinement until the desired property is attained. First, a diverse genetic library is generated and subsequently screened for the desired property. The enzyme variant with the most significant improvement in the desired property (hit) is then used as new parental variant in a following round (**Fig. 1**). The introduction of mutations can occur hereby randomly or targeted at specific sites. Techniques for genetic diversification include random mutagenesis, error prone PCR (*ep*PCR), DNA shuffling or staggered extension processes. The foremost advantage of directed evolution using one of those techniques is that there is no need for prior knowledge of the protein structure or its function. The biggest drawback, however, is the large number of possible mutations, that can occur within a single enzyme, also referred to as *the numbers problem*.[170] For instance, fully saturating all amino acid residues in a 300-amino-acid-long enzyme with the 20 canonical amino acids would yield 20^{300} enzyme variants, surpassing the estimated quantity of stars in the universe. Consequently, even if a library containing such a multitude of variants were constructed, only a fraction of the genetic diversity could be feasibly screened. Therefore, strategies to reduce the number of variants are employed, but in any case a robust high-throughput-screening (HTS) is needed to screen through a large number of possible variants.

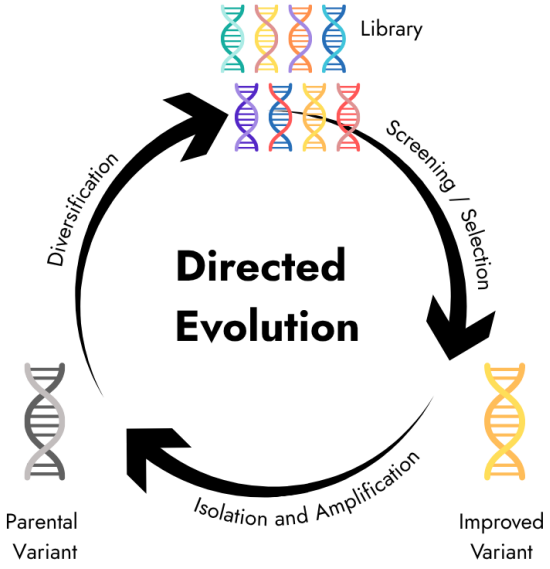


Figure 1 Principle of directed evolution

Flow cytometry and chip-based microfluidic assays are screening methods allowing for a screening of > 10^6 h^{-1} variants.[171,172] Microtiter plates (MTPs) are widespread screening formats and allow analysis of up to 10^4 variants per day, agar plate based assays are suitable for library

sizes of up to 10^5.[173] Another method combining mutagenesis and screening is in vivo continuous evolution, consisting of continuous cyclces of gene diversification and selection within a cell without intervention. By increasing the selection pressure while accumulating mutations in the gene on interest, desired properties can be evolved throughout several generations.[174,175]

In a campaign to improve the activity of an UPO variant derived from *Agrocybe aegerita* in organic solvents, two gerenerations of random mutagenesis (*ep*PCR) and in vivo DNA shuffling were performed, followed by one round of in vivo site-directed recombination. The final variant of the campaign carried nine mutations and displayed a 23-fold enhanced activity in 30% acetonitrile (vol/vol) and activity in acqueous acetone, methanol and dimethyl sulfoxide mixtures compared to the starting enzyme. The screening for the desired enhanced activity was performed in MTPs, in each round around 3500 variants were screened.[176]

One way of reducing the number of variants that need to be screened is targeting specific areas and residues on the basis of structural or functional knowledge to create smaller, "smart" libraries that contain a higher probability for hits. Such areas are for example the active site for altering activity or selectivity or substrate entrance channels to allow or hinder certain molecules from entering the active site. This approach is often labeled *semi-rational design*.

It was utilized to evolve a P450 enzyme variant (P411) to perform enantio-, regio- and chemoselective intermolecular alkylation of sp^3 C-H bonds through carbene C-H insertion. The campaign started with a total turnover number (TTN) of 13 and consisted of sequential rounds of site-saturated mutagenesis of residues in the active site pocket and on loops and other flexible regions of the protein as well as residues possessing a nucleophilic side chain and the removal of an enzyme domain. After 15 rounds of directed evolution the best enzyme variant showed 140-fold activity improvement and excellent stereoselectivity (enantiomeric ratio (e.r.) 96.7:3.3).[177] This example shows that enzyme engineering can go beyond improving enzymes natural functions, directed evolution can also be used to create new, abiological functionalization unknown to nature.

Another approach to tackle the numbers problem is reducing the amino acid alphabet, that is available for introducing mutations. In another study displaying semi-rational design, a limonene epoxide hydrolase (LEH) from *Rhodococcus erythropolis* was utilized to desymmetrically hydrolyse cyclohexene oxide. Wild type LEH leads to 4% enantiomeric excess (ee) of (*S*,*S*)-cyclohexane-1,2-diol. Guided by structural analysis, 10 residues were chosen for diversification and it was noticed, that all of them have a hydrophobic character. To reduce the screening effort the residues were grouped into three groups and only three amino acids—valine, phenylalanine and tyrosine—were selected to saturate the residues in each group simultaneously, which reduced the screening effort to a few hundred samples. The campaign resulted after one round in a (*S*,*S*)-selective variant and after three rounds in a (*R*,*R*)-selective variant each displaying 97-99% ee.[178]

A critical aspect of designing semi-rational evolutionary experiments is the selection of residues to be targeted for mutation. The modification of multiple residues typically does not

result in a linear change in enzyme properties; instead, it often exhibits epistatic behavior, producing a complex fitness landscape.[179,180] Epistatic interactions play a major role in shaping protein structure and function, reflecting the non-additive effects of multiple mutations on key parameters such as stability, activity, and selectivity. This non-linearity presents a significant challenge for accurately predicting mutational outcomes and understanding structure-function relationships. However, epistasis also offers a powerful opportunity to induce substantial improvements in desired properties by mutating amino acids that are spatially proximal within the protein.[181]

A semi-rational engineering technique exploiting epistatic effects is the combinatorial active-site saturation test (CAST) which is often combined with iterative saturation mutagenesis (ISM). For CASTing pairs of spatially close amino acid residues are chosen, whose side chains are oriented toward the binding site of the WT enzyme and pairwise randomized. Following, the best hit of this first round of screening is used as template for the following round, leading to an iterative improvement of enzyme properties.[182,183] The three methods CASTing, ISM and a reduced amino acid alphabet were combined to improve the (*S*)-enantioselectivity of a liase from *Pseudomonas aeruginosa* (PAL) used for hydrolytic kinetic resolution of a racemic ester. In a previous study a combination of epPCR, saturation mutagenesis and DNA shuffling was used to evolve PAL, after screening 50,000 variants the E-factor of the enzyme could be improved from $E = 1.2$ to $E = 51$ while introducing six mutations. The new apporach combining CASTing, ISM and a reduced amino acid alphabet resulted in an enzyme variant with three mutations displaying $E = 594$ after screening 10,000 variants utilizing strong epistatic effects.[184]

The availability of enzymes engineered through directed evolution has opened new synthetic routes for the production of key pharmaceuticals and industrial chemicals, often with the added benefit of reduced environmental impact. For her pioneering contributions to the field of directed evolution Frances Arnold was arwarded the Nobel Prize in Chemistry in 2018.

Rational Design

Rational design is an engineering strategy that relies on detailed knowledge of an enzyme's structure and function. Typically, specific amino acid residues—or stretches of residues—are selected and targeted for mutation or deletion to achieve a desired functional outcome. The choice of these residues is guided by either sequence-based analyses, such as multiple sequence alignment with well-characterized homologous proteins, or structure-based approaches, traditionally grounded in experimental methods like X-ray crystallography.

The earliest rational design studies employed site-directed mutagenesis to investigate the contribution of individual residues to enzymatic activity and stability (i.e., structure-function relationships)[185,186]. Subsequent efforts focused on enhancing enzyme robustness by rigidifying flexible regions[187], introducing disulfide bridges[188,189], or engineering surface salt bridges[190].

Although early rational design efforts relied on relatively simple structure-function hypothe-

ses and manual residue selection, the approach has become increasingly sophisticated through the integration of computational methods. Today, rational design is closely intertwined with computational workflows that enhance both precision and efficiency. These include tools such as sequence alignment, homology modeling, and computational structure prediction—most notably via AlphaFold[191]—which now guide mutation strategies with unprecedented accuracy. This integration has also enabled the development of algorithms capable of predicting the functional consequences of engineered variants, effectively bridging traditional structure-based design with data-driven approaches. Given this convergence, many rational design strategies are best understood within the broader context of computational enzyme engineering. The following section explores these tools in detail, highlighting their applications in both classical rational design and machine learning-guided methodologies.

1.3.3 Computational Tools Transform Enzyme Engineering

The integration of computational methods into enzyme engineering has transformed both the scale and scope of what can be achieved through enzyme design. Fundamental tools such as multiple sequence alignment and structural modeling enable the identification of conserved residues, functional motifs, and flexible regions. These analyses are further enhanced by molecular dynamic (MD) simulations and structure prediction algorithms like AlphaFold, which provide atomic-level insights even in the absence of experimental structures. More recently, machine learning and AI-driven models have opened new possibilities for predicting beneficial mutations, modeling protein fitness landscapes, and designing enzymes with optimized properties—all while minimizing experimental workload. This section outlines the core computational strategies that are shaping the future of enzyme engineering.

Sequence-Based Enzyme Engineering

Sequence-based computational design allows the enzyme design or engineering from protein sequences, preventing the need for structural information. Instead insights from billions of years of natural enzyme evolution are utilized. The development of next-generation sequencing techniques lead to a wealth of sequence data, which can be employed for phylogenetic analysis.[192] Two of the most widespread sequence-based compuational methods are Consensus Design (CD) and Ancestral Sequence Reconstruction (ASR). CD is based on Multiple Sequence Allignment (MSA) of hundreds of homologous protein sequences. Within the MSA the distribution frequency of amino acids is calculated to find "consensus" amino acids. Common amino acids at specific positions are assumed to be more stable and evolutionary conserved, leading to potential effects on enzyme stability, activity and selectivity.[193–195]

The concept of ASR is based on the assumption, that ancestral enzymes were less specific and rather generalists, capable of catalyzing broader reaction scopes and in return exhibited higher

thermostability.[196] Based on the phylogentetic relationship between modern homologs ancestral sequences are reconstructed after applying a statistical model calculating sequenes at internal nodes of the phylogenetic tree.[197]

ASR was for example utilized to resurrect ancestral P450 enzymes from the pre-Cambrian era, which displayed significantly improved thermostability—they were able to withstand $\sim$ 30 °C higher temperatures and $\geq$100 times longer incubation times than their extant forms. When comparing the anccestral CYP3_N1, which is predicted to have existed in the vertebrates $\geq$450 million years ago, with the well-characterized human CYP3 A4, the ancestor showed not only improved thermostability, but maintained comparable activity toward known CYP3 substrates proving that thermostable proteins can be devised using sequence data alone.[198]

Structure-Based Enzyme Engineering

Structure-based enzyme engineering relies on identifying key residues within the substrate-binding pocket or access channels that are essential for substrate recognition, orientation, and catalysis. Traditionally, enzyme structures have been determined using experimental methods such as X-ray crystallography and cryo-electron microscopy, both of which are time-consuming and resource-intensive. Prior to recent breakthroughs in structure prediction, homology modeling was the dominant approach for obtaining structural models when no experimental structure was available.[199,200]

The major bottleneck of predicting protein structures directly from amino acid sequences has now been substantially overcome by deep learning-based approaches. Algorithms such as AlphaFold[191] and RoseTTAFold[201] leverage evolutionary and structural data at large scale to generate highly accurate structural predictions, dramatically expanding the availability of enzyme models for structure-based design.

At the core of structure-based engineering is the understanding of substrate positioning and its interactions within the active site. Static approaches such as molecular docking[202,203] model how a substrate fits into the binding pocket, while dynamic simulations like molecular dynamics (MD)[204,205] capture enzyme flexibility, side-chain rearrangements, and time-resolved atomic motions. Catalytic mechanisms can be investigated using quantum mechanics/molecular mechanics (QM/MM) simulations,[206] providing insight into the transition states and reaction energetics. Additionally, free energy calculations[207,208] allow quantification of binding affinities and identification of residues with the highest contributions to substrate stabilization—key targets for mutagenesis.

Beyond analysis, computational tools such as Rosetta and RosettaDesign[209–211] use physical energy functions to design new or altered enzyme structures, allowing remodeling of loops, binding pockets, or entire active sites, up to the level of *de novo* enzyme design.

These computational techniques aim to predict mutations that alter the local chemical microenvironment—hydrogen bond networks, steric clashes, and electrostatic interactions—

thereby modulating enzymatic performance. After identifying residues of interest, in silico mutagenesis and ranking based on predicted binding energies or stability enables virtual screening, dramatically reducing experimental load by narrowing down to a small set of promising variants.

For instance, in a study engineering phosphoserine aminotransferase (SerC) to catalyze the deamination of L-homoserine, MD simulations and binding free energy decomposition were used to identify critical binding-site residues. Subsequent in silico site-saturation mutagenesis and virtual screening identified variants with significantly improved activity toward the new substrate and abolished activity toward the native one. The best mutant exhibited a 4.2-fold increase in catalytic efficiency compared to wild type.[212]

In another example, an adapted CASCO (Catalytic Selectivity by Computational Design)[213] workflow was applied to invert the stereoselectivity of a thermostable alcohol dehydrogenase. Using MD simulations, Rosetta design calculations, and Boltzmann-weighted binding energy predictions, four variants with 6-8 active-site mutations were identified. Three variants successfully inverted the enantioselectivity of acetophenone reduction from 99% (*S*)-1-phenylethanol to 99% (*R*)-1-phenylethanol, while maintaining catalytic activity and broad substrate scope.[214]

Modern virtual screening strategies often draw on lessons from directed evolution. Frameworks such as FRISM (Focused Rational Iterative Site-Specific Mutagenesis),[215] CADEE (Computer-Aided Directed Evolution of Enzymes),[216] and EnzyHTP[217] automate mutant generation and screening based on MD simulations, QM/MM analysis, and energy-based scoring, substantially reducing the number of variants that must be tested in the lab.

The increasing availability of accessible software tools and web-based platforms now enables the use of sophisticated computational design methods by researchers without extensive training in computational chemistry, further integrating structure-based approaches into enzyme engineering workflows.

A study aiming to promote the peroxygenase activity of P450 enzymes by engineering their water tunnels to facilitate H_2O_2 access to the heme center within the active site utilized the CAVER Web 1.0 web server.[218] This online plattform is an interactive tool for analysing protein tunnels and channels built on top of the widely used tunnel detection tools Caver 3.02[219] and CaverDock 1.0. Only needing a crystal structure as input, the tool identified key residues at the bottle neck and entrance of different the water tunnels, which were subsequently mutated to small or polar hydrophilic side-chains. The approach was tested with several P450s, especially the H_2O_2-driven activity of two native NADH-dependent P450s (CYP199A4 and CYP153A$_{M.aq}$) increased significantly (by >183-fold and >15-fold, respectively).[220]

Molecular modeling techniques such as MD simulations, free energy calculations, and QM/MM analysis play a dual role in structure-guided enzyme design: initially to pinpoint candidate residues for modification, and subsequently to rationalize the functional consequences of engineered variants. By revealing the molecular basis of altered catalytic properties, these

methods enhance our mechanistic understanding and provide a rational basis for further rounds of directed evolution or computational design.[221–223]

Combining Sequence- and Structure-Based Information

Some strategies and tools do not rely only on sequence- or structure-based information but utilize both in combination to achieve even better predictions for beneficial mutations. One of those tools is PROSS (Protein Repair One-Stop Shop), which tackles the challenge of marginal stability of many natural enzymes—a major bottleneck for the usage of proteins in basic and applied research.[224] The algorithm first performs a MSA from which it computes a position-specific substituion matrix (PSSM), representing the chance of observing any of the 20 amino acids at each position of the protein. Amino acid residues within the active site and especcially ligands for the catalytic reaction remain hereby undisturbed, non or rarely observed mutations are excluded. In a second step Rosetta calculations are applied to determine the energy difference of each single allowed mutation within the wild type structure and the wild type protein. This defined "space of potentially stabilizing mutations" is finally combined and Rosetta cominatorial sequence design calucations are used to find the potentially most stable designs for experimental testing.[225,226]

PROSS was used for example to find enzyme designs with 17-45 mutations improving the yield (4-fold) and thermostability (by 4 °C) of the metastable envelope glycoportein (Env) of HIV-1, which is hypothesized to improve induction of broadly neutralizing antibodies.[227] In another study human acetylcholinesterase (hAChE), an enzyme mediating synaptic transmission, was engineered with PROSS, leading to a variant with 51 mutations containing improved core packing, surface polarity and backbone rigidity, expressing in 2000-fold higher levels in *E. coli* compared to the wild type and exhibited 20 °C higher thermostability without change in enzymatic properties or active-site configuration.[226] The PROSS algorithm was utilized within this work to engineer novel peroxygenases for expression in fast-growing yeast, which proved challenging with the wild type enzymes (see Chapter III).

Another enzyme engineering web server based tool utilising both sequence- and structural data is FuncLib, which focuses on utilising epistatic effects of simultaneous mutations within the active-site while providing a small, diverse set of designs for experimental testing. The first step of the algorithm contains again a MSA and the generation of a PSSM, however focusing on selected residues within the active-site. From this filtered set, multipoint mutants are modeled and refined that comprise 3-5 mutations relative to wild type in Rosetta, including backbone and side-chain minimization, to enable radical mutations such as from small to large amino acid side chains. All multipoint mutants are ranked according to their predicted stability and finally designs are excluded that differe by fewer than two active-site mutations to ensure a high diversity.[228] FuncLib was successfully used within this work to create a diverse set of UPO designs for the oxyfunctionalization of different terpene substrates (see Chapter V).

2 Aim of the Thesis

The overarching goal of this thesis was to develop a systematic framework for the secretion and targeted optimization of unspecific peroxygenases (UPOs) as versatile biocatalysts for selective oxyfunctionalization reactions. UPOs are particularly attractive because they can catalyze diverse C–H oxyfunctionalizations using hydrogen peroxide as the sole cosubstrate. Despite this potential, the broader application of UPOs is limited by challenges such as difficult heterologous expression, restricted substrate scope, and limited control over activity and selectivity. Most substrates investigated in this work provide challenges with respect to regio- and stereoselectivity, including compounds of pharmaceutical relevance, highlighting both the limitations and the catalytic potential of this enzyme class.

This thesis seeked to address these limitations by exploring how modern computational and experimental approaches can be combined to support rational enzyme engineering. Recent advances in computational chemistry and structural biology, including improved accessibility through user-friendly tools, opened new opportunities for guiding enzyme design and reducing experimental effort, even for researchers without extensive computational expertise. The work investigated how these emerging methods can complement classical experimental strategies, enabling systematic optimization of UPOs for improved secretion, activity, selectivity, and multi-substrate performance.

This motivation led to the following guiding research questions:

- How can novel UPOs, including those initially not secretable, be reliably expressed and functionally screened?
- How do the results of minimal, algorithm-assisted design libraries compare to those of more extensive, structure-guided and MD-supported directed evolution campaigns, and what trade-offs exist between library size, experimental effort, and final enzyme performance?
- How can computationally assisted design strategies be applied to improve enzyme performance across multiple substrates, and what are the limits of multi-substrate engineering?
- How can available structural, functional, and expression data be used to decide on an effective strategy for enzyme engineering including computational and experimental methods?

The aim of this thesis was to provide a transferable methodological framework and a broadly applicable set of strategies for UPO engineering. By integrating computational and experimental approaches, the work seeks to expand the accessible UPO toolbox, demonstrate that key limitations can be addressed, and provide guidance for the design of future enzyme engineering campaigns, applicable for both UPOs and other monooxygenases. In this way, the work aimed to take steps toward bringing these versatile catalysts closer to practical use and to lay a platform for future exploration of their potential.

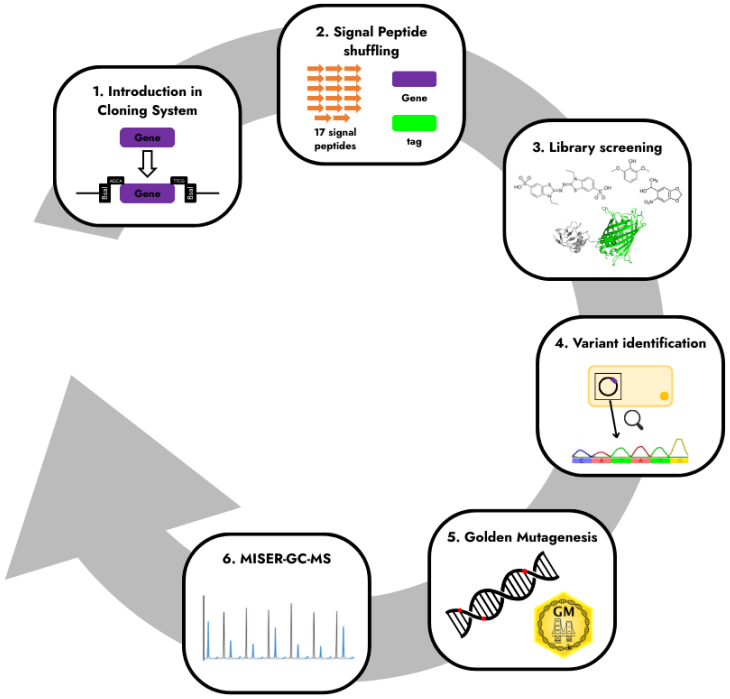
3 | Chapter I

This chapter has been published as:

Secretion and directed evolution of unspecific peroxygenases in *S. cerevisiae*

by: Niklas Dietz, Li Wan, **Judith Münch**, Martin J. Weissenborn

in: *Methods in Enzymology* **2023**, *693*, Elsevier; doi: https://doi.org/10.1016/bs.mie.2023.09.013



This chapter describes the methods and protocols employed for secretion and directed evolution of unspecific peroxygenases in *S. cerevisiae*. Most of the described methods are utilized in the following chapters, from library generation via signal-peptide shuffling and mutation, through yeast cultivation and enzyme secretion in batch and microtiter plate format to activity analysis via colorimetric assays and MISER-GC-MS. Especially the MISER-GC-MS protocols describe detailed the workflow for substrate-independent high-throughput screening for engineering an unspecific peroxygenase and were the basis for the directed evolution campaign described in Chapter IV.

CHAPTER TEN

Secretion and directed evolution of unspecific peroxygenases in *S. cerevisiae*

Niklas Dietz, Li Wan, Judith Münch, and Martin J. Weissenborn*
Institute of Chemistry, Martin-Luther-University Halle-Wittenberg Weinbergweg 22, Halle (Saale), Germany
*Corresponding author. e-mail address: martin.weissenborn@chemie.uni-halle.de

Contents

Methods in Enzymology, Volume 693
ISSN 0076-6879, https://doi.org/10.1016/bs.mie.2023.09.013

Abstract

Yeast-based secretion systems are advantageous for engineering highly interesting enzymes that are not or barely producible in *E. coli*. The herein-presented production setup facilitates high-throughput screening as no cell lysis is required. All techniques are described in detail, with access to freely available online tools and all vectors have been made available on the non-profit plasmid repository AddGene. We describe the method for UPOs as a model enzyme, showcasing their secretion, detection, and evolution using *S. cerevisiae*. Additional material to transfer this to *P. pastoris* has been published by our group previously (Püllmann & Weissenborn, 2021).

1. Introduction

Fungal unspecific peroxygenases (UPOs) have gained increasing attention from researchers for their ability to oxyfunctionalize a wide spectrum of substrates in the last decade (Beltran-Nogal et al., 2022; Hofrichter & Ullrich, 2014; Hofrichter, Kellner, Pecyna, & Ullrich, 2015; Münch, Püllmann, Zhang, & Weissenborn, 2021). They are particularly interesting because of their large abundance with more than 4800 annotated UPO sequences and their broad substrate scope with more than 400 determined substrates. (Beltran-Nogal et al., 2022; Kinner, Rosenthal, & Lutz, 2021) UPOs utilize hydrogen peroxide (H_2O_2) as both oxygen and electron source and are thus independent of cofactors like NAD(P)H or complex electron transfer chains.

The challenging production of UPOs has been the main bottleneck for further application for a long time. (Püllmann et al., 2021) Therefore, Alcalde et al. and our lab developed UPO production setups in *Saccharomyces cerevisiae* and *Pichia pastoris*. These hosts can perform necessary post-translational modifications like glycosylation and are able to secrete the proteins out of the cells thereby facilitating the downstream purification and enzyme engineering. (Molina-Espeja et al., 2014; Püllmann & Weissenborn, 2021; Püllmann et al., 2021) Although several methods exist to engineer *E. coli* to

gain protein secretory ability, (Kleiner-Grote, Risse, & Friehs, 2018) there is no readily accessible standard strain, and tedious optimization through genetic engineering is usually needed. On the other hand, *S. cerevisiae* with a natural post-translational modification system and an endogenous protein secretion system, has been used as the host for secretory protein production for decades. Many mature strains, tools, and methods have been developed to optimize the protein secretion in *S. cerevisiae* with different applications. (Tippelt & Nett, 2021) Moreover, *S. cerevisiae* is phylogenetically close to the original host of UPOs, which makes it easier to produce functional recombinant UPOs. Therefore, we chose *S. cerevisiae* as the host to produce and engineer UPOs.

In this chapter, we present the methods used in our laboratory to produce and engineer UPOs. All methods are easily applicable to any laboratory and the required tools and plasmids are readily available. The first thing we describe is the signal peptide shuffling, which enables the secretion of the target protein and has been proven in numerous projects to be pivotal for protein production. We discuss assay systems with the highlight of substrate-independent protein detection and quantification using splitGFP. With suitable secretion and activity assays enzymes on hand, we subsequently focus on preparing protein libraries. We have developed a rapid and highly reliable mutagenesis tool termed Golden Mutagenesis enabling the site-directed mutagenesis of several positions simultaneously within one day. Lastly, to enable the directed evolution to be applied in a substrate-independent manner, we have altered a standard GC–MS to enable the analysis of a 96-well microtiter plate within one hour.

We hope that this book chapter will provide the reader with all the required knowledge to perform the secretion and engineering of many exciting enzymes in *S. cerevisiae* (Fig. 1).

2. Generation of a signal peptide library

As UPOs are secreted enzymes, the signal peptide plays a crucial role in the production. The heterologous expression and secretion in *S. cerevisiae* of an unspecific peroxygenase from *Agrocybe aegerita* (*Aae*UPO) could be increased 1110-fold by introducing nine mutations (variant PADA-I). Four of these mutations are in the signal peptide. Just these four mutations result in a 27-fold increase in secretion levels (Molina-Espeja et al., 2014). Our group established a system, which also focuses on the signal peptide. Instead of

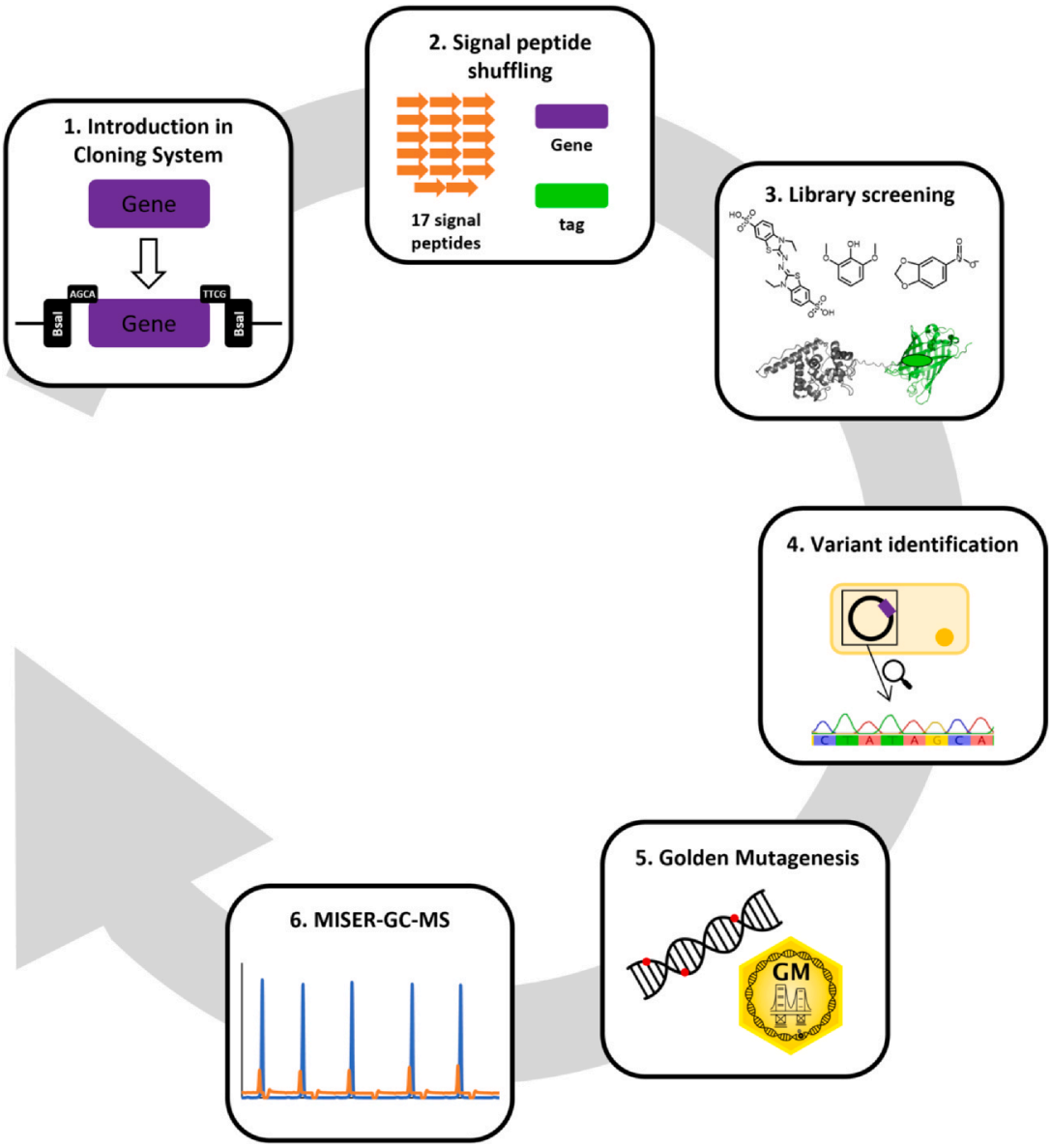


Fig. 1 Workflow for the secretion and directed evolution of UPOs.

introducing mutations in the signal peptide, the whole signal peptide is exchanged. With this approach, we could double the secretion level of *Aae*UPO compared to the evolved PaDa-I variant and enabled the production of three novel UPOs, which were not expressed before (Püllmann et al., 2021). The whole system relies on a Golden Gate Cloning system, which enables the efficient assembly of variants with different signal peptides.

We here describe the protocol for rapidly generating a signal peptide library with a set of 17 signal peptides for the expression and secretion of enzymes in *S. cerevisiae* using Golden Gate Cloning. All plasmids including the signal peptides are available at AddGene [Yeast Secrete and Detect (Kit #1000000166)].

2.1 Equipment

- PCR Thermocycler machine (Labcycler basic, SensoQuest)
- PCR tubes (Sarstedt)
- Thermomixer (Thermomixer comfort, Eppendorf)
- Spectrophotometer (BioSpectrometer, Eppendorf)

2.2 Buffer, strain, plasmids and reagents

- gene of signal peptides in pAGM9121 (Addgene #153491–#153507)
- pAGM9121_TwinStrep-GFP11 (Addgene #153514) (C-terminal tag)
- pAGM9121 (Addgene #51833) (backbone vector Level 0)
- pAGT572_Nemo2.0 (backbone vector Level 1)
- UPO gene of interest
- *Bbs*I (BpiI) (5 U/μL, Thermo Fisher Scientific)
- T4 DNA Ligase (1–3 U/μL, Promega—*it is crucial to use exactly this ligase*)
- 10× T4 Ligase Buffer (Promega)
- *Bsa*I (10 U/μL, New England Biolabs)
- NucleoSpin Plasmid Kit (Machery-Nagel)
- *Escherichia coli* DH10B chemically competent cells (Thermo Fisher Scientific), genotype: *F– mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ (ara-leu)7697 galU galK λ– rpsL(StrR) nupG*
- Sterile SOC medium [2% (w/v) tryptone, 0.5% (w/v) yeast extract, 10 mM magnesium chloride, 10 mM sodium chloride, 2.5 mM potassium chloride, 20 mM glucose, pH 7.0]
- LB-Agar plates [1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) sodium chloride, 1.5% (w/v) agar, pH 7.0] supplemented with 100 μg/mL spectinomycin, 50 μg/mL X-Gal, and 150 μM IPTG
- LB-Agar plates [1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) sodium chloride, 1.5% (w/v) agar, pH 7.0] supplemented with 100 μg/mL carbenicillin, 50 μg/mL X-Gal, and 150 μM IPTG
- Sterile LB medium [1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) sodium chloride, pH 7.0]

2.3 Procedure

2.3.1 Preparing your gene of interest for implementation into the modular cloning system

Timing: Planning

1. To implement the UPO gene of interest in the modular cloning system, the DNA has to be modified. For that, there are two possibilities:

a. Ordering the gene as dsDNA flanked by the compatible 4 bp-overhangs and restriction enzyme recognition sites (Fig. 2), or
b. Amplification of the gene via PCR from a suitable template. In that case, primers need to be designed according to Fig. 2.

In both cases, make sure that the DNA is free of any further BsbI or *Bsa*I recognition sites. You can use the "domestication function" of our online tool https://msbi.ipb-halle.de/GoldenMutagenesisWeb/ to design primers for their one-step removal.

2.3.2 Preparing Level 0 parts

Timing: Day 1–3

1. Set up the Golden Gate Cloning reaction as follows (Table 1).
   To calculate the necessary amount of plasmid or gene, use this formula:

$$\mu l\ to\ pipet = \frac{desired\ amount\ of\ plasmid\ (in\ fmol) \times size\ of\ the\ plasmid\ (in\ bp)}{1520 \times concentration\ of\ your\ plasmid\ (in\ ng \times \mu l^{-1})}$$

Fig. 2 Sequence of primers/flanked sites to introduce a gene in the modular cloning system.

Table 1 Golden Gate master mix for preparing Level 0 plasmids.

Components	Amount
10× T4 Ligase Buffer	1.5 μL
pAGM9121(acceptor plasmid)	10 fmol
Gene of interest (dsDNA or fragment)	50 fmol
BsbI	0.5 μL
T4 DNA Ligase	1.0 μL
ddH_2O	Add to 15 μL

2. Take the PCR tubes to the thermocycler and start the following Golden Gate Cloning program (Table 2).
3. After finishing the reaction, thaw an aliquot (50 μL) of chemically competent *E. coli* DH10B cells on ice for 5 min.
4. Mix the cells with the whole Golden Gate reaction mixture and incubate them for 15 min on ice. In between, set a thermomixer to 42 °C and pre-warm 250 μL SOC medium.
5. Perform the heat shock for 90 s at 42 °C. Afterward, place the cells back on ice for 3 min.
6. Add 250 μL pre-warmed SOC medium and incubate the cells at 37 °C for 1 h under shaking in a thermomixer device.
7. After incubation, plate 100 μL cell suspension on an LB agar plate (+spectinomycin, +X-Gal, +IPTG). Incubate the cells at 37 °C overnight.
8. Check for positively assembled plasmids by blue-white selection. White colonies: the assembly was performed correctly; blue colonies: no correct insertion of the target gene.
9. A single, clearly separated, white colony is used to inoculate 5 mL LB medium (+100 μg/mL spectinomycin). The culture is grown overnight at 37 °C under shaking.
10. Isolate the plasmid using the NucleoSpin Plasmid Kit according to the manufacturer's instructions. Determine the plasmid concentration using a Nanodrop spectrophotometer.
11. Sequence the plasmid to ensure the right nucleotide sequence of the gene.

2.3.3 Creation of signal peptide library

Timing: Day 3–5.

12. After confirmation of the right sequence, the signal peptide library can be prepared. For that set up a Golden Gate reaction as follows Table 3.

Table 2 Golden Gate cloning program for digestion and ligation of Level 0 parts.

Step	Time	Temperature
1	3 h	37 °C
2	20 min	80 °C
3	∞	12 °C

Table 3 Master Mix for Golden Gate Cloning of the signal peptide library.

Components	Amount
10× T4 Ligase Buffer	1.5 µL
pAGT572_Nemo2.0 (acceptor plasmid)	10 fmol
pAGM9121_signalpeptide mix	3 fmol of each, in total 50 fmol
pAGM9121_gene of interest	50 fmol
pAGM9121_TwinStrep-GFP11 (C-terminal tag)	50 fmol
*Bsa*I	0.5 µL
T4 DNA Ligase	1.0 µL
ddH_2O	Add to 15 µL

13. Transfer the tubes to the thermocycler and run the following program (Table 4).
14. After finishing, use the reaction mixture to transform *E coli* DH10B cells as described in point 4–8 of this part with slight modifications. As the plasmid harbors a carbenicillin resistance, use an LB agar plate supplemented with carbenicillin, X-Gal, and IPTG. After plating out a part of the cell suspension, the rest is used to directly inoculate 5 mL LB medium (+100 µg/mL carbenicillin) to preserve the diversity of the library. Incubate both the agar plate and the liquid culture overnight at 37 °C under shaking for the liquid culture.
15. The cloning efficiency is checked by blue-white selection. Isolate the plasmid library using the NucleoSpin Plasmid Kit according to the manufacturer's instructions. Determine the plasmid concentration using a Nanodrop spectrophotometer.
16. Sequence the library to confirm the diversity of the signal peptide. Store the library at −20 °C until further use.

2.4 Notes

1. The described protocol typically leads to a cloning efficiency above 90% regarding the blue-white selection. If the efficiency is significantly lower, we recommend repeating the Golden Gate reaction.

Table 4 Golden Gate cloning program for the signal peptide library.

Step	Time	Temperature
1	2 min	37 °C
2	5 min	16 °C
		Repeat step 1–2 for at least 35 cycles
3	20 min	80 °C
4	∞	12 °C

2. As the volumes of plasmid/gene for the Golden Gate Cloning are often very small, we recommend preparing dilutions prior to the reactions. For example, prepare a solution with concentrations of 10 fmol/μL (acceptor plasmid) or 50 fmol/μL (gene), and add 1 μL of each dilution to the reaction afterwards.

3. Library cultivation

High-throughput screening relies on a space- and time-efficient production of variants. In this part, we describe the cultivation of *S. cerevisiae* in 96-well microtiter plates. The *E. coli*-derived plasmid library is transformed into *S. cerevisiae* with a high transformation efficiency so that the whole diversity of the library is covered. After transformation, the protein production is performed in 96-well microtiter plates. In that case, we can cultivate more than 90 variants at the same time in a space-efficient format. As secreted enzymes, UPOs have a big advantage for high-throughput screening, as the peroxygenase-containing supernatant can be used directly for screening. No further cell lysis is necessary.

3.1 Transformation in *S. cerevisiae*

3.1.1 Equipment

- Vortex Genie 2 (Scientific industries)
- Thermomixer (Thermomixer comfort, Eppendorf)
- Centrifuge (Centrifuge 5424, Eppendorf)
- Incubator (Multitron, Infors)

3.1.2 Buffer, strains, plasmids and reagents

- Sterile 50% (v/v) polyethylene glycol 4000
- Sterile 100 mM Tris–HCl 10 mM EDTA pH 7.44
- Sterile 1 M lithium acetate pH 7.5
- Sterile ddH_2O
- 100% dimethyl sulfoxide (DMSO)
- Salmon sperm DNA (Sigma-Aldrich)
- *Saccharomyces cerevisiae* INVSc1 chemically competent cells, genotype: *MATa his3D1 leu2 trp1-289 ura3-52 MAT his3D1 leu2 trp1-289 ura3-52*
- SC Drop-out plate (1.7 g/L SC Drop Complements Kaisers Mixture without uracil, 7 g/L yeast nitrogen base without amino acids, 2% (w/v) Glucose, 2% (w/v) agar)
- UPO of interest in pAGT572_Nemo2.0
- pAGT572_Nemo2.0

3.1.3 Procedure

Timing: Day 1.

1. Thaw 10 μL salmon sperm DNA at room temperature. Additionally thaw an aliquot (50 μL) chemically competent *S. cerevisiae* INVSc1 cells and your plasmids on ice.
2. Prepare the plating buffer [40% (v/v) polyethylene glycol 4000; 100 mM lithium acetate; 1 mM EDTA; 10 mM Tris–HCl pH 7.4] using the sterile stock solutions. For one transformation you will need 600 μL buffer. Set a thermomixer to 30 °C.
3. Mix 100 ng plasmid with the salmon sperm DNA. Repeat the same procedure with the plasmid pAGT572_Nemo2.0 (negative control) and if possible, a characterized UPO (positive control).
4. Transfer the whole mix to the aliquot of cells and flip the tube several times.
5. Add 600 μL plating buffer and mix using a Vortex Genie 2 (intensity 6 of 10) until the cells are evenly distributed.
6. Place the tube into the thermomixer and incubate the cell for 30 min at 30 °C under shaking at 850 rpm.
7. After incubation, add 70 μL DMSO and mix by inverting several times. Place the tube back in the thermomixer and set the temperature to 42 °C. Incubate the cells for 15 min at 42 °C without shaking.
8. Spin down the cells (11,000*g*, 30 s) and discard the supernatant.

9. Resuspend the cell pellet in 350 µL sterile ddH_2O. Plate 100 µL suspension on an SC Drop-out plate without uracil.
10. Incubate the plate at 30 °C for 2–3 days until white colonies appear.

3.2 Cultivation in 96-well microtiter plate

3.2.1 Equipment

- Sterile toothpicks
- Incubator (Multitron, Infors)
- Universal clamp CR1800 (Enzyscreen)
- Polystyrene transparent square 96-half-deepwell microplate CR1496c (Enzyscreen)
- Sandwich cover CR1396b (Enzyscreen)
- Sterile Erlenmeyer flask or Falcon tube
- Centrifuge Avanti JXN-26 (Beckmann Coulter)
- Rotor JS-5.3 (Beckmann Coulter)
- Microtiter plate carrier "368914" (Beckmann Coulter)

3.2.2 Buffer, strains, plasmids and reagents

- Sterile ddH_2O
- Sterile 20% (w/v) Galactose
- Sterile 1.0 M potassium phosphate buffer pH 6.0
- Sterile 0.1 M magnesium sulfate
- 10× SC Drop-out stock without uracil (35 g yeast nitrogen base without amino acids, 8.6 g SC Drop Complements Kaisers Mixture without uracil in 500 mL ddH_2O)
- Sterile 20 g/L hemoglobin stock
- 100% Ethanol absolute
- 70% Ethanol
- Chloramphenicol
- SC Drop-out plates from transformation (see Section 3.1)

3.2.3 Procedure

Timing: Day 1.

1. Sterilize the microtiter plates and the sandwich cover under the sterile bench with 70% ethanol and let them dry.
2. Prepare the medium [1× SC Drop-out stock solution without uracil; 2% (w/v) galactose; 71 mM potassium phosphate buffer pH 6.0; 3.2 mM magnesium sulfate; 3.3% (v/v) ethanol; 50 mg/L hemoglobin; 25 µg/L

chloramphenicol] from the sterile stock solutions in an Erlenmeyer flask or a falcon tube. You will need approximately 25 mL medium per microtiter plate.

3. Fill each well with 220 µL medium.
4. Inoculate each well with a single, clearly separated yeast colony using toothpicks. At this step avoid cross-contamination of neighboring wells. To cover the entire diversity of the library, inoculate at least three times the number of wells as the number of theoretical distinct variants. Example: For a signal peptide library with 17 signal peptides, inoculate at least 51 wells. Do not forget to inoculate wells with pAGT572_Nemo2.0 which serves as negative control and if possible, a positive control. Additionally, keep one well without inoculation.
5. Cover the microtiter plate with the sandwich cover and mount the plate into the incubator. Incubate the plate for 72 h at 30 °C under shaking (230 rpm).

Timing: Day 3.

6. Separate the cells from the peroxygenase-containing supernatant by centrifugation (3400 rpm; 50 min; 4 °C).
   The supernatant can be used directly for screening.

3.2.4 Note

1. If you plan to detect the UPO activity based on peroxidase activity e.g., with ABTS or DMP, we recommend excluding hemoglobin from the cultivation media. Hemoglobin can also perform the reaction with ABTS and DMP that leads to a strong background signal, when it is not completely metabolized during the cultivation. Without hemoglobin, it is easier to distinguish between the activity of the variant and the background signal. *S. cerevisiae* is naturally able to produce hemoglobin.
2. We recommend to have at least four positive controls and three negative controls on one microtiter plate. One of the negative controls should only contain medium and not be inoculated with any cells as control for sterility. Discard the plate if any cells grow in the empty control well. A suitable negative control is the backbone vector.

4. Colorimetric screening methods

The most accessible high-throughput screening methods are colorimetric assays. The enzyme of interest converts hereby a substrate into a

product, which can be detected by absorbance or fluorescence measurement. Different colorimetric assays are established for UPOs. Three of these are described (Sections 4.1–4.3), and can be performed directly after cultivating different enzyme variants in a microtiter plate (see Section 3).

To enable the substrate-independent detection of novel enzymes, we developed a tag system based on the green fluorescent protein (GFP) for yeast. We used a protocol of Santos-Aberturas et al. (Santos-Aberturas, Dörr, Waldo, & Bornscheuer, 2015) and adapted it for the usage of enzyme-containing supernatant from *S. cerevisiae*. The approach relies on the formation of superfolder GFP (sfGFP). One β-sheet (GFP11, 16 amino acids) of the sfGFP is C-terminally-tagged to the enzyme of interest, combined with a TwinStrep-tag, forming in total a TwinStrep-GFP11-tag. During the assay the complementary part (GFP1–10) is added, so that the full sfGFP assembles and emits fluorescence. This assay enables an activity-independent specific, quantitative and high-throughput feasible protein detection.

4.1 DMP assay

4.1.1 Equipment

- Tissue culture test plate (TPP)
- Tecan Spark 10M (Tecan)

4.1.2 Buffer and reagents

- 2,6-Dimethoxyphenol (DMP)
- Hydrogen peroxide (H_2O_2)
- 1 M potassium phosphate buffer pH 6.0

4.1.3 Procedure

1. Transfer 40 µL supernatant from the cultivation plate to the tissue culture test plate.
2. Set up a Tecan Spark 10 M absorbance method with the following instruction:
   - kinetic mode
   - Absorption wavelength: 469 nm
   - Measurement interval: 30 s
   - Measurement time: 30 min
3. Add 160 µL screening solution (final: 100 mM potassium phosphate buffer pH 6.0, 1 mM H_2O_2, 3 mM DMP) to each well.

4. Start the measurement immediately after adding the screening solution.
5. Analyze the data regarding the slope of the linear part of the measurement curve.

4.2 ABTS assay

4.2.1 Equipment

- Tissue culture test plate (TPP)
- Tecan Spark 10M (Tecan)

4.2.2 Buffer and reagents

- 2,2′-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS)
- Hydrogen peroxide (H_2O_2)
- 1 M Citrate buffer pH 4.0

4.2.3 Procedure

1. Transfer 40 µL supernatant from the cultivation plate to the tissue culture test plate.
2. Set up a Tecan Spark 10M absorbance method with the following instruction:
   - Kinetic mode
   - Absorption wavelength: 418 nm
   - Measurement interval: 30 s
   - Measurement time: 30 min
3. Add 160 µL screening solution (final: 100 mM citrate buffer pH 4.0, 1.69 mM H_2O_2, 300 µM ABTS) to each well.
4. Start the measurement immediately after adding the screening solution.
5. Analyze the data regarding the slope of the linear part of the measurement curve.

4.3 NBD assay

4.3.1 Equipment

- Tissue culture test plate (TPP)
- Tecan Spark 10M (Tecan)

4.3.2 Buffer and reagents

- 5-Nitro-1,3-benzodioxole (NBD)
- Acetone
- Hydrogen peroxide (H_2O_2)
- 1 M potassium phosphate buffer pH 7.0

4.3.3 Procedure

1. Transfer 40 µL supernatant from the cultivation plate to the tissue culture test plate.
2. Set up a Tecan Spark 10 M absorbance method with the following instruction:
   - Kinetic mode
   - Absorption wavelength: 425 nm
   - Measurement interval: 30 s
   - Measurement time: 30 min
3. Add 160 µL screening solution [final: 100 mM potassium phosphate buffer pH 7.0, 1 mM H_2O_2, 5% (v/v) acetone, 300 µM NBD] to each well.
4. Start the measurement immediately after adding the screening solution.
5. Analyze the data regarding the slope of the linear part of the measurement curve.

4.4 splitGFP assay

4.4.1 Equipment

- 96-well Nunc MaxiSorp Fluorescence plate (Thermo Fisher Scientific)
- Tecan Spark 10M (Tecan)
- Orbital shaker
- Water bath

4.4.2 Buffer and reagents

- TNG Buffer [100 mM Tris–HCl, 100 mM NaCl; 10% (v/v) Glycerin pH 7.4]
- BSA in TNG solution [0.5% (w/v) Bovine Serum Albumin Fraction V in TNG Buffer]
- Aliquot (10 mL) sfGFP1-10 inclusion bodies, prepared according to Santos-Aberturas et al. (2015)

4.4.3 Procedure

Timing: days before measurement.

1. Prepare sfGFP1-10 as described previously (Santos-Aberturas et al., 2015).

Timing: measurement day

2. For blocking, fill each well of a 96-well Nunc MaxiSorp Fluorescence plate with 180 µL BSA in TNG solution and incubate the plate for 20 min under shaking.
3. Discard BSA solution and let the plate dry upside down.

4. Transfer 20 µL supernatant to the fluorescence plate.
5. Set up Tecan Spark 10 M fluorescence method with the excitation set to 485 nm and the emission wavelength to 535 nm. Set the gain to well calculated gain. The well for gain calculation should be set to the negative control (backbone plasmid).
6. Thaw an aliquot sfGFP 1–10 in a water bath. For the measurement solution, dilute it with 10 mL ice-cold TNG Buffer.
7. Add 180 µL measurement solution to each well.
8. Start the fluorescence measurement and store the data as starting point (t_0).
9. Incubate the plate for 1–4 nights at 4 °C. After incubation, measure the fluorescence with the same settings as for the first measurement. Store the data as endpoint (t_{end}).
10. Identify your best expressing variant by subtracting the t_0 from the t_{end} values and comparing them to the negative control.

4.5 Notes

1. Before isolating the plasmid of the best variant, we recommend to repeat the analysis with the best performing variants in replicates (at least triplicates). Resuspend the cells in the cultivation plate with an inoculation loop. Streak out the resuspended cells on SC Drop-out plates without uracil and incubate the plate overnight at 30 °C. After incubation, you can cultivate the cells in a microtiter plate as described under Section 3.
2. To minimize standard deviations, we recommend using a multichannel pipette or a Platemaster (Gilson).

5. Isolation of episomal plasmids from *S. cerevisiae*

The described screening is performed blindly, which means we do not know which mutations are included in the enzyme variant displaying the most interesting properties. After screening and identifying the best variant for our parameter of choice, the sequence is revealed. The plasmids of the selected variants need to be isolated from *S. cerevisiae*. As yeast cells, in contrast to *E. coli*, have a cell wall, the usage of commercially available Kits like NucleoSpin Plasmid Kit from Macherey-Nagel is not sufficient. We describe the isolation of episomal plasmids from *S. cerevisiae*. The isolation is performed in two parts: First, the cell wall is enzymatically digested. Second, the plasmid can be isolated using a plasmid isolation kit.

5.1 Equipment

- Thermomixer (Thermomixer comfort, Eppendorf)
- Incubator (Multitron, Infors)
- Centrifuge (centrifuge 5424, Eppendorf)
- Spectrophotometer (BioSpectrometer, Eppendorf)
- Inoculation loop

5.2 Reagents and strains

- 10 mM ethylenediaminetetraacetic acid (EDTA) pH 8.0
- Sorbitol buffer (1.2 M sorbitol, 10 mM Calcium chloride, 100 mM Tris–HCl pH 7.5, 35 mM β-mercaptoethanol)
- Lyticase from *Arthrobacter luteus* (Sigma-Aldrich)
- NucleoSpin Plasmid Kit (Macherey Nagel)
- *Escherichia coli* DH10B chemically competent cells (Thermo Fisher Scientific), genotype: *F– mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ (ara-leu)7697 galU galK λ– rpsL(StrR) nupG*
- Sterile SOC medium (2% (w/v) tryptone, 0.5% (w/v) yeast extract, 10 mM magnesium chloride, 10 mM sodium chloride, 2.5 mM potassium chloride, 20 mM Glucose, pH 7.0)
- LB-Agar plates [1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) sodium chloride, 1.5% (w/v) agar, pH 7.0] supplemented with 100 μg/mL carbenicillin
- Sterile LB media [1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) sodium chloride, pH 7.0]
- SC Drop-out plates [1.7 g/L SC Drop Complements Kaisers Mixture without uracil, 7 g/L yeast nitrogen base without amino acids, 2% (w/v) Glucose, 2% (w/v) agar]
- Sequencing primers for pAGT572_Nemo2.0

5.2.1 Procedure

Timing: Day 1.

1. Resuspend the cells of your best variant in the remaining supernatant using an inoculation loop. Streak out the cells on SC Drop-out plates and incubate the plate at 30 °C overnight.

Timing: Day 2–4

1. Use an inoculation loop to scrape off some cells from the plate. Resuspend the cells in 1 mL 10 mM EDTA pH 8.0.

2. Pellet your cells (5000*g*, 10 min) and discard the supernatant.
3. Resuspend the cells in 600 µL freshly-prepared sorbitol buffer. Add 200 U lyticase and incubate the mix for 3 h (30 °C, 600 rpm in a thermomixer) for cell wall digestion.
4. After incubation, pellet your cells by centrifugation (2000*g*, 10 min) and discard the supernatant.
5. Prepare the plasmids using the NucleoSpin Plasmid Kit starting with the alkaline lysis step following the manufacturer's instructions. In the final step, the plasmid is eluted with 25 µL elution buffer.
6. The whole eluate is used to transform one aliquot of *E. coli* DH10B (described in Section 2) plating the whole transformation mix on selective LB agar plates (+100 µg/mL carbenicillin).
7. The following day, inoculate 5 mL LB media (+100 µg/mL carbenicillin) with a single colony and incubate the culture at 37 °C overnight under shaking.
8. After incubation, isolate the plasmid using the NucleoSpin Plasmid Kit according to the manufacturer's instructions. Determine the plasmid concentration using a spectrophotometer.
9. Sequence your plasmid to identify your best variant.

6. Golden Mutagenesis

After addressing the problem of the challenging heterologous production of UPOs, we further focus on engineering these promising enzymes using directed evolution to increase activity, regioselectivity or stereoselectivity. A common technique for directed evolution is Iterative Saturation Mutagenesis (Acevedo-Rocha, Hoebenreich, & Reetz, 2014; Reetz & Carballeira, 2007), which relies on an efficient simultaneous alteration of specific protein residues. Due to the high laboratory effort of directed evolution campaigns the cloning technique needs to be fast, reproducible and cost-effective. The commercial kit QuikChange® is often used to do so. Here we present an alternative method, which is based on Golden Gate Cloning coined Golden Mutagenesis (Püllmann et al., 2019), allowing the simultaneous saturation of multiple residues in one step. To further increase the applicability of Golden Mutagenesis, we offer a fully-automated primer-design online tool https://msbi.ipb-halle.de/GoldenMutagenesisWeb/.

6.1 Equipment

- Sterile toothpicks
- Equipment for Agarose Gel Electrophoresis
- Gel imager
- PCR thermocycler machine (Labcycler basic, SensoQuest)
- PCR tubes (Sarstedt)
- Spectrophotometer (BioSpectrometer, Eppendorf)

6.2 Buffer, strains, enzymes etc.

- pAGT572_Nemo2.0
- gene which should be mutated
- T4 DNA Ligase (1–3 U/μL, Promega)
- 10× Ligase Buffer (Promega)
- *Bsa*I (10 U/μL, New England Biolabs)
- dNTPs (10 mM Mix, Thermo Fisher Scientific)
- Phusion Polymerase (Thermo Fisher Scientific)
- 5× Phusion Green HF Buffer (Thermo Fisher Scientific)
- Taq-Polymerase
- 5× GoTaq Buffer (Promega)
- Agarose (AppliChem)
- TAE-Buffer (40 mM Tris-Acetate; 2 mM EDTA, pH 8.0)
- Roti®GelStain (Carl Roth)
- GeneRuler 1 kb DNA ladder (Thermo Fisher Scientific)
- 75% DMSO
- NucleoSpin Plasmid Kit (Macherey-Nagel)
- NucleoSpin Gel and PCR Clean-up Kit (Macherey-Nagel)
- *Escherichia coli* DH10B chemically competent cells (Thermo Fisher Scientific), genotype: *F– mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ (ara-leu)7697 galU galK λ– rpsL(StrR) nupG*
- Sterile SOC medium [2% (w/v) tryptone, 0.5% (w/v) yeast extract, 10 mM magnesium chloride, 10 mM sodium chloride, 2.5 mM potassium chloride, 20 mM glucose, pH 7.0]
- LB-Agar plates [1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) sodium chloride, 1.5% (w/v) agar, pH 7.0] supplemented with 100 μg/mL carbenicillin, 50 μg/mL X-Gal, and 150 μM IPTG
- Sterile LB medium [1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) sodium chloride, pH 7.0]
- Sequencing primers for pAGT572_Nemo2.0

6.3 Procedure

6.3.1 Design of primers

Timing: Planning.

1. Open our online tool available at https://msbi.ipb-halle.de/GoldenMutagenesisWeb/.
2. Upload the gene sequence of interest in the tool.
3. In the tab configuration choose pAGT572_Nemo from the pre-existing configurations and deselect the Point "Use Level 0 to go in Level 2" (yellow part).
4. In the preview and selection tab choose the mutations.
5. Check primers regarding self-dimerization or hairpin formation. Order the primers.

6.3.2 Preparation of fragments

Timing: day 1.

1. Prepare a primer stock solution of 100 µM. Use this stock solution for preparing a 10 µM working solution.
2. Set up a PCR as mentioned in Table 5.
3. Transfer the tubes to the thermocycler and start the PCR protocol shown in Table 6.
4. During the PCR, prepare an agarose gel. The agarose concentration should be ranging between 1% (w/v) Agarose (fragments > 700 bp) and 2.5% (w/v) Agarose (fragments down to 100 bp). Use Roti® GelStain for detection.

Table 5 PCR reaction mix for fragment generation.

Compound	Amount
5× Phusion Green HF Buffer	10 µL
DMSO (75%)	2 µL
dNTPs (10 mM)	1 µL
Gene of interest	100 ng
Forward primer (10 µM)	1 µL
Reverse primer (10 µM)	1 µL
Phusion	0.5 µL
ddH_2O	Add to 50 µL

Table 6 PCR method for generating of fragments.

Step	Time	Temperature
1	1 min	98 °C
2	15 s	95 °C
3	30 s	60 °C
4	90 s per kb	72 °C
		Repeat step 2–4 for at least 35 cycles
5	10 min	72 °C
6	∞	12 °C

5. After the PCR is finished, load the gel with 5 μL of your PCR reaction. Also, load the DNA ladder to determine the size of your fragments. Run the gel and analyze it using a gel imager.
6. If the size of the fragments is correct, purify them using the NucleoSpin Gel and PCR Clean-up Kit following the instructions. Determine the concentration of the fragments at the spectrophotometer.

6.3.3 Cloning

Timing: Day 1–3.

1. Set up the Golden Gate cloning reactions as follows (Table 7).
2. Put the reactions into a thermocycler. For 1–2 fragments, use the program according to Tables 8, for 3 or more fragments, use the program mentioned in Table 9.
3. Use the whole cloning mix to transform chemically competent *E. coli* DH10B as described in Section 2. Streak out 100 μL transformation mixture on selective LB-Agar plates (+carbenicillin; +X-Gal; +IPTG). Use the rest to inoculate 5 mL LB medium (+50 μg/mL carbenicillin) and incubate overnight at 37 °C.
4. You can check the cloning efficiency by blue-white selection, as the acceptor plasmid harbors a LacZ gene. Recombinant cells appear white, and cells containing the acceptor plasmid appear blue. Prepare the library from the liquid culture using the NucleoSpin Plasmid Kit.
5. Store the plasmids at −20 °C until further use.

Table 7 Master Mix for Golden Gate Cloning of fragments.

Compound	Amount
10× Ligase buffer	1.5 µL
pAGT572_Nemo2.0 (acceptor plasmid)	10 fmol
Fragments	50 fmol each
*Bsa*I	0.5 µL
T4 DNA Ligase	1 µL
ddH_2O	Add to 15 µL

Table 8 Golden Gate Cloning method for 1–2 fragments.

Step	Time	Temperature
1	3 h	37 °C
2	20 min	80 °C
3	∞	12 °C

Table 9 Golden Gate Cloning method for ≥ 3 fragments.

Step	Time	Temperature
1	2 min	37 °C
2	5 min	16 °C
		Repeat step 1–2 for at least 35 cycles
3	20 min	80 °C
4	∞	12 °C

6.3.4 Quick-quality-control

Timing: day 3.

6.3.4.1 Check size by Colony PCR

1. Check the correct size of the insert by colony PCR. For that prepare a Master Mix according to Table 10.

Table 10 Master Mix colony PCR.

Compound	Amount
5× GoTaq Buffer	10 µL
DMSO (75%)	2 µL
dNTPs	1 µL
Forward primer (10 µM)	1 µL
Reverse primer (10 µM)	1 µL
Taq polymerase	0.5 µL
ddH_2O	Add to 50 µL

2. Fill every PCR tube with 50 µL Master Mix.
3. Scrape off a single white bacterial colony and resuspend the cells in the master mix using a toothpick.
4. Put the tubes into the thermocycler and start the program according to Table 11.
5. Analyze the PCR reactions with a 1% (w/v) agarose gel with Roti®GelStain as gel stain. Use the GeneRuler1 kb DNA ladder for the determination of the fragment size (Fig. 3A).

6.3.4.2 Sequence for diversity

1. Sequence the plasmid with compatible sequencing primers.
2. Check the results at the mutation sites. Fig. 3B shows you how it should look in case of a saturation library using the "22c trick" (Kille et al., 2013).
3. Upload the.ab1 file to our online tool to get pie charts showing the base pair distributions at specific positions (Fig. 3B).

7. Performing reactions in 96-well microtiter plate

To fulfill the requirements of a high-throughput protocol also for non-colorimetric substrates, we developed a reaction setup in 96-well microtiter plates for further analysis by GC–MS. As previously, the enzyme-containing supernatant from the cultivation in microtiter plates (see Section 3) is used. The following protocol describes the workflow

Table 11 PCR program for colony PCR.

Step	Time	Temperature
1	1 min	98 °C
2	15 s	95 °C
3	30 s	60 °C
4	90 s per kb	72 °C
		Repeat step 2–4 for at least 35 cycles
5	10 min	72 °C
6	∞	12 °C

performed for engineering *Mth*UPO toward the conversion of β-ionone (Münch et al., 2023). The general aspects of the method can be applied to every substrate but have to be adapted and optimized.

7.1 Equipment

- Incubator (Multitron, Infors)
- Universal clamp CR1800 (Enzyscreen)
- Sandwich cover CR1200 (Enzyscreen)
- Centrifuge Avanti JXN-26 (Beckmann Coulter)
- Rotor JS-5.3 (Beckmann Coulter)
- Microtiter plate carrier "368914" (Beckmann Coulter)
- 96-well glass coated microplates (1.2 mL, U-shape, 7 mm diameter) (Fisher Scientific, product nr. 16469375)
- Polypropylene square 96-deepwell microplates (CR1496, Enzyscreen)
- WebSeal Mat, 96 round well flat base (7 mm diameter) (Fisher Scientific, product nr. 15142739)

7.2 Buffer, strains, enzymes etc

- Enzyme library produced in *S. cerevisiae* in 96-well microtiter plate (see Section 3)
- 1 M Potassium phosphate buffer pH 7
- β-Ionone (substrate)
- α-bisabolol (internal standard)

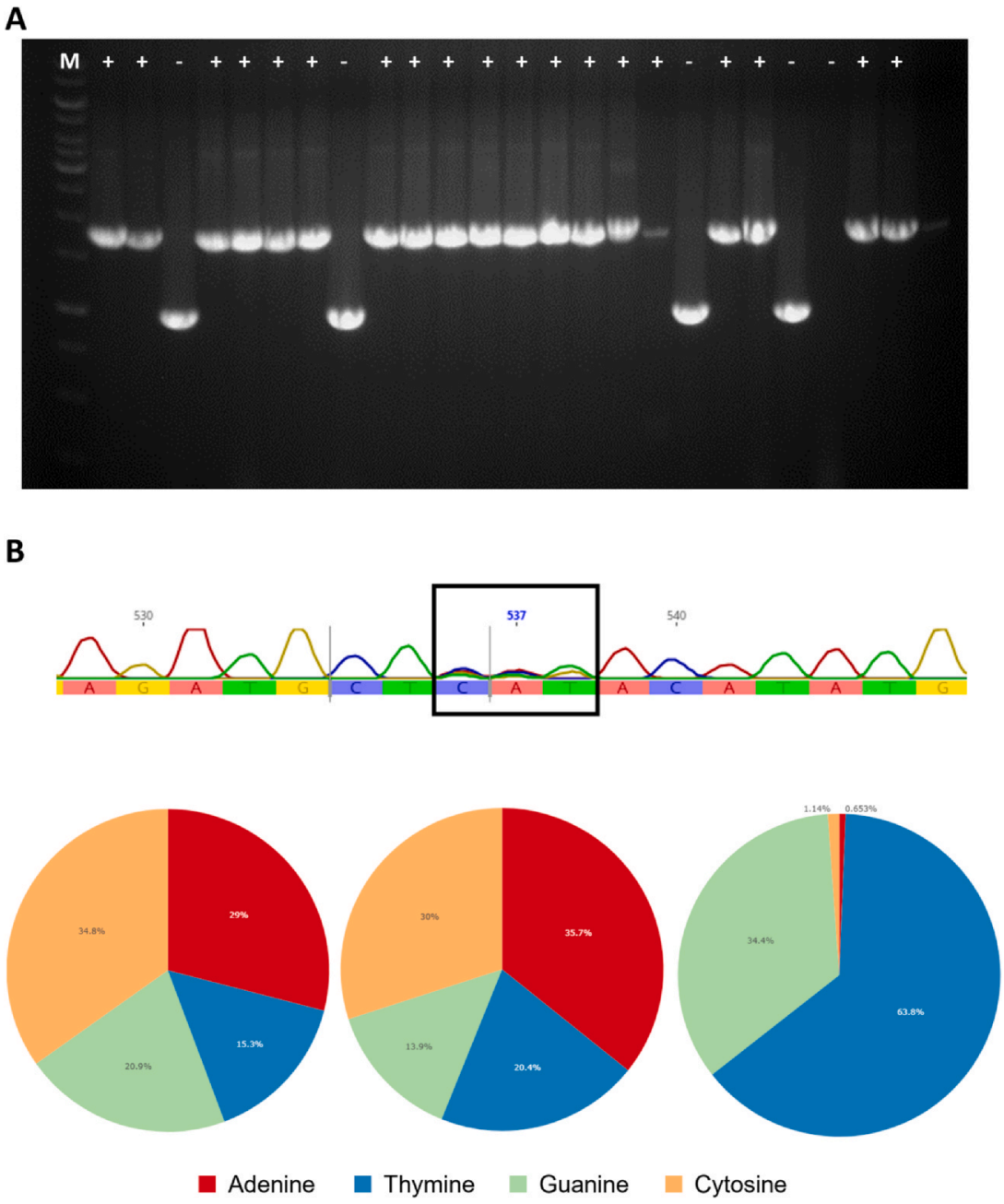


Fig. 3 Example of Quick-Quality-Control. This figure displays an example for the Quick-Quality-Control by colony PCR (A) and sequencing (B). (A) shows an agarose gel (1%) of a colony PCR. The DNA was stained by Roti®GelStain and detected by a gel imager. The expected fragment size was 1700 bp. (B) shows the sequencing of a mutant library at position 154 of the *Mth*UPO gene prepared using the "22c trick" (Kille et al., 2013). The upper part demonstrates the analysis of the sequencing in Geneious Prime. The lower part displays the base distribution provided by the Quick-Quality-Control function of our online tool. M: marker (GeneRuler 1 kB DNA ladder), +: correct assembly, −: incorrect assembly.

- Ethyl acetate (solvent) (GC grade)
- Acetone (co-solvent)
- Hydrogen peroxide

7.3 Procedure

Timing: after cultivation.

1. Transfer 100 µL supernatant to a polypropylene deepwell microtiter plate.
2. Add 400 µL reaction master mix per well (final concentrations: 100 mM potassium phosphate buffer pH 7.0, 1.0 mM β-ionone, 1.0 mM H_2O_2, 5% (v/v) acetone).
3. Cover the plate with the sandwich cover and perform the reaction in the incubator at 30 °C under shaking (300 rpm) for 1 h.
4. After the reaction, add 500 µL ethyl acetate containing 250 µM α-bisabolol as internal standard per well. For extraction, incubate the plate at 300 rpm for 30 min.
5. To separate the phases, centrifuge the plate for 5 min at 3000 rpm.
6. After centrifugation, transfer 300 µL of the organic phase to the glass-coated 96-well microtiter plate and seal the plate with the WebSeal Mat. The sample can be analyzed by MISER-GC–MS.

7.4 Notes

1. The described protocol was optimized for the conversion of β-ionone by *Mth*UPO. (Münch et al., 2023) The method has to be adapted and optimized for the substrate of choice.
2. To minimize standard deviations, we recommend using a multichannel pipette or a Platemaster (Gilson) for distributing the master mix, adding the extracting solvent and removing the organic phase for further analysis.
3. To avoid extensive evaporation, it is important to tightly seal the plate and minimize the time in which the plate is not covered.

8. MISER-GC–MS

As directed evolution is highly substrate dependent, the closer one can work to the original substrate, the better. To assay the chemo- and regioselectivity of UPOs toward a specific substrate, we developed *multiple injection in a single experimental run* method for GC–MS (MISER-GC–MS),

which is capable of screening the activity of 96 mutants toward up to seven analytes in around one hour. (Knorrscheidt et al., 2021a, 2021b) In this method, the autosampler is controlled by a separate software (official software of the supplier), which enables the multiple injection of samples in a single isothermal GC sequence. The ions of interest are analyzed in a SIM method and quantified relative to the internal standard. All of those settings are possible for a regular GC–MS instrument equipped with an autosampler. We describe the general workflow of establishing a MISER-GC–MS method, but the method itself has to be adapted to the substrate of interest. The described method was established and optimized for engineering *Mth*UPO toward the conversion of β-ionone (Münch et al., 2023). The general aspects of the protocol can be applied to every substrate but have to be adapted and optimized.

8.1 Reagents and equipment

8.1.1 Reagents

- Glass-coated microtiter plate with organic phase from the reaction (see Section 7)
- β-Ionone (substrate)
- 4-Hydroxy-β-ionone (product)
- α-bisabolol (internal standard)
- Ethyl acetate (GC grade)
- 100 mM potassium phosphate buffer pH 7.0

8.1.2 Equipment

- Polypropylene square 96-deepwell microplates (CR1496, Enzyscreen)
- 96-well glass-coated microplates (1.2 mL, U-shape, 7 mm diameter) (Fisher Scientific, product nr. 16469375)
- GC vials with lids
- WebSeal Mat, 96 round well flat base (7 mm diameter) (Fisher Scientific, product nr. 15142739)
- sandwich cover CR1200 (Enzyscreen)
- GC–MS-QP2010 Ultra (Shimadzu)
  - Helium as carrier gas
  - MS: dual-stage turbomolecular pumps + quadrupole
- auto sampler AOC-5000 Auto Injector (Shimadzu)
- PAL MTHolder (CTC Analytics AG)
- Cycle Composer (Version 1.6.0, CTC Analytics AG)

- GC column OPTIMA 5 MS Accent (25 m × 0.20 mm, 0.2 μm df) (Macherey-Nagel, product nr. 725810.25)
- Split/splitless uniliner inlets w/wools (3.5 mm, 5.0 × 95 mm), for Shimadzu GC (Restek, product nr. 21719)
- Centrifuge Avanti JXN-26 (Beckmann Coulter)
- Rotor JS-5.3 (Beckmann Coulter)
- Microtiter plate carrier "368914" (Beckmann Coulter)
- Vortex Genie 2
- Centrifuge (Centrifuge 5424, Eppendorf)

8.2 Ions selection for SIM mode

8.2.1 Procedure

Timing: before first MISER run.

1. Prepare a sample containing 500 μM β-ionone, 500 μM 4-hydroxy-β-ionone and 500 μM α-bisabolol in ethyl acetate.
2. Analyze 1 μL of the mixture by GC–MS in SCAN mode with the following settings:
   - Interface temperature: 290 °C
   - Injection temperature: 280 °C
   - Ion source temperature: 250 °C
   - Ionization voltage: 70 V
   - Temperature program:
     70 °C
     50 °C/min to 200 °C
     100 °C/min to 300 °C hold 5 min
3. Identify ions that have high relative abundance and are unique to the substrate, the product, and the internal standard by comparing the mass spectra (Table 12).

Table 12 Identified *m/z*.

	Molecule	*m/z* for SIM
Substrate	β-Ionone	177
Product	4-Hydroxy-β-ionone	208
Internal standard	α-Bisabolol	204

8.2.2 Note

1. For quantification and minimization of standard deviations the internal standard should have a similar extraction behavior and GC retention time as the desired product of your reaction.
2. The selected ion should be exclusive to only the substrate, the product or the internal standard and must have relative high abundance, and preferably high m/z, to minimize the interference from solvent ions.
3. Reduce injection temperature to 200 °C in case of analyzing epoxides.

8.3 Verification of linearity of internal standard in extraction and MS response

8.3.1 Procedure

Timing: before first MISER run.

1. Transfer 400 μL of reaction buffer (potassium phosphate buffer pH 7.0) to a reaction vial. Add 1 mL of ethyl acetate containing different concentrations of α-bisabolol (e.g. 100 μM, 500 μM and 1000 μM).
2. Perform the extraction by mixing using a Vortex Genie2. Afterward, separate the phases by centrifugation (1 min, 5000 rpm).
3. Transfer the organic phase to a GC vial and analyze it using the GC–MS in SIM mode with m/z of 204. Use the same parameters as mentioned in Section 8.2.
4. Plot the peak areas against the concentration of α-bisabolol. And define the linearity by performing linear regression.

8.3.2 Note

1. If the relation is not linear, the concentration of the internal standard should be adjusted to find the linear range or choose a different internal standard.

8.4 Preparation of GC–MS device for MISER experiments

8.4.1 Procedure

Timing: before first MISER measurement.

1. Open the program "GCMS Real Time Analysis".
2. Go to "Instrument"—"System configuration" and remove the autosampler "AOC-5000" from the Analytical Line. Confirm the selection with "Set". Close the program.
3. Open the program "PAL Cycle Composer with Macro Editor"

4. Go to "Method Editor". Select the Method (.PME) and Macro file (.pma). After selection, the Method and Macro file, click on "Replace" at the left bottom of the window to ensure that the correct Method and Macro files are loaded.
5. Open the program "GCMS Real Time Analysis". It is important that the "GCMS Real Time Analysis" is always opened after "PAL Cycle Composer with Macro Editor" starts. This ensures that the Autosampler is controlled by the "PAL Cycle Composer with Macro Editor", instead of "GCMS Real Time Analysis".
6. Open or create a MISER Batch File (.qgb) (Fig. 4) in "GCMS Real Time Analysis". Since the autosampler is controlled by "PAL Cycle Composer", choose any sampler file or leave it blank if your software allows it.
7. Start the GC–MS in "GCMS Real Time Analysis" program. Wait until all the settings are loaded and all the parameters are stable before continuing to the next step.
8. Change to the "PAL Cycle Composer" program. Ensure that the correct Method (.PME) and Macro (.pma) are loaded. If everything is set, start the autosampler by clicking the "Run" button and click "Okay" in the pop-up window.
9. After the autosampler starts to move, immediately press the green "Start" button on the GC–MS device to start the temperature program. This step is important for synchronizing the autosampler and the GC–MS.

8.4.2 Note

1. After selecting the Method and Macro file all modifications should be performed in the selected Macro file with "PAL Cycle Composer with Macro Editor". As for detailed tutorial of the software "PAL Cycle Composer with Macro Editor", please refer to the manual.
2. Different instruments could have different methods to decouple the autosampler from their default control software. Please refer to the instrument producer for detailed instructions.
3. Be aware that a batch file for MISER measurement only consists of one row (Fig. 4).

	Client	Sample ID	Sample Name	Data File	Sample Type	Vial#	Method File	Tuning File
1	CLIENT	20220422	Round 1 Plate C2	H_001.qgd	0:Unknown	1	MISER230_Split50_100min.qgm	_tune_F1_220419.qgt

Fig. 4 Example of Setup in GCMS Real Time Analysis for a MISER experiment.

8.5 Optimization of MISER GC–MS method

8.5.1 Procedure

Timing: Before measurement of library.

1. Prepare a 1 mL sample of 500 μM β-ionone, 500 μM 4-hydroxy-β-ionone and 500 μM α-bisabolol in ethyl acetate.
2. Compose a Macro file for at least 10 injections as in Fig. 5. The times of injections is determined by the repetitive execution of the module "GET_SAMPLE(), PUT_SAMPLE() and CLEAN_SYR()".
3. Write a GC method file for the isocratic temperature profile of the MISER-GC–MS method. Run your composed GC method and Macro file to optimize the MISER method. During the optimization, change one parameter at a time. The parameters have different effects (see Table 13) and have to be adjusted to find an optimized method (Table 14).

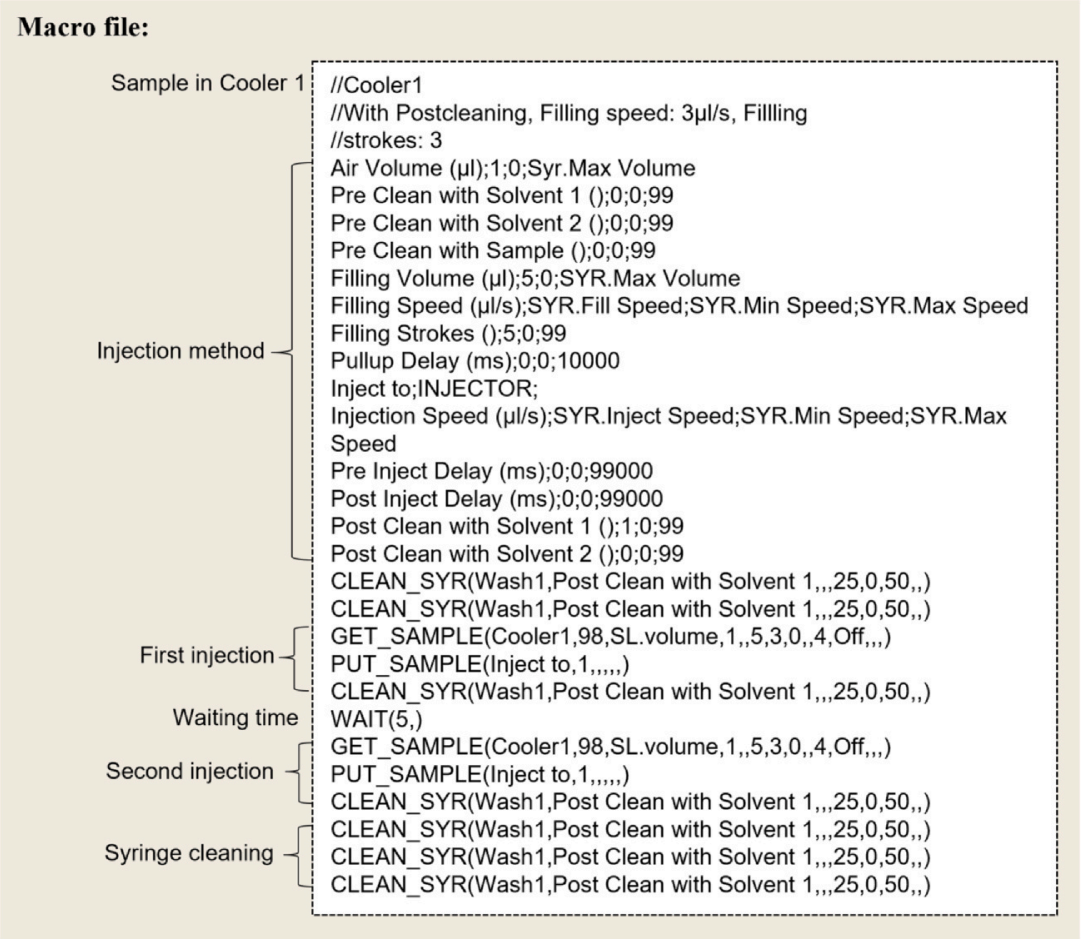


Fig. 5 Example Macro file.

Table 13 Parameter for optimization of MISER method.

Oven temperature	Defines how fast the molecules elute from the column. Should be set as high as possible, a separation of the different substrates, products and internal standard on the GC column is not necessary. However, peaks with the same m/z cannot overlap.
Split ration	Defines how much of your sample is inserted to the column. It should be set as high as possible to have sharp peaks with good resolution but without compromising the intensity of the signal in the chromatogram.
Pressure/flow/linear velocity	Have effects on the resolution and velocity of elution. Can be manipulated if you have problems with resolution.
Post-cleaning steps	Defines the number of washing steps. It is essential to avoid cross-contamination. Increasing the number of washing steps extends the time between two injections and the total run time.
Filling speed	Defines how fast the syringe is loaded and essential for application of the same volume for every injection. Can be manipulated if the standard deviations are high. In the case of viscous samples, decrease the filling speed.
Waiting time	Defines how long you wait before the next injection. Should be adjusted that all molecules are eluted before the next injection to avoid overlapping between elution peaks and the solvent peak of the next injection.

8.5.2 Note

1. It is vital to avoid the overlap between the product and internal standard peak of one injection and the solvent peak of the next injection. The overlap will cause a major error in the integration of the target peak, and thus lead to high standard deviations. The interval time between injections can be adjusted by the number of post-cleaning steps and the "WAIT ()" command.

Table 14 Optimized parameters for the analysis of the β-ionone conversion by *Mth*UPO.

	Settings	Parameters
GC–MS method (controlled by GCMS Real Time Analysis)	Isocratic oven temperature	230 °C
	Split ratio	50
	Pressure/Total Flow	23.1 mL/min
	Linear velocity	28.7 cm/s
Autosampler method (controlled by PAL Cycle Compose)	Post-cleaning steps	3
	Filling Speed	3 μL/s
	Waiting time	5 s
	Resulting interval time	56 s

2. All internal standard peaks in a chromatogram from an optimized MISER method should have similar height and area over the full measurement (standard deviation below 3%).
3. Command GET_SAMPLE(Cooler1,98,SL.volume,1, ,5,3,0, ,4,Off, , ,) specifies the position and method of getting samples. PUT_SAMPLE (Inject to, 1, , , , ,) specifies the sample being injected to injector 1 (please be aware whether you have multiple injectors). and CLEAN_SYR(Wash1,Post Clean with Solvent 1, , ,25, 0,50, ,) specifies the cleaning procedure after injection.
4. The GC–MS temperature sequence should be long enough to run all the injections.
5. For substances that are hard to flush out of the GC column, a column heating step can be added after a certain number of injections (e.g. after 12 injections are eluted, heat the column with 200 °C/min to 300 °C, hold for 5 min, then decrease to 230 °C and start another injection). Besides changing the temperature program for the GC–MS, "WAIT (HEATING TIME)" must be added after the correspondent numbers of injection.

8.6 Determination of standard deviation of MISER GC–MS method

8.6.1 Procedure

Timing: before library screening.

1. Prepare three test runs as described in Table 15 for determining the standard deviation. All samples should be tested in 96-well glass-coated microplates as 96 replicates.
2. Run the developed MISER program (Fig. 6) and analyze the standard deviation of each sample.
3. As result the standard deviations over the whole plate should be <3% for RUN A, <5% for RUN B and <10% for RUN C.

8.6.2 Note

1. If any injection peak (internal standard) is missing, check whether it is caused by a faulty autosampler method, a too low sample volume/missing sample in the glass-coated microtiter plate or incorrect syringe filling before sample injection. Optimize accordingly.
2. For minimizing the standard deviation caused by pipetting, we recommend using a multichannel pipette or Platemaster (Gilson) for distributing solutions over the plate whenever possible.
3. To avoid extensive evaporation, it is important to tightly seal the plate and minimize the time in which the plate is not covered.

Table 15 Test runs for determination of the standard deviation.

RUN A	1. Prepare a mixture of 500 μM β-ionone, 4-hydroxy-β-ionone and α-bisabolol in ethyl acetate. You will need approximately 30 mL. 2. Distribute 300 μL of the mixture per well to a 96-well glass-coated microtiter plate.
RUN B	1. Prepare a reaction mix of 500 μM β-ionone and 500 μM 4-hydroxy-β-ionone in 100 mM potassium phosphate buffer pH 7.0 (5% v/v) acetone. You will need approximately 50 mL. 2. Add 500 μL per well of a 96-well deepwell microtiter plate. 3. Perform the extraction as described in Section 7.
RUN C	1. Express *Mth*UPO wild type in *S. cerevisiae* in a microtiter plate (Section 3). 2. Perform the reaction and extraction (Section 7).

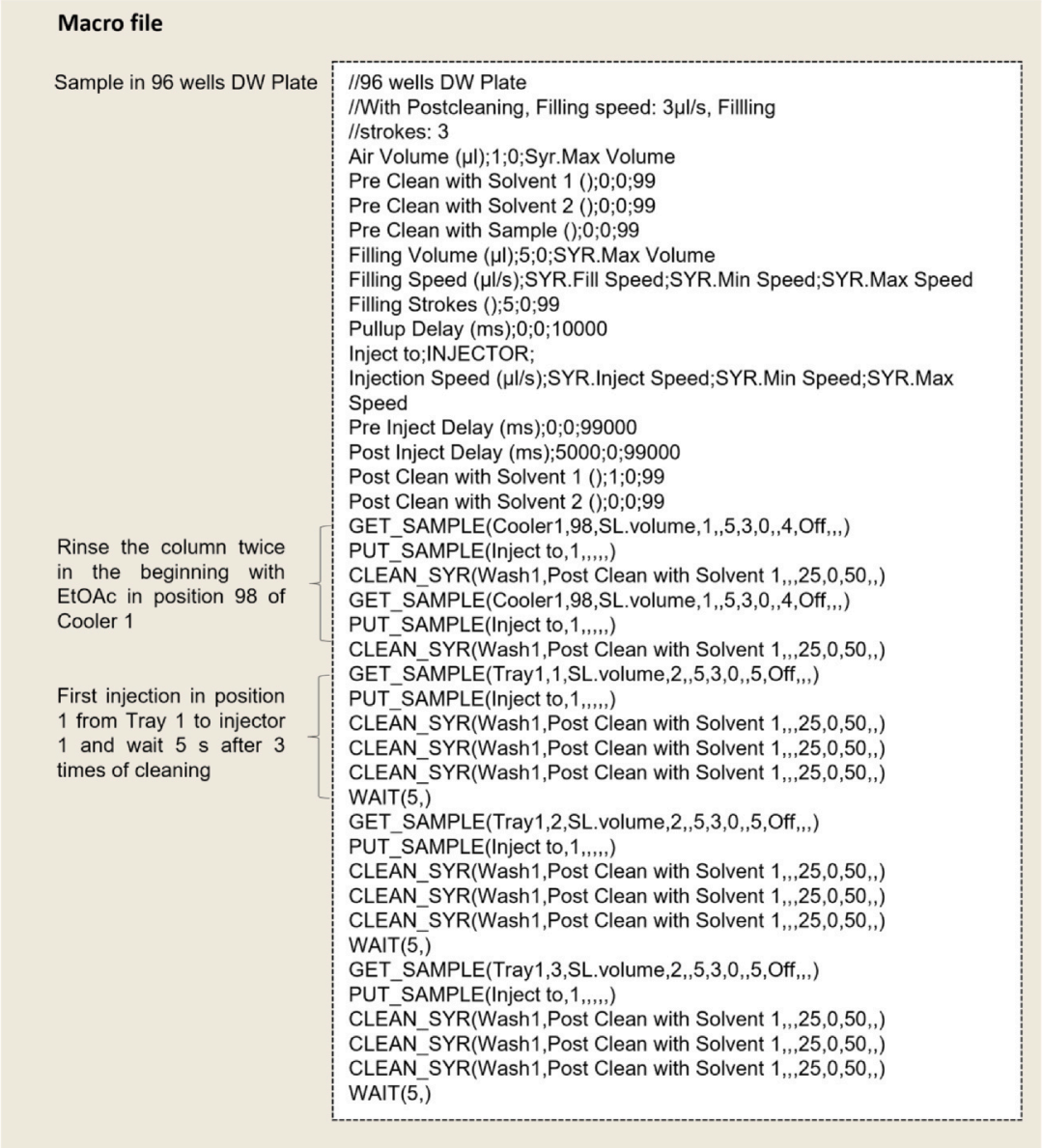


Fig. 6 Example Macro file with optimized parameters.

4. The concentration of the internal standard should be within the same range as the product concentration to minimize standard deviations. If necessary, adjust the concentration.

8.7 MISER GC–MS method for high-throughput screening

8.7.1 Procedure

1. Cultivate your library as described in Section 3. Each plate should contain 88 wells with different mutants, five wells with parental variant, two wells with strain with empty vector (negative control), and one well without cells.

2. Perform the reaction and extraction as described in Section 7.
3. Run your optimized MISER method.
4. Analyze the data to find mutants with increased activity (see Section 8.8).

8.8 Data evaluation

The data evaluation is performed in "GCMS Postrun Analysis" and a spreadsheet program like Microsoft Excel.

8.8.1 Procedure

1. Check if all samples are successfully measured. Open the output data file with the data analysis software (GCMS Postrun Analysis). Display only the m/z of internal standard and check if there are 96 regular peaks with similar heights and even intervals (Fig. 7A).
2. First, determine the peak area of the lowest product peak by manual peak integration. Set this value as the threshold for automatic peak integration. Repeat the step for the internal standard peaks (Fig. 7B).
3. Check if all 96 internal standard peaks and 96 product peaks are integrated. If one or both peaks are missing for one injection, integrate the correspondent chromatogram area anyway, to ensure that the position of the 96-well plate and the sampling sequence correspond to each other.
4. Copy the peak areas to a spreadsheet program of your choice, like Microsoft Excel
5. Divide the product peak by the internal standard peak in each well and compare them to the negative controls (cell with backbone plasmid) and parental variant.

8.8.2 Note

1. Hits found with the MISER method should be verified at least in triplicates. To do so, a streak out of the selected variant using cells from the original cultivation microtiter plate (Section 3) is prepared. Afterward, cells are cultivated as described in Section 3. Reactions and measurements can be performed either using the MISER method or in single runs. Plasmids from confirmed hits can be isolated and sequenced to reveal the mutations (see Section 6).

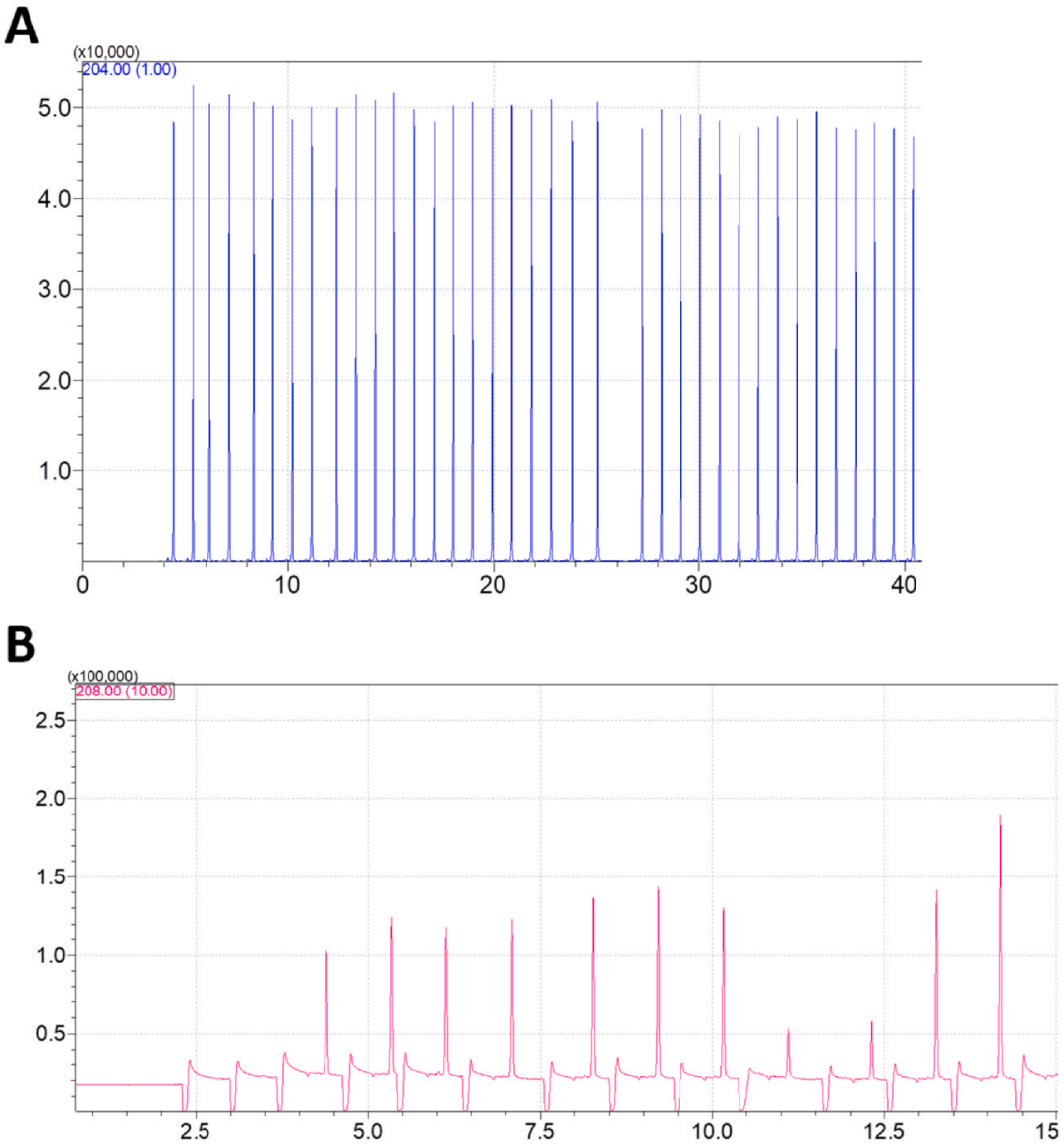


Fig. 7 Example chromatograms. (A) display a part of the chromatogram for *m/z* of 204 which is correlated to α-bisabolol (internal standard). (B) show a part of chromatogram for the product 4-hydroxy-β-ionone with a *m/z* of 208. Both chromatograms are measured using the MISER program described by Münch et al. (2023).

2. It is essential to maintain the strict correspondent relationship between the position of the 96-well plate and the sampling sequence to find the right mutant.
3. It is essential to check peak integration manually. Product and internal standard peaks should have even intervals. Peaks with longer time gaps indicate missing peak(s) in between or inactive variants.
4. A hit with improved activity should have at least 1.5-fold activity compared with the parental strain.

9. Summary and conclusion

This chapter provides a workflow for secretion and directed evolution of unspecific peroxygenases. Starting from a gene fragment, the enzymes were introduced into a modular cloning system, which is used to perform a signal peptide shuffling to improve the secretion of the enzymes in *S. cerevisiae*. The shuffling can be performed with 17 distinct signal peptides to provide a wide variety. All elements used in the signal-peptide-shuffling are available on the non-commercial platform AddGene. To screen the library, we provide a method for a space- and time-efficient cultivation of *S. cerevisiae* in microtiter plate with the possibility to directly use the supernatant for analysis. The analysis of the expression is described for three standard activity-based and one activity-independent colorimetric assays, which were adapted for the use of supernatant from *S. cerevisiae*. After identifying the best variant, the plasmids are isolated by a modified instruction of a commercially available plasmid kit. The described techniques were used to increase the secretion of established UPOs and expression of novel enzymes. (Püllmann & Weissenborn, 2021; Püllmann et al., 2021). The methods can also be applied for the secretion *P. pastoris* and improved by a combination of a promotor and signal peptide shuffling. In addition to UPOs, the feasibility of application to other enzyme classes such as lipases and laccases was demonstrated. (Püllmann & Weissenborn, 2021). All cloning parts for the transfer into *P. pastoris* are also available on AddGene.

After successful secretion of enzymes, the enzyme can be used for industrial applications. For that purpose, the enzymes can be engineered by directed evolution. We provide a method for efficient mutagenesis on multiple sites. The protocol can be performed in one day and has an efficiency of up to 100%, which is a big advantage to other mutagenesis methods like QuikChange. Additionally, we provide an online tool for automated primer design, so that no further knowledge in primer design is necessary. The successful application of this method was demonstrated in multiple projects. (Knorrscheidt et al., 2020; Knorrscheidt et al., 2021a; Münch et al., 2023).

For screening such a mutant library independently from colorimetric assays we describe the possibility to use a normal GC–MS for high-throughput screening. By using the technique of *multiple injection in single experiment run* (MISER) it is possible to screen a 96-well plate in approximately one hour. The method has been successfully used in multiple high-throughput screening campaigns. (Knorrscheidt et al., 2020; Knorrscheidt et al., 2021a; Münch et al., 2023) We also demonstrate an

improved protocol for simultaneous detection of up to three substrates and six products from one reaction mixture. (Knorrscheidt et al., 2021b).

Acknowledgment

L.W. and J.M. thank the German Research Foundation (DFG, project ID 43649874, TP A05, RTG 2670) for generous funding and N.D. thanks the German Research Foundation (DFG, project ID 505185500) for their support. J.M. thanks the Bundesministerium für Bildung und Forschung (Maßgeschneiderte Inhaltsstoffe 2, 031B0834A).

References

Acevedo-Rocha, C. G., Hoebenreich, S., & Reetz, M. T. (2014). Iterative saturation mutagenesis: A powerful approach to engineer proteins by systematically simulating Darwinian evolution. In E. M. J. Gillam, J. N. Copp, & D. Ackerley (Eds.). *Directed evolution library creation: Methods and protocols* (pp. 103–128). New York: Springer. ⟨https://doi.org/10.1007/978-1-4939-1053-3_7⟩.

Beltran-Nogal, A., Sanchez-Moreno, I., Mendez-Sanchez, D., de Santos, P. G., Hollmann, F., & Alcalde, M. (2022). Surfing the wave of oxyfunctionalization chemistry by engineering fungal unspecific peroxygenases. *Current Opinion in Structural Biology, 73*, 102342. https://doi.org/10.1016/j.sbi.2022.102342.

Hofrichter, M., Kellner, H., Pecyna, M. J., & Ullrich, R. (2015). *Fungal unspecific peroxygenases: Heme-thiolate proteins that combine peroxidase and cytochrome P450 properties*. Springer International Publishing, 341–368. ⟨https://doi.org/10.1007/978-3-319-16009-2_13⟩.

Hofrichter, M., & Ullrich, R. (2014). Oxidations catalyzed by fungal peroxygenases. *Current Opinion in Chemical Biology, 19*, 116–125.

Kille, S., Acevedo-Rocha, C. G., Parra, L. P., Zhang, Z.-G., Opperman, D. J., Reetz, M. T., & Acevedo, J. P. (2013). Reducing codon redundancy and screening effort of combinatorial protein libraries created by saturation mutagenesis. *ACS Synthetic Biology, 2*(2), 83–92. https://doi.org/10.1021/sb300037w.

Kinner, A., Rosenthal, K., & Lutz, S. (2021). Identification and expression of new unspecific peroxygenases - Recent advances, challenges and opportunities. *Front Bioeng Biotechnol, 9*, 705630. https://doi.org/10.3389/fbioe.2021.705630.

Kleiner-Grote, G. R. M., Risse, J. M., & Friehs, K. (2018). Secretion of recombinant proteins from *E. coli*. *Engineering in Life Sciences, 18*(8), 532–550. https://doi.org/10.1002/elsc.201700200.

Knorrscheidt, A., Püllmann, P., Schell, E., Homann, D., Freier, E., & Weissenborn, M. J. (2020). Identification of novel unspecific peroxygenase chimeras and unusual YfeX axial heme ligand by a versatile high-throughput GC-MS approach. *ChemCatChem, 12*(19), 4788–4795. https://doi.org/10.1002/cctc.202000618.

Knorrscheidt, A., Soler, J., Hünecke, N., Püllmann, P., Garcia-Borràs, M., & Weissenborn, M. J. (2021a). Accessing chemo- and regioselective benzylic and aromatic oxidations by protein engineering of an unspecific peroxygenase. *ACS Catalysis, 11*(12), 7327–7338. https://doi.org/10.1021/acscatal.1c00847.

Knorrscheidt, A., Soler, J., Hünecke, N., Püllmann, P., Garcia-Borràs, M., & Weissenborn, M. J. (2021b). Simultaneous screening of multiple substrates with an unspecific peroxygenase enabled modified alkane and alkene oxyfunctionalisations. *Catalysis Science & Technology, 11*(18), 6058–6064. https://doi.org/10.1039/D0CY02457K.

Molina-Espeja, P., Garcia-Ruiz, E., Gonzalez-Perez, D., Ullrich, R., Hofrichter, M., & Alcalde, M. (2014). Directed evolution of unspecific peroxygenase from Agrocybe aegerita. *Applied and Environmental Microbiology, 80*(11), 3496–3507. https://doi.org/10.1128/aem.00490-14.

Münch, J., Püllmann, P., Zhang, W., & Weissenborn, M. J. (2021). Enzymatic hydroxylations of sp3-carbons. *ACS Catalysis, 11*(15), 9168–9203. https://doi.org/10.1021/acscatal.1c00759.

Münch, J., Soler, J., Hünecke, N., Homann, D., Garcia-Borràs, M., & Weissenborn, M. J. (2023). Computational-aided engineering of a selective unspecific peroxygenase toward enantiodivergent β-Ionone hydroxylation. *ACS Catalysis, 13*(13), 8963–8972. https://doi.org/10.1021/acscatal.3c00702.

Püllmann, P., Knorrscheidt, A., Münch, J., Palme, P. R., Hoehenwarter, W., Marillonnet, S., ... Weissenborn, M. J. (2021). A modular two yeast species secretion system for the production and preparative application of unspecific peroxygenases. *Communications Biology, 4*(1), 562. https://doi.org/10.1038/s42003-021-02076-3.

Püllmann, P., Ulpinnis, C., Marillonnet, S., Gruetzner, R., Neumann, S., & Weissenborn, M. J. (2019). Golden mutagenesis: An efficient multi-site-saturation mutagenesis approach by Golden Gate cloning with automated primer design. *Scientific Reports, 9*(1), 10932. https://doi.org/10.1038/s41598-019-47376-1.

Püllmann, P., & Weissenborn, M. J. (2021). Improving the heterologous production of fungal peroxygenases through an episomal Pichia pastoris promoter and signal peptide shuffling system. *ACS Synthetic Biology, 10*(6), 1360–1372. https://doi.org/10.1021/acssynbio.0c00641.

Reetz, M. T., & Carballeira, J. D. (2007). Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. *Nature Protocols, 2*(4), 891–903. https://doi.org/10.1038/nprot.2007.72.

Santos-Aberturas, J., Dörr, M., Waldo, G. S., & Bornscheuer, U. T. (2015). In-depth high-throughput screening of protein engineering libraries by split-GFP direct crude cell extract data normalization. *Chemistry & Biology, 22*(10), 1406–1414.

Tippelt, A., & Nett, M. (2021). Saccharomyces cerevisiae as host for the recombinant production of polyketides and nonribosomal peptides. *Microbial Cell Factories, 20*(1), 161. https://doi.org/10.1186/s12934-021-01650-y.

4 | Chapter II

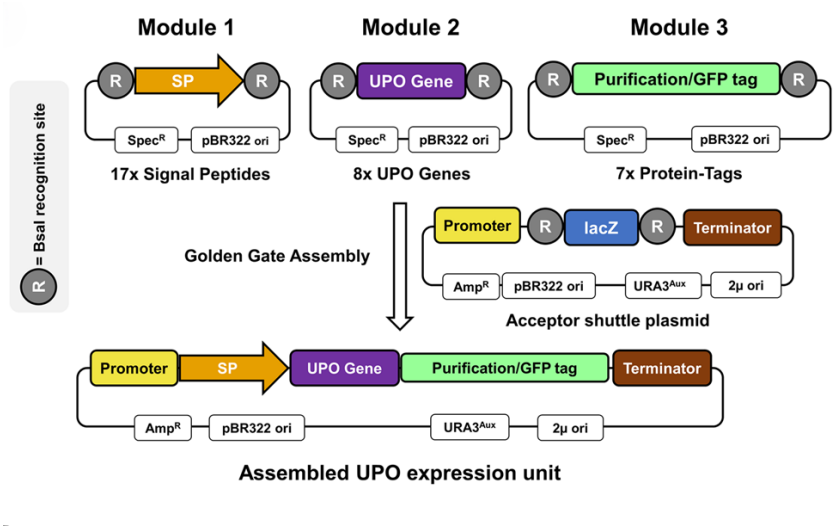
This chapter has been published as:

A modular two yeast species secretion system for the production and preparative application of unspecific peroxygenases

by: Pascal Püllmann, Anja Knorrscheidt, **Judith Münch**, Paul R. Palme, Wolfgang Hoehenwarter, Sylvestre Marillonnet, Miguel Alcalde, Bernhard Westermann, Martin J. Weissenborn

in: *Communications Biology* **2021**, *4*, 562; doi: https://doi.org/10.1038/s42003-021-02076-3

Supplementary information can be assessed free of charge at the article's landing page.



This chapter describes the construction of a modular Golden Gate system for the secretion of unspecific peroxygenases (UPOs) in *Saccaromyces cerevisiae* and *Pichia pastoris* as heterologous hosts. This system is used throughout all further projects within this dissertation. In this chapter it was applied and combined with a modular signal peptide shuffling approach to secrete five UPOs, four of them for the first time in yeast. Two of these UPOs are a newly discovered and not prior annotated as UPO. One of them, *Mth*UPO, was successfully implemented in a preparative scale reaction, the enantioselective hydroxylation of the pharmaceutical relevant building block phenethylamine, as proof-of-concept. The herein discovered *Mth*UPO is later used in two further engineering project within this thesis.

communications biology

ARTICLE

https://doi.org/10.1038/s42003-021-02076-3 OPEN

A modular two yeast species secretion system for the production and preparative application of unspecific peroxygenases

Pascal Püllmann[1], Anja Knorrscheidt[1], Judith Münch[1], Paul R. Palme[1], Wolfgang Hoehenwarter[1], Sylvestre Marillonnet[1], Miguel Alcalde[2], Bernhard Westermann[1,3] & Martin J. Weissenborn[1,3]✉

Fungal unspecific peroxygenases (UPOs) represent an enzyme class catalysing versatile oxyfunctionalisation reactions on a broad substrate scope. They are occurring as secreted, glycosylated proteins bearing a haem-thiolate active site and rely on hydrogen peroxide as the oxygen source. However, their heterologous production in a fast-growing organism suitable for high throughput screening has only succeeded once—enabled by an intensive directed evolution campaign. We developed and applied a modular Golden Gate-based secretion system, allowing the first production of four active UPOs in yeast, their one-step purification and application in an enantioselective conversion on a preparative scale. The Golden Gate setup was designed to be universally applicable and consists of the three module types: i) signal peptides for secretion, ii) UPO genes, and iii) protein tags for purification and split-GFP detection. The modular episomal system is suitable for use in *Saccharomyces cerevisiae* and was transferred to episomal and chromosomally integrated expression cassettes in *Pichia pastoris*. Shake flask productions in *Pichia pastoris* yielded up to 24 mg/L secreted UPO enzyme, which was employed for the preparative scale conversion of a phenethylamine derivative reaching 98.6 % *ee*. Our results demonstrate a rapid, modular yeast secretion workflow of UPOs yielding preparative scale enantioselective biotransformations.

[1] Leibniz Institute of Plant Biochemistry, Halle (Saale), Germany. [2] Department of Biocatalysis, Institute of Catalysis, CSIC, Madrid, Spain. [3] Institute of Chemistry, Martin-Luther-University Halle-Wittenberg, Halle (Saale), Germany. ✉email: martin.weissenborn@ipb-halle.de

Fungal unspecific peroxygenases (UPOs) have recently emerged as novel versatile hydroxylation biocatalysts. They solely rely on hydrogen peroxide as cosubstrate reaching impressive total turnover numbers for sp^3-carbon hydroxylation of up to 300000[1–4]. Further UPO catalysed reactions include aromatic hydroxylation, heteroatom oxidation, halogenation and carbon–carbon double bond epoxidation[5]. Due to their high oxyfunctionalisation versatility and activity, while solely requiring hydrogen peroxide and being independent of auxiliary electron transport proteins and expensive cofactors, UPOs have attracted keen interest in the biocatalysis field[5–7]. Current challenges include suboptimal regio- and enantioselectivities towards specific substrates and the extremely limited panel of available UPOs, impeding broader screening setups to find a suitable catalyst to perform a particular reaction. There is an estimated number of more than 4000 putative UPO genes currently annotated and widely spread within the fungal kingdom representing just a small fraction of the available genetic diversity[8].

To provide further insight into the natural function of UPOs as well as broadening the available substrate scope, it is crucial to access more enzymes from diverse phylogenetic backgrounds. Recent studies reported on the conversion of testosterone by a UPO derived from an ascomycetous mould, a reaction that could not be performed by any other UPO thus far[9]. This new activity further emphasises the limitations of the small available UPO panel. It would be further highly desirable to heterologously produce UPOs utilising fast-growing standard laboratory hosts such as bacteria or yeast. These organisms would facilitate protein engineering and allow directed evolution campaigns for tailoring UPOs towards desirable traits such as increased stability, regio- and enantioselectivity.

Although substantial work has been invested into the heterologous expression of the firstly discovered *Agrocybe aegerita* UPO (*Aae*UPO) using the yeast *Saccharomyces cerevisiae*, sufficient protein amounts of 8 mg/L were obtained as the result of an intensive directed evolution campaign[10]. This fundamental work led to several successful UPO application studies on the conversion of a range of substrates from agrochemicals to pharmaceuticals[11–14]. The yeast secretion variant PaDa-I (hereinafter: *Aae*UPO*) was adapted subsequently for large-scale protein production by utilising the methylotrophic yeast *Pichia pastoris* (syn. *Komagataella phaffii*) reaching recombinant protein titres of 217 mg/L within a bioreactor setup[15].

The successful production was achieved by the introduction of nine amino acid exchanges. Four of these were localised within the 43 amino acid signal peptide (SP), which orchestrates protein secretion in the natural fungal host as well as in *S. cerevisiae*. The engineered signal peptide combined with the wild-type *Aae*UPO enzyme resulted in a 27-fold increase in protein secretion yield highlighting the paramount importance of the respective signal peptide for heterologous production as already shown by others[16–21]. Recent studies report the production of UPOs in *E. coli*[22,23]. However, it remains elusive whether these recombinant peroxygenases harbour comparable activities and stabilities if compared to UPOs produced in eukaryotic hosts. The reported expression yields are substantially lower compared to *S. cerevisiae* raising the question, whether enough functional protein could be produced for laboratory evolution campaigns.

Golden Gate cloning has proven to be an invaluable synthetic biology tool enabling seamless assembly of gene fragments utilising type IIs restriction enzymes[24–32]. By using type IIs restriction enzymes, defined 4 base pair sticky overhangs can be created for reassembly. These overhangs can be easily specified by PCR, allowing a sequence defined, efficient and seamless assembly of nine and more gene fragments in a one-pot and one-step digestion-ligation manner[25,32,33].

For the detection of the target protein secretion in small volumetric amounts of yeast supernatant, a sensitive, high-throughput suitable, and protein-specific assay would be highly beneficial. Previously reported split-GFP (green fluorescent protein) systems, which rely on tagging the protein of interest with a short amino acid peptide tag and subsequent GFP reconstitution, present an ideal tool for this task[34,35].

In this study, we envisioned a tripartite Golden Gate-based modular system. This system consists of the modules 'signal peptide', 'UPO gene' and 'protein-tag' (Fig. 1A). The 'protein-tag' module combines affinity-based purification as well as the enzyme quantification by split-GFP. This *S. cerevisiae* expression system gave rise to a rapid workflow starting from UPO genes to heterologously produced and purified UPOs within 2 to 4 weeks (Supplementary Fig. 2).

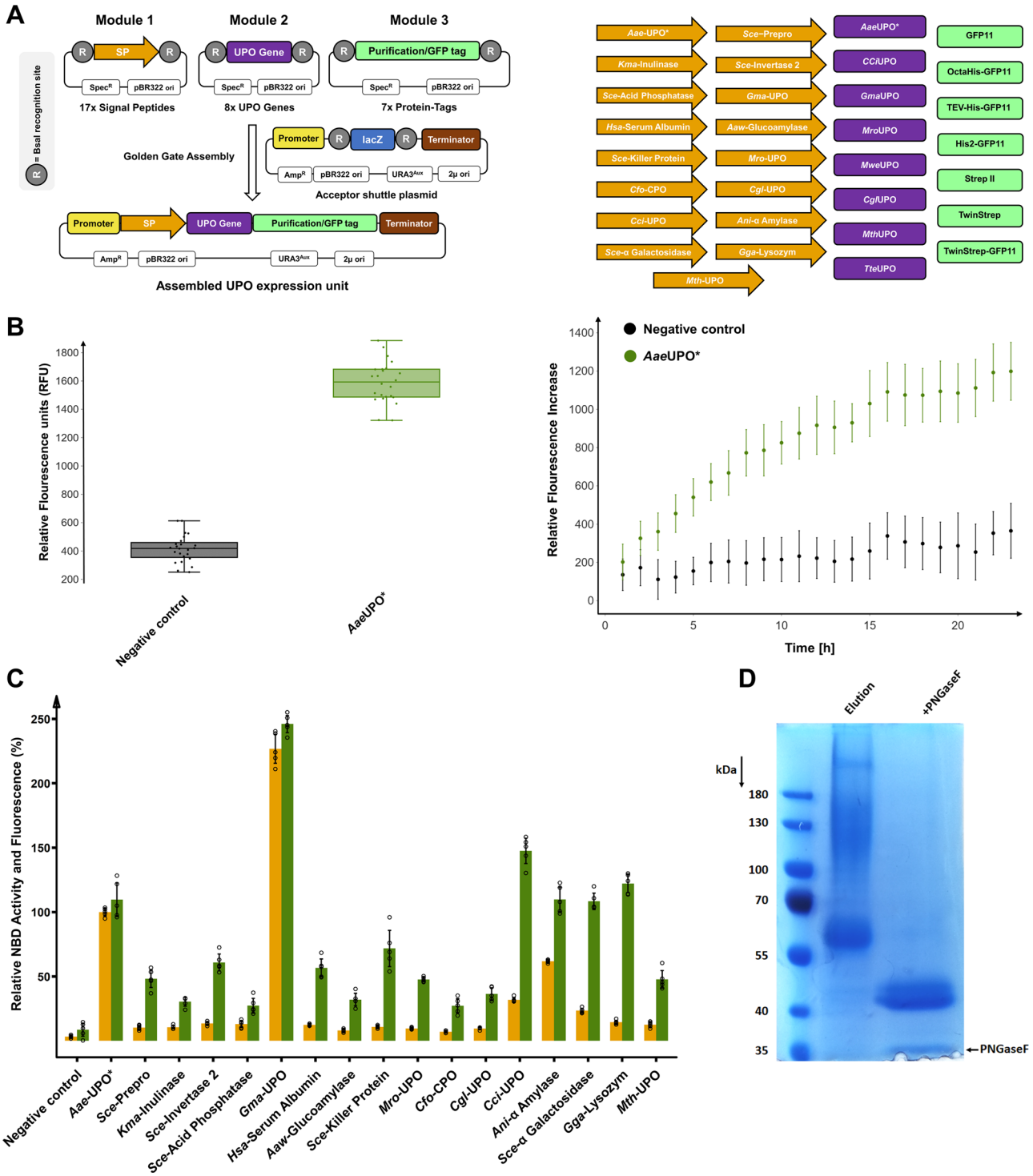
To unlock access to higher protein amounts, we designed two fully compatible episomal and one integrative plasmid for UPO production in the methylotrophic yeast *Pichia pastoris*. In total, four active UPOs were heterologously produced in yeast for the first time. The high recombinant UPO yields using *P. pastoris* enabled the enantioselective hydroxylation of a phenethylamine derivative on a preparative scale.

Results

The modular Golden Gate UPO expression system. Three modules were designed for pre-defined assembly into an episomal *S. cerevisiae* shuttle expression plasmid. We created 32 modules (Fig. 1A) consisting of 17 signal peptides (Module 1), 8 UPO genes (Module 2) and 7 protein-tags (Module 3) derived from a broad phylogenetic background as summarised in Table 1. Module 3 is employed for affinity-based enzyme purification and/or split-GFP-based protein quantification. To verify the envisioned system for protein quantification, the *C*-terminal GFP11 detection tag (Module 3) was assembled with the previously evolved UPO signal peptide *Aae*-UPO* (Module 1) and the engineered peroxygenase *Aae*UPO* (Module 2)[10,36]. Target gene expression is controlled by a GAL 1.3 promoter, which is repressed in the presence of glucose and strongly induced by galactose, being a truncated version of the widely used GAL1 promoter. The successful split-GFP assay was validated by a significant fluorescence response in the sample with the secreted protein (Fig. 1B).

Module 1, exhibiting 17 distinct signal peptides (SP), is the pivotal part for guiding protein secretion. The diverse signal peptide library consists of sequences originating from *S. cerevisiae*, further yeast organisms, basidiomycetes, ascomycetes and animals (Supplementary Table 4). Seven signal peptide sequences originate from (putative) UPOs and a closely related chloroperoxidase (*Cfu*CPO). To demonstrate the importance of the signal peptide, we assembled the *Aae*UPO* gene (Module 2) and the GFP11 tag (Module 3) with each of the 17 signal peptides (Module 1). UPO secretion levels were monitored by enzymatic activity using the 5-nitro-1,3-benzodioxole (NBD)[37] assay as well as split-GFP detection (Fig. 1C).

All constructs showed significant secretion levels and enzymatic activities. The signal peptides *Cci*-UPO, *Ani*-α Amylase, *Sce*-α Galactosidase and *Gga*-Lysozyme led to similar protein concentrations as the evolved signal peptide *Aae*-UPO*. The signal peptide *Gma*-UPO resulted in a more than doubled activity and secretion of the *Aae*UPO* enzyme relative to the evolved *Aae*-UPO* signal peptide (220% increase). This observation is particularly impressive considering that the signal peptide *Aae*-UPO* was evolved for the optimised secretion of *Aae*UPO* in *S. cerevisiae* by subjecting it to several rounds of directed evolution[10]. The signal peptide *Gma*-UPO originates from the



putative *Galerina marginata* UPO (*Gma*UPO). When correlating normalised enzymatic activity and split-GFP-based fluorescence values of the signal peptide library, in most cases, higher fluorescence levels than activity values were measured. This observation indicates the occurrence of differing *Aae*UPO* enzyme variants depending on respective signal peptide cleavage. This could be due to the great diversity of the utilised signal peptides likely resulting in differing *N*-termini and affecting the enzymatic activity of the processed enzyme.

To give rise to a general, one-step protein purification protocol for UPOs, Module 3 was further extended to allow for simultaneous affinity-based protein purification and GFP11-based fluorescence detection. Several versions of the GFP11 tag in combination with Strep®- or Hexa/Octahistidine-affinity tags were generated and tested (Supplementary Table 5)[38,39]. We used the protein tags with the previously identified beneficial combination of signal peptide (*Gma*-UPO, Module 1) and UPO (*Aae*UPO*, Module 2) and identified the TwinStrep-GFP11 protein tag as most suitable. This tag consists of a double 8 amino acid Strep II tag (Twin-Strep®-tag)[40] and a *C*-terminal GFP11 sequence. Comparison of the modules GFP11 and TwinStrep-GFP11 revealed unaltered enzymatic activities but a significantly higher fluorescence response for the TwinStrep-GFP11 construct (Supplementary Fig. 3). This difference is

Fig. 1 The Golden Gate system consisting of the modules signal peptide, UPO gene and protein-tag and its functional verification regarding split-GFP assay, signal peptide shuffling and purification using the model UPO *Aae*UPO* in *S. cerevisiae*. **A** *Left*: Concept of the modular Golden Gate system as a tripartite system, consisting of signal peptide (SP; contains ATG start codon), UPO gene (lacking start and stop codon) and *C*-terminal Tag (contains stop codon). *Right*: Overview of the individual parts of the modular shuffling systems, containing 17 signal peptides, 8 UPO genes and 7 *C*-terminal protein tags. Detailed sequence information of all parts can be found in Supplementary Tables 4 and 5. **B** Quantification of the UPO secretion in *S. cerevisiae* using the split-GFP system. Two constructs were utilised for testing, namely a previously derived yeast secretion variant of *Aae*UPO (*Aae*UPO*) and further including a *C*-terminal GFP11. The acceptor shuttle plasmid (pAGT572_Nemo 2.0) was used as negative control. *Left:* biological replicates ($n = 24$) of *Aae*UPO* and the negative control were screened within the split-GFP assay. Relative fluorescence units (RFU) were measured at 0 and 72 h after adding GFP1-10. Values are shown as boxplots (*Aae*UPO*: median = 1589, s.d. 8.9%; negative control: median = 416, s.d. 22.6%) with individual data points shown as dots. *Right:* Continuous fluorescence measurements (24 h; 23 time points) of each construct. Data are mean of fluorescence − background (background = first measurement after 1 h) ± s.d. of biological replicates ($n = 24$). **C** Screening of the constructed signal peptide shuffling library utilising *Aae*UPO* as reference protein. Values for 5-nitro-1,3-benzodioxole (NBD) conversion (orange bars) and fluorescence by split-GFP assay (green bars) were normalised to the previously used *Aae*UPO SP* -*Aae*UPO* construct (100%). Data are mean ± s.d. of biological replicates ($n = 5$). Primary data are displayed within the Source data file. Detailed information on the origin and the sequence of the signal peptides can be found in Supplementary Table 4. **D** SDS-PAGE analysis of *Aae*UPO* after one step TwinStrep tag® purification, utilising the designed TwinStrep-GFP11 purification/detection combination tag. Additionally, *Aae*UPO* was subjected to enzymatic deglycosylation by PNGaseF and analysed (right lane).

probably due to better accessibility of the terminal GFP11 portion since the overall size of the tag is increased (27 vs. 59 amino acids), and several flexible linkers are included. SDS PAGE analysis revealed the successful one-step purification of the mature protein *Aae*UPO* (Fig. 1D). The positioning of an N-terminal Strep II protein-tag revealed greatly diminished UPO activity (Supplementary Fig. 4) and was therefore not further investigated.

Utilisation of the modular system for the heterologous production of novel UPOs. To demonstrate that the modular system can provide quick access to produced UPOs, we choose seven UPO genes to be expressed in *S. cerevisiae* while three being undescribed putative UPOs. Four UPOs were previously described and produced in their natural hosts—*Marasmius rotula* UPO (*Mro*UPO)[41], *Marasmius wettsteinii* UPO (*Mwe*UPO)[8], *Chaetomium globosum* UPO (*Cgl*UPO)[9]—or heterologously expressed in an *Aspergillus oryzae* strain (*Coprinopsis cinerea* UPO (*Cci*UPO))[42].

Two putative UPO sequences were selected based on sequence alignments and data bank searches using the short-type peroxygenase *Cgl*UPO as a template. Two sequences were retrieved, originating from fungi classified as thermophilic: *Myceliophthora thermophila* (*Mth*UPO) and *Thielavia terrestris* (*Tte*UPO)[43], bearing 72% and 51% sequence identity to *Cgl*UPO, respectively (Supplementary Table 10). The predicted long-type UPO gene *Gma*UPO is derived from the basidiomycete *Galerina marginata* and was selected based on its high sequence identity (71 %) with *Aae*UPO*.

All genes were introduced as modules (Module 2) into the Golden Gate system and subjected to random shuffling utilising all 17 signal peptides (Module 1).

Out of the seven UPO genes, six could be secreted by *S. cerevisiae* in combination with at least two signal peptides (Fig. 2A). *Cci*UPO exhibited no secretion with any of the signal peptides. *Mwe*UPO and *Gma*UPO were identified through the split-GFP assay, but no activity was detected using the colorimetric 2,6-dimethoxyphenol (DMP) assay[12]. *Mwe*UPO, *Mro*UPO and *Cgl*UPO were the only UPOs, which showed the highest activities with their endogenous signal peptides, *Mro*UPO and *Mwe*UPO sharing the identical native signal peptide. *Mth*UPO and *Tte*UPO showed remarkable secretion levels within the microtiter plate setup, leading to 17-fold (*Mth*UPO) and 50-fold (*Tte*UPO) split-GFP signal intensities above background level. A high signal peptide promiscuity was observed for *Mth*UPO and *Tte*UPO with at least 5 and 8 suitable signal peptides, respectively (Supplementary Figs. 5 and 6).

Purification and characterisation of the identified UPOs. All successfully secreted UPOs in combination with their best signal peptides were equipped with the TwinStrep-GFP11 tag, produced in 1 L shake flask scale, and purified by affinity chromatography. The occurrence and primary sequence of each UPO was confirmed by tryptic digest and mass spectrometric peptide analysis (Supplementary Tables 6 and 8). *Aae*UPO* analysis revealed the amino acids 'EPGLPP' being the first detectable residues at the *N*-terminus in accordance with previous results[15]. This finding indicates that the employed signal peptide *Gma*-UPO leads to a comparable cleavage pattern as the evolved *Aae*-UPO* signal peptide. The split-GFP response and the NBD activity also exhibited the same ratio for both signal peptides (Fig. 1C), which further strengthens the point of a similar cleavage pattern. *Mro*UPO and *Mwe*UPO were produced utilising their native signal peptide (*Mro*-UPO SP). Fragments derived from the signal peptide *Mro*-UPO (11 amino acids for *Mro*UPO and 9 amino acids for *Mwe*UPO) were identified by MS analysis, suggesting a different cleavage pattern compared to the natural host[8]. Obtained *N*-termini of *Gma*UPO and *Mth*UPO are in good agreement with the predicted cleavage sites based on alignments with the enzymes *Aae*UPO* and *Cgl*UPO, respectively. The *N*-terminus of *Cgl*UPO could not be resolved. For *Tte*UPO, a peptide fragment of 10 amino acids of the utilised signal peptide (*Sce*-Prepro) was identified.

*Gma*UPO and *Mwe*UPO were not further studied as the purified enzymes did not exhibit any activity towards the colorimetric peroxygenase substrates DMP and NBD. For these enzymes, we were not able to obtain pure elution samples for subsequent measurements of native as well as CO differential absorption spectra.

Biochemical parameters were therefore determined for *Mro*UPO, *Cgl*UPO, *Mth*UPO and *Tte*UPO. UV absorption profiles showed the expected characteristic peroxygenase haem-thiolate features. A Soret band with a maximum around 420 nm (*Mro*UPO: 419 nm; *Cgl*UPO: 418 nm; *Mth*UPO: 420 nm and *Tte*UPO: 419 nm) and two Q-bands in the range of 537 to 546 and 569 to 573 nm (Fig. 2B)[2] were detected. *Cgl*UPO revealed a broader Soret band shape as well as less pronounced Q-bands. The respective carbon monoxide complexes exhibited absorption maxima around 444 nm (Supplementary Fig. 7).

Protein purity and glycosylation were analysed by SDS-PAGE. Native deglycosylation was performed using PNGaseF (Supplementary Fig. 8). All obtained molecular weights after deglycosylation were in approximate agreement with the calculated weight based on the primary sequence and peptide analysis by mass spectrometry. *Mro*UPO exhibited a defined band at approx. 42

Table 1 Origin of utilised UPO genes and signal peptides for target protein secretion.

Descriptor	Type	Organism	Descriptor	Type	Organism
*Aae*UPO*	Secretion engineered UPO	*Agrocybe aegerita*	*Cfo*-CPO	Chloroperoxidase signal peptide	*Caldariomyces fumago*
*Gma*UPO	Wild-type UPO	*Galerina marginata*	*Cci*-UPO	UPO signal peptide	*Coprinopsis cinerea*
*Cci*UPO	Wild-type UPO	*Coprinopsis cinerea*	*Sce*-α Galactosidase	α Galactosidase signal peptide	*Saccharomyces cerevisiae*
*Mro*UPO	Wild-type UPO	*Marasmius rotula*	*Sce*-Prepro	α factor signal peptide	*Saccharomyces cerevisiae*
*Mwe*UPO	Wild-type UPO	*Marasmius wettsteinii*	*Sce*-Invertase 2	Invertase 2 signal peptide	*Saccharomyces cerevisiae*
*Cgl*UPO	Wild-type UPO	*Chaetomium globosum*	*Gma*-UPO	UPO signal peptide	*Galerina marginata*
*Mth*UPO	Wild-type UPO	*Myceliophthora thermophila*	*Aaw*-Glucoamylase	Glucoamylase signal peptide	*Aspergillus awamori*
*Tte*UPO	Wild-type UPO	*Thielavia terrestris*	*Mro*-UPO	UPO signal peptide	*Marasmius rotula*
Aae-UPO*	Engineered signal peptide of *Aae*UPO	*Agrocybe aegerita*	*Cgl-UPO*	UPO signal peptide	*Chaetomium globosum*
Kma-Inulinase	Inulinase signal peptide	*Kluyveromyces marxianus*	*Ani*-α Amylase	α Amylase signal peptide	*Aspergillus niger*
Sce-Acid Phosphatase	Acid phosphatase signal peptide	*Saccharomyces cerevisiae*	*Gga*-Lysozym	Lysozyme signal peptide	*Gallus gallus*
Hsa-Serum Albumin	Serum albumin signal peptide	*Homo sapiens*	*Mth*-UPO	UPO signal peptide	*Myceliophthora thermophila*
Sce-Killer Protein	Killer protein signal peptide	*Saccharomyces cerevisiae*			

kDa that was slightly shifted towards lower molecular weight after deglycosylation. *Cgl*UPO revealed a smeared band in the range of 55–130 kDa. Deglycosylation led to the occurrence of two distinct protein bands of approx. 37 and 33 kDa indicating different protein subtypes. *Mth*UPO and *Tte*UPO showed an intensive smeared band in the range of 55–200 kDa. This smear was converted into distinct protein bands upon deglycosylation with approx. 38 kDa and 36 kDa for *Mth*UPO and *Tte*UPO, respectively.

To gain insights into the impact of the glycosylation on the activity of the respective enzyme, UPOs were deglycosylated in the native state and assessed for their activity towards NBD (Supplementary Fig. 9). The enzymatic activity of *Mro*UPO was in comparison least affected with a decrease of approx. 80%, whereas in the case of *Cgl*UPO no activity could be obtained after deglycosylation. The activity was substantially impaired as well for *Tte*UPO and *Mth*UPO, leading to a complete loss and approx. 85% decrease, respectively, in enzymatic activity.

We next evaluated the pH-dependencies of the enzymes using NBD as a substrate (Fig. 2C). *Mro*UPO, *Mth*UPO and *Tte*UPO exhibited a similar profile with maximum activity at slightly acidic conditions (pH 5), whereas *Cgl*UPO's activity optimum was detected at pH 7. *Tte*UPO showed a broader tolerance towards lower pH values, retaining medium (pH 3; 40%) and high activity (pH 4.0; 80%) at acidic conditions. The obtained values for *Mro*UPO and *Cgl*UPO are in good agreement with previous data obtained with homologously produced enzyme[9,41].

Enzymatic epoxidation and hydroxylation experiments. The heterologously produced UPOs were tested towards their substrate specificity and activities by investigating three distinct reaction types: aromatic hydroxylation (sp^2-carbon), alkene epoxidation and the benzylic hydroxylation (sp^3-carbon) of phenylalkanes with varying alkyl chain lengths from two to five carbons (Fig. 3). All reactions were performed under the same conditions and assessed for the achieved turnover number (TON) within one hour. Substantially differing behaviour could be observed between *Aae*UPO* and the novel heterologously produced UPOs regarding substrate conversion, specific product formation and stereoselectivity. *Aae*UPO* proved to be the only enzyme displaying a high specificity for single hydroxylation of naphthalene leading to 1-naphthol (92% of the formed product, Fig. 3A). The other UPOs exhibited a strong tendency for further oxidation leading to the dione product 1,4-naphthoquinone. The epoxidation of styrene (Fig. 3B) was efficiently catalysed by *Aae*UPO* (4580 TON) in combination with a poor stereoselectivity (2% *ee*). *Cgl*UPO exhibited comparable epoxidation activities (4110 TON) and an enantioselectivity of 44% *ee*. For *Mth*UPO, TON decreased to 1100 but revealed the highest stereoselectivity (45% *ee*). The studies of the benzylic hydroxylation of phenylalkanes—ranging from phenylethane to phenylpentane—confirmed the known preference of *Aae*UPO* towards short alkane chain length (Fig. 3C)[3]. Starting from 4500 turnovers for the conversion of phenylethane and deteriorating to no product formation and only traces of benzylic hydroxylation using phenylbutane and phenylpentane, respectively (for other product formations see Supplementary Fig. 17). *Cgl*UPO and *Mth*UPO exhibited an inverted trend with increasing product formations for longer alkyl chain lengths, exhibiting the lowest activity for the phenylethane hydroxylation.

The highest activity was detected in both cases using phenylbutane (*Cgl*UPO: 1670 TON, *Mth*UPO: 1490 TON) with only slightly decreased activity for phenylpentane as a substrate and the only significant side-product being the further oxidation of the benzylic alcohol to the corresponding ketone

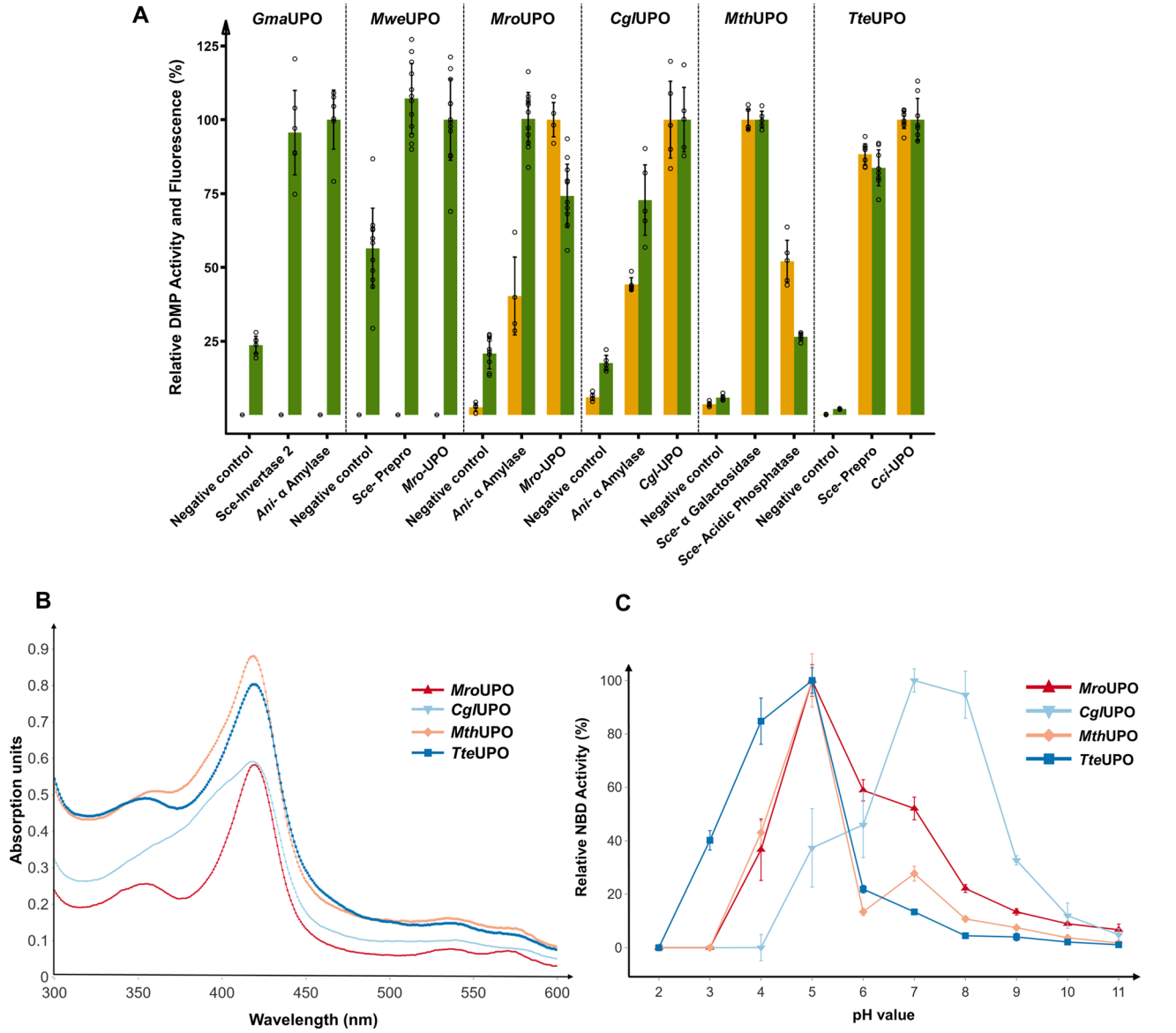


Fig. 2 Through signal peptide shuffling identified novel UPO construct and their analysis of UV absorption spectra and pH profiles. A Golden Gate signal peptide shuffling was applied for the testing of described and putative UPO genes, and the two best signal peptide/UPO gene combinations are displayed. *Gma*UPO, *Mwe*UPO, *Mro*UPO and *Cgl*UPO were screened in combination with a GFP11-tag. *Mth*UPO and *Tte*UPO were screened using the TwinStrep-GFP11 protein tag. UPO enzyme activity was determined by monitoring the conversion of 2,6-dimethoxyphenol (DMP) to coerulignone. The highest average fluorescence (split-GFP) and conversion values (DMP) within one enzyme panel were set to 100%, and the other values normalised accordingly. Data are mean ± s.d. of biological replicates ($n \geq 4$). Corresponding primary data are displayed within the Source data file. **B** UV-Vis absorption spectra of the purified peroxygenases *Mro*UPO, *Cgl*UPO, *Mth*UPO and *Tte*UPO in the wavelength range between 300 and 600 nm (measurement interval: 1 nm). **C** pH profiles of *Mro*UPO, *Cgl*UPO, *Mth*UPO and *Tte*UPO catalysed enzymatic conversion of 5-nitro-1,3-benzodioxole (NBD) to 4-nitrocatechol. The highest mean activity of a respective enzyme was set to 100% and the other values normalised accordingly. Data are means ± s.d. of measurements performed in triplicates. Corresponding primary data are displayed within the Source data file.

(Supplementary Fig. 17). *Tte*UPO showed a similar preference towards long-chain phenylalkanes with the highest TON for phenylpentane conversion (500 TON). *Tte*UPO represented the only UPO with a significant specificity towards the formation of the *S* alcohol enantiomer for phenylpropane and phenylbutane. For phenylpentane, it revealed the formation of the opposite alcohol enantiomer than the other tested UPOs as well.

Expanding the modular UPO secretion system to *Pichia pastoris*. The methylotrophic yeast *P. pastoris* (syn. *Komagataella phaffii*) constitutes an attractive heterologous production host with a steadily growing toolbox of valuable synthetic biology parts such as plasmids, promoters and signal peptides[44,45]. *P. pastoris* can reach high cell densities, efficiently perform post-translational modifications such as glycosylation and disulfide-linkage and offers a set of strong and tightly regulated promoters for target gene expression. Amongst other factors, these properties render *P. pastoris* a widely used eukaryotic host for the large-scale industrial production of therapeutic proteins and enzymes[46]. We investigated the adaptation of the modular system for use in *P. pastoris*. Therefore, two novel episomal *P. pastoris* expression plasmids were designed and assembled. They contain a previously described autonomously replicating sequence coined panARS, which confers episomal stability and a hygromycin B marker gene

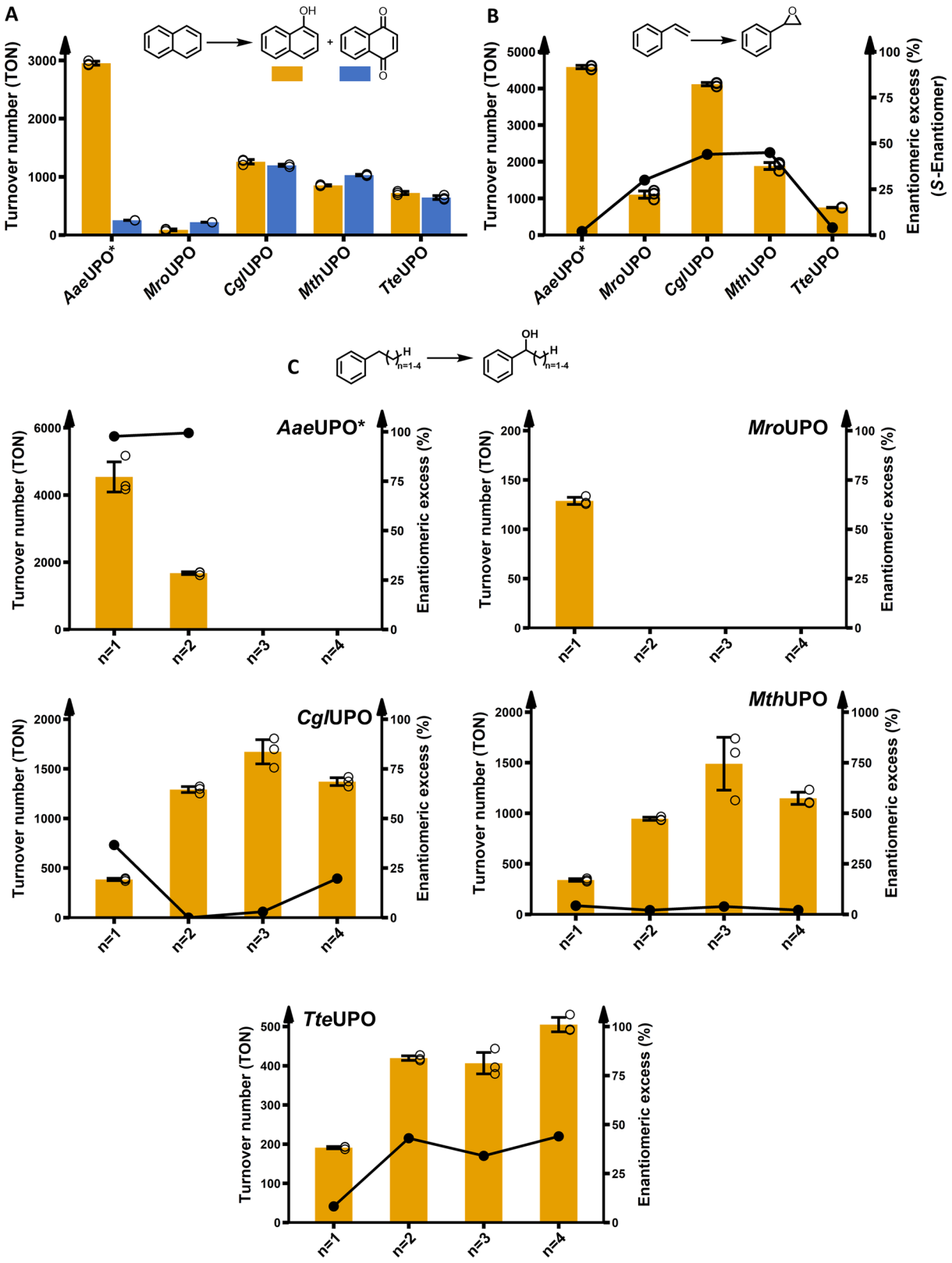
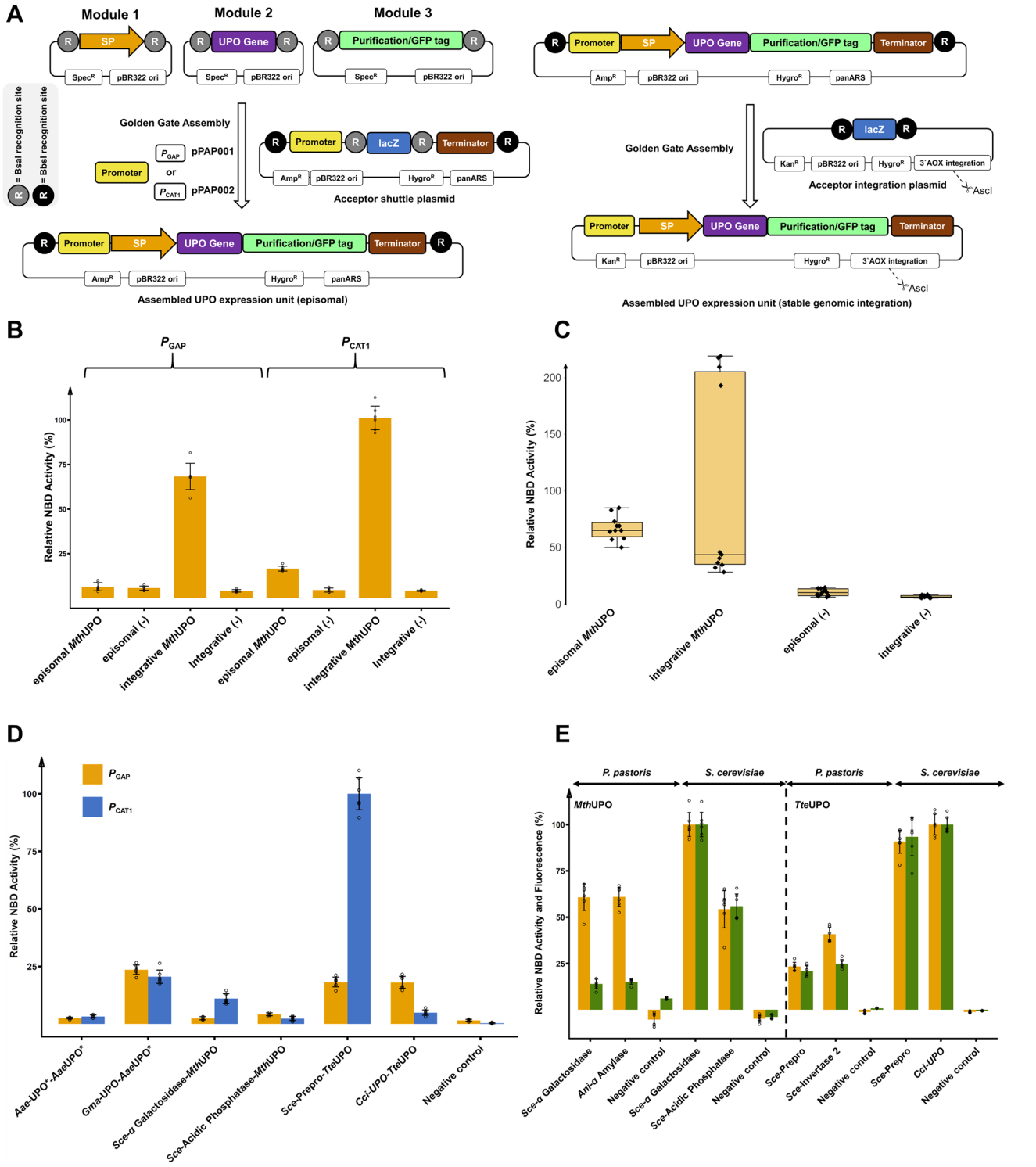


Fig. 3 Enzymatic activity assessment of the peroxygenases regarding aromatic hydroxylation, epoxidation and sp^3-carbon hydroxylation. All reactions were performed for 1 h using 1 mM of substrate. Bar charts display the obtained turnover numbers (TON) within one hour. The lines correspond to the enantiomeric excess %. Data are mean ± s.d. measurements derived from biological triplicates with individual data points shown as circles. See supplementary information for further details. **A** Conversion of naphthalene to naphthol and 1,4-naphthoquinone. **B** Conversion of styrene to styrene oxide. **C** A homologous row of phenylethane, phenylpropane, phenylbutane and phenylpentane hydroxylation, respectively, focusing on hydroxylation of the benzylic carbon. The alcohol enantiomer is indicated by an (R) or (S). The exact enantiomer for phenylpentane was not determined. See Supplementary Fig. 17 for occurrence of side-products. For *Mro*UPO catalysed conversion of phenylethane no enantioselectivity could be determined.



for antibiotic selection[47–49]. The constructed episomal plasmids differ by the employed promoter: the strong constitutive glyceraldehyde-3-phosphate dehydrogenase promoter (P_{GAP}, plasmid pPAP001) and the recently described strong methanol inducible catalase 1 promoter (P_{CAT1}, pPAP002)[50]. The plasmids were constructed to allow direct implementation of the tripartite modular UPO secretion system, consisting of Module 1 (signal peptide), Module 2 (UPO gene) and Module 3 (*C*-terminal tag, Fig. 4A; left). To further allow the genomic integration to generate stable *P. pastoris* cell lines for antibiotic-free large-scale enzyme production in shake flasks or fermenters, a third plasmid (pPAP003) was constructed. The episomal plasmids are designed to enable direct transfer of the identified best transcription unit (promoter-signal peptide-gene-tag-terminator) combination to the integration plasmid. This transfer requires only an additional Golden Gate assembly reaction using the restriction enzyme BbsI (Fig. 4A; right).

We tested all *P. pastoris* plasmids using the newly discovered peroxygenase *Mth*UPO in combination with the *Sce*-α Galactosidase signal peptide. The constructs proved to be functional and led to an NBD conversion signal (Fig. 4B). P_{GAP}-based secretion was generally lower in comparison to P_{CAT1}, and the episomal

Fig. 4 The compatible modular Golden Gate setup utilising episomal and integrative *P. pastoris* plasmids and its application. A *Left*: Overview of the designed episomal *P. pastoris* screening setup. All previously created basic modules are compatible to be used within this system. Two episomal plasmids were designed harbouring the constitutive strong promoter P_{GAP} or the strong inducible promoter P_{CAT1}. *Right:* Identified gene constructs can be directly transferred in a one-pot Golden Gate reaction (+BbsI) from the episomal plasmid to an integrative plasmid. After linearisation (AscI digest) this plasmid can be integrated into the genomic 3'AOX region of *P. pastoris*. **B** Comparison of relative activities of 5-nitro-1,3-benzodioxole (NBD) conversion for different *P. pastoris* constructs bearing the tripartite combination of α Galactosidase signal peptide-*Mth*UPO-TwinStrep-GFP11. P_{GAP} bearing constructs were screened utilising Glucose (1.5 % (w/v)) as sole carbon source. P_{CAT1} bearing constructs were screened utilising a dual feeding strategy (0.5% (v/v) glycerol and 1.5% (v/v) methanol) as primary and inducible carbon sources. The highest expression mean is set to 100% and all data normalised. Data are mean ± s.d. of biological replicates ($n = 6$) originating from streak outs of one previously screened colony of the respective construct. **C** Comparison of relative activities of NBD conversion of P_{CAT1}-based constructs bearing the tripartite combination of α Galactosidase signal peptide-*Mth*UPO-TwinStrep-GFP11. Box plots of biological replicates ($n = 11$) of individual *P. pastoris* colonies for each construct. The highest expression mean is set to 100% and all data normalised (episomal *Mth*UPO: median = 65, s.d. 15.0%; integrative *Mth*UPO: median = 44, s.d. 83.3%; episomal (-): median = 10, s.d. 28.2%; integrative (-): median = 6, s.d. 16.4%). **D** Comparison of relative activities of NBD conversion for different episomal *P. pastoris* constructs (6 biological replicates each) using the indicated signal peptide-UPO combinations as well as a TwinStrep-GFP11 tag. P_{GAP} (yellow bars) and P_{CAT1} (blue bars). The highest expression is set to 100%, and all data are normalised accordingly. Data are mean ± of biological replicates ($n = 6$). **E** Direct comparison of episomal UPO production of the two best signal peptide-UPO combinations for *Mth*UPO and *Tte*UPO as identified by a previously performed signal peptide shuffling approach in both yeast species. Episomal *P. pastoris* expressions utilising P_{CAT1}. The highest mean expression and activity for each enzyme is set to 100%, and all data are normalised. Data are mean ± s.d. of biological replicates ($n = 6$). NBD conversion activity (orange) and relative fluorescence units (green). All primary data are displayed within the Source data file.

P_{GAP} *Mth*UPO activity was not distinguishable from the negative control. The integrative plasmids outperformed their episomal counterparts significantly with a factor of approx. 5 for P_{CAT1}. A similar observation however in a varying degree was made testing the enzymes *Aae*UPO* and *Tte*UPO (Supplementary Fig. 10), indicating that the integrative constructs are in general promoting higher UPO secretion levels than their episomal counterparts.

To gain insights into interclonal variabilities of UPO secretion, episomal and integrative plasmids were transformed into *P. pastoris*. Individual colonies were cultivated and tested for UPO secretion. The episomal construct showed diminished mean activity but a substantially lower clonal variability than the integrative plasmid when tested with NBD (Fig. 4C). This high variability of the secretion level for the integrative plasmid is presumably due to divergent numbers of copy insertions into the *P. pastoris* genome, which is a commonly occurring feature in this organism[51–53], and might also lead based on the Hygromycin B selection marker to different colony sizes (Supplementary Fig. 11).

To investigate and compare the secretion levels of episomal P_{GAP} and P_{CAT1} harbouring plasmids, twelve constructs were generated harbouring the peroxygenases *Aae*UPO*, *Mth*UPO and *Tte*UPO. All promoter combinations (P_{GAP} and P_{CAT1}) and the two previously identified signal peptides were constructed in combination with the respective UPO gene and analysed for NBD activity. All constructs resulted in a significant NBD conversion (Fig. 4D). The previously observed 220% improved *Aae*UPO* secretion in *S. cerevisiae* by combining *Aae*UPO* with the signal peptide *Gma*-UPO was found to be even more pronounced using the episomal *P. pastoris* constructs (P_{CAT1}: 620%). Besides the striking influence of the promoter on the secretion level, also t he combination of the signal peptide and the promoter proved to be pivotal. For *Tte*UPO, using the promoter P_{CAT1} in combination with the *Sce*-Prepro signal peptide led to the highest detected activity with a 20-fold higher signal compared to the *Cci*-UPO signal peptide. The same signal peptide variations employing the P_{GAP} promoter, however, resulted in similar secretion levels. This demonstrates besides the crucial role of the chosen signal peptide (Figs. 1C and 2A) an additionally pivotal synergistic influence of the promoter/signal peptide combination on the UPO secretion.

To gain insights into the different signal peptide preferences for secretion in *P. pastoris*, the signal peptide shuffling approach was repeated in *P. pastoris* using *Mth*UPO and *Tte*UPO and choosing the episomal P_{CAT1} bearing plasmid (Supplementary Figs. 12 and 13). For *Mth*UPO the signal peptides *Sce*-α Galactosidase and *Ani*-α Amylase proved to be most suitable, and *Sce*-Prepro and *Sce*-Invertase 2 were identified as top hits for *Tte*UPO. Interestingly, *Sce*-Invertase 2 has not been identified amongst the top hits in *S. cerevisiae* whereas the best signal peptide (*Cci*-UPO) for secretion in *S. cerevisiae* (Fig. 4D) was not identified in the *P. pastoris* screen.

To compare episomal *S. cerevisiae* and *P. pastoris* secretion, the two best performing constructs for *Mth*UPO and *Tte*UPO were selected. This species comparison (Fig. 4E) indicates that the episomal *S. cerevisiae* secretion is superior to the episomal *P. pastoris* production. In the case of *Mth*UPO, both *P. pastoris* constructs led to approx. 60% of NBD conversion in comparison to the most suitable *S. cerevisiae* construct, while already exhibiting higher NBD conversion rates than the second most suitable signal peptide for secretion in *S. cerevisiae* (*Sce*-Acidic Phosphatase). The split-GFP fluorescence assay revealed a diminished response for the *P. pastoris* setup relative to the *S. cerevisiae* constructs. Regarding *Tte*UPO, the best *P. pastoris* construct (*Sce*-Invertase 2) led to approx. 40% of relative NBD conversion when compared to the best *S. cerevisiae* construct (*Cci*-UPO). For *Tte*UPO the split GFP assay followed a linear pattern when comparing species, without revealing a diminished response for *P. pastoris*.

Comparison of shake flask UPO production in *P. pastoris* and *S. cerevisiae*. By using the constructed integrative plasmid pPAP003 and the P_{CAT1} promoter, stable *P. pastoris* cell lines were constructed to produce five UPOs: *Aae*UPO*, *Mro*UPO, *Cgl*UPO, *Mth*UPO and *Tte*UPO (Fig. 5A). Utilising *P. pastoris* as host led to substantially higher production titres in all cases, except for *Tte*UPO. The rather low yields of *Mro*UPO and *Cgl*UPO produced in *S. cerevisiae* could be increased substantially when using *P. pastoris* (*Mro*UPO: 3-fold, *Cgl*UPO: 15-fold). The *Mth*UPO production yield was improved 5-fold. Regarding *Tte*UPO, the product titre was decreased in *P. pastoris* by approx. 20%, however, still maintaining an overall high yield. The production titres of *S. cerevisiae* derived *Tte*UPO (17 mg/L), and *P. pastoris* derived *Mth*UPO (24 mg/L) are the highest yields for shake flask cultivation of recombinant fungal peroxygenases reported thus far. The transfer of the expression system to a fed-batch bioreactor might further elevate protein titres due to the higher cell densities achievable. In previous work, this transfer into a bioreactor resulted in 27-fold improved product titre of

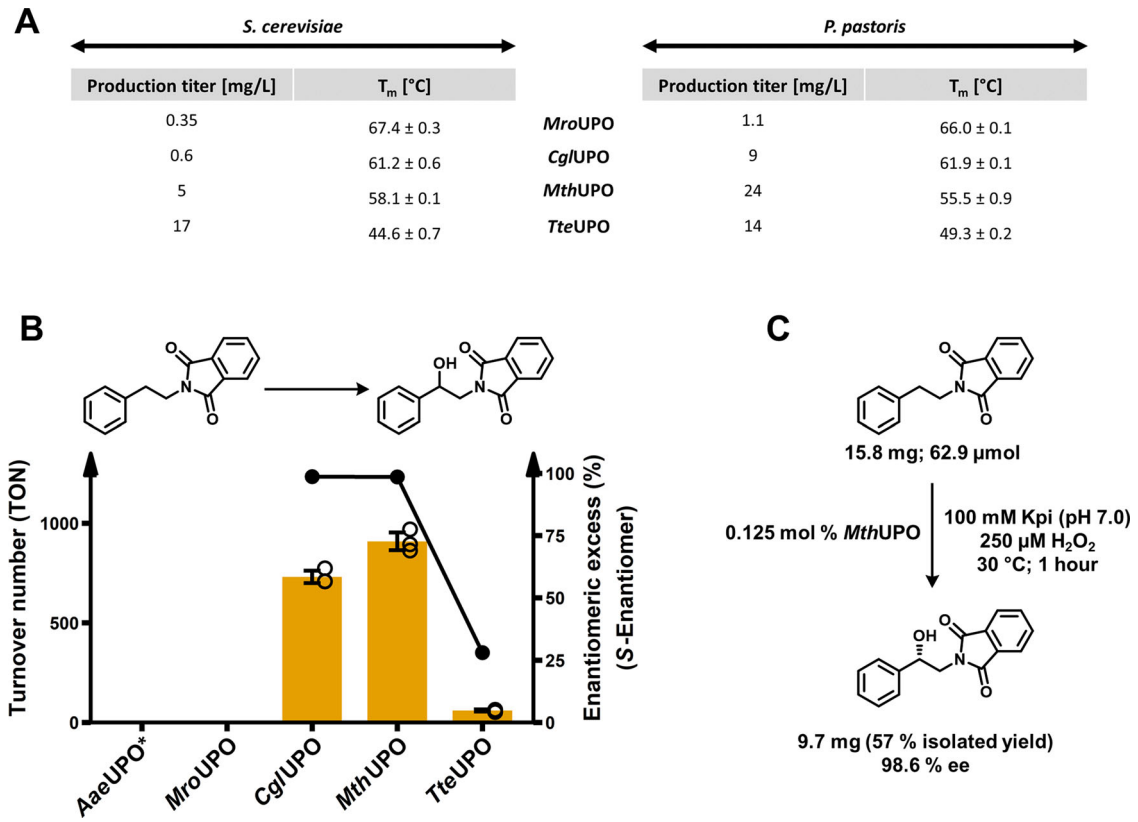


Fig. 5 Expression yields and thermostabilities of UPOs derived from the different yeast systems and conversion of a phenethylamine derivative on analytical and preparative scale. A Comparison of volumetric production titre of recombinant UPOs in shake flask scale (1 L) between *S. cerevisiae* (episomal construct) and *P. pastoris* (integrative construct) as obtained after ultrafiltration of the respective culture supernatant. UPOs were produced and secreted utilising their natural signal peptide (*Mro*UPO and *Cgl*UPO) or a previously identified suitable exogenous signal peptide *Mth*UPO (*Sce*- α Galactosidase) and TteUPO (*S. cerevisiae*: *Cci*-UPO; *P. pastoris*: *Sce*-Prepro). For all *P. pastoris* production setups the methanol inducible promoter P_{CAT1} was utilised. Thermal denaturation midpoints (T_m) for the four UPOs produced in both organisms were determined in biological triplicates using purified protein samples (in 100 mM potassium phosphate; pH 7.0) using differential scanning fluorimetry (DSF). **B** Bar chart showing turnover number within one hour for the benzylic hydroxylation of *N*-phthaloyl-phenethylamine by *P. pastoris* produced *Aae*UPO*, *Mro*UPO, *Cgl*UPO, *Mth*UPO and *Tte*UPO. Turnover data are mean ± s.d. of measurements made in triplicates. TON determined by GC-MS and *ee*% by chiral HPLC (Supplementary Figs. 18–20). **C** Preparative scale conversion of *N*-phthalimide protected phenethylamine using *P. pastoris* produced *Mth*UPO.

217 mg/L for *Aae*UPO*[15]. All proteins were purified using the attached TwinStrep-tag and analysed by SDS-PAGE (Supplementary Fig. 14). Highly pure enzyme preparations were obtained after one-step TwinStrep purification. Based on the successful production in both organisms, thermostability values (denaturation midpoint; T_m) of the four UPOs were assessed using differential scanning fluorimetry (Fig. 5A). The obtained values of the respective UPOs derived from both organisms proved to be alike to a variation of 0.7 to 4.7 °C. The highest thermostability values were obtained for *Mro*UPO with 67.4 and 66.0 °C produced by *S. cerevisiae* and *P. pastoris*, respectively. The two UPOs derived from thermophilic fungi, *Mth*UPO and *Tte*UPO, exhibited no superior thermostability when compared to the closest related enzyme *Cgl*UPO. *Tte*UPO revealed the lowest thermostability in the tested group with 44.6 and 49.3 °C for *S. cerevisiae* and *P. pastoris*, respectively.

Enantioselective hydroxylation of an *N*-protected phenethylamine on a preparative scale. To gain insights into the ability of the enzymes to convert industrially relevant molecules in an enantioselective manner, we selected *N*-protected phenethylamine as a substrate. The hydroxylation of phenethylamine derivatives at the benzylic position provides access to a plethora of pharmaceutically important classes like beta-blockers and sympathomimetics[54].

The peroxygenases *Aae*UPO*, *Mro*UPO, *Cgl*UPO, *Mth*UPO and *Tte*UPO were produced in *P. pastoris*, purified, and assessed for their activity on *N*-phthaloyl-phenethylamine. *Aae*UPO* and *Mro*UPO exhibited no formation of the benzylic alcohol product, and *Tte*UPO performed 60 TON within an hour while achieving an enantioselectivity of 28% *ee* (Fig. 5B, Supplementary Figs. 18 and 20). *Cgl*UPO and *Mth*UPO revealed the highest activities with 730 and 908 TON, respectively, within the 1 h reaction setup (Supplementary Fig. 18). *Mth*UPO showed an over-oxidation to the ketone amounting to 222 TON (Supplementary Fig. 19). The enantioselectivity of alcohol formation proved to be excellent for *Cgl*UPO and *Mth*UPO with 98.7% *ee* and 98.6% *ee* (Supplementary Fig. 20).

Harnessing the high production titre of *Mth*UPO in *P. pastoris* (24 mg/L) in combination with the previously observed good substrate conversion and high enantioselectivity we aimed for the proof-of-concept enantioselective synthesis of (*S*)-(+)-2-*N*-Phthaloyl-1-phenylethanol on a preparative scale (Fig. 5C). In a first upscaling reaction (300 mL total volume) 0.125 mol% of *Mth*UPO derived without further purification from concentrated *P. pastoris* supernatant were used as catalyst loading. The upscaled reaction (30 °C; 1 h) led to the synthesis of 9.70 mg (S)-(+)-2-*N*-Phthaloyl-1-phenylethanol (57% purified yield) and an enantiomeric excess of 98.6% (Supplementary Fig. 20).

Discussion

Fungal unspecific peroxygenases (UPOs) have gained substantial interest as versatile hydroxylation catalyst since their initial discovery 16 years ago[2]. The most significant limitation for the wider application of UPOs arguably remains their heterologous production utilising a fast-growing host. Thus far, only one UPO could be produced and engineered within a system amenable to high throughput: the *S. cerevisiae* secretion variant *Aae*UPO*[10].

Building on the therein developed expression setup, we started our endeavour to construct a versatile UPO secretion system. The constructed Golden Gate-based platform consists of a signal peptide library (Module 1), UPO genes (Module 2) and protein-tags (Module 3). This format enabled the first report of successful yeast secretion of six UPOs—two of them (*Mth*UPO and *Tte*UPO) derived from genome and secretome data had not been characterised as UPOs before[43]. The whole expression platform could be subsequently transferred to *P. pastoris*, resulting in excellent UPO expression yields allowing for preparative scale hydroxylation reactions.

Since the only enzyme out of the panel that could not be produced (*Cci*UPO) belongs to the group of long-type UPOs, and it previously took considerable effort to engineer the long-type UPO *Aae*UPO towards secretion in yeast, one could argue that the heterologous production of long-type UPOs seems to be more challenging[10]. In fact, *Mro*UPO, *Cgl*UPO, *Mth*UPO and *Tte*UPO, which could be initially produced as non-engineered wild-type enzymes in yeast and characterised within the scope of this work, all belong to the class of short-type UPOs. Recent work in our laboratory suggests that gene shuffling of long-type UPOs can offer a viable option to obtain a library of active and structurally diverse long-type UPOs expanding the panel of available recombinant peroxygenases[55].

The hypothesised pivotal role of the employed signal peptides for the successful secretion of UPOs was manifested within this study. Yeast secretion signal peptides generally consist of three main parts: (1) an *N*-terminal positively charged domain, (2) a hydrophobic core and (3) a cleavage site region exhibiting mostly uncharged residues[56]. The length and amino acid composition of the signal peptides, however, can vary substantially. This variation is also reflected by the selected, diverse signal peptide pool in the present paper, which ranges from several predicted and reported UPO signal peptides, commonly utilised endogenous *S. cerevisiae* signal peptides up to human (*Hsa*-Serum Albumin) and bird-derived (*Gga*-Lysozym) sequences. This set includes peptides ranging from 17 to 87 amino acids in lengths and exhibits substantial sequence alterations to provide a high diversity. Apart from the UPO-derived signal peptides, this signal peptide pool was selected based on previous reports on recombinant protein secretion in *S. cerevisiae*[56,57]. In case of *Aae*UPO* the native signal peptide was evolved by means of MORPHING[36] resulting in a 27-fold increased total secretion through addition of the four mutations F12Y/A14V/R15G/A21D[10].

Exchanging this evolved signal peptide sequence with the corresponding *Gma*-UPO sequence resulted in a further 2.2-fold improved secretion in *S. cerevisiae* and even 6-fold enhancement in *P. pastoris*. The signal peptides *Aae*-UPO* and *Gma*-UPO display a high sequence identity and similarity of 50 and 73%, respectively. *Gma*-UPO exhibits two additional amino acids, and these insertions are the two hydrophobic amino acids alanine and leucine before position 12. These amino acid additions might be pivotal for the increased secretion level. Interestingly, the overall suitability of other signal peptides based on specific enzyme activity proved to be low (Fig. 1C), even though in many cases comparable split-GFP values were obtained when comparing respective signal peptides to the evolved *Aae*-UPO* signal peptide. This low promiscuity towards functional signal peptides points towards the occurrence of differing *N*-termini of the mature enzyme. A hypothesis, that has been indirectly stated before in the context of this evolved enzyme[58] as introduced mutations within the mature protein (V57A and V75I) are argued to preserve the natural *N*-terminus (EPGLPPPGPL) of the mature protein when produced in yeast in connection with the evolved signal peptide. This is contrary to the wild type *Aae*UPO enzyme, where *N*-terminal proteolysis (EPG↓LPPPGPL) occurs[59]. Using the *Gma*-UPO signal peptide for production, we could verify the natural *N*-terminus occurrence (starting AEPGLPP) by peptide analysis (Supplementary Table 8); therefore the cleavage pattern of this signal peptide in yeast seems to be comparable to the evolved signal peptide used in previous studies[10,13,15,58,60]. The hypothesis of low promiscuity also holds in the episomal production attempts in *P. pastoris*, where all identified active constructs exhibited the signal peptide *Aae*-UPO* or *Gma*-UPO, respectively.

The case of *Aae*-UPO* might suggest that the closest related signal peptide in terms of sequence and length demonstrates the highest secretion rates. Regarding *Mth*UPO this only might be true in terms of sequence length. The wild-type signal peptide consists of 17 amino acids, and the most suitable orthologous SPs have a similar length: *Sce*-α Galactosidase (19 aa), *Ani*-α Amylase (20 aa) and *Sce*-Acid Phosphatase (18 aa). However, the sequence similarity of the most suitable signal peptides *Sce*-α Galactosidase and *Ani*-α Amylase is solely 3 and 29%, respectively. In general, a high promiscuity of *Mth*UPO towards a diverse set of signal peptides was observed, leading to the identification of five (*S. cerevisiae*) and eight (*P. pastoris*) suitable signal peptides (Supplemental Figs. 5 and 12). In case of *Tte*UPO, no wild-type signal peptide was retrieved, and this UPO proved to be highly promiscuous, showing no preference for a signal peptide length (17–87 amino acids) or sequence composition whatsoever. Out of the pool of 17 signal peptides, nine proved to be highly suitable (Supplementary Figs. 6 and 13), when probing *S. cerevisiae* and *P. pastoris*.

The subsequent transfer of beneficial signal peptide-gene combination to *Pichia pastoris* proved successful. Nonetheless, subtle differences and preferences were shown by a signal peptide shuffling approach for *Mth*UPO and *Tte*UPO in *P. pastoris*. Interestingly, the α-factor leader signal peptide (*Sce*-Prepro), which is often used as a gold standard signal peptide for target protein secretion in *P. pastoris*[45,46], was only identified among the top hits in combination with *Tte*UPO.

In summary almost all the 17 tested signal peptides proved to be highly functional in combination with at least one UPO gene. However, predicting a suitable signal peptide for recombinant protein secretion remains challenging. Previous work reported on the engineering of improved signal peptides by means of directed evolution to produce UPOs[10], laccases[19], aryl-alcohol oxidases[21] and single-chain antibodies[18] in *S. cerevisiae*. These improvements seem to be highly depended on the attached protein and are therefore not per se applicable to other non-related proteins. Novel approaches such as machine learning-based design of signal peptides might help to rationalise the use of SPs, but still need to be transferred to eukaryotic systems[61]. We therefore decided to build a rather diverse signal peptide panel, which can be rapidly assembled using the modular Golden Gate system and tested in a high-throughput manner in a 96-well plate setup. By employing our envisioned modular signal peptide shuffling system, we were able to further improve the production of previously described UPOs (*Aae*UPO*) and obtain multiple suitable signal peptide-gene combinations to produce novel wild-type UPOs (*Mth*UPO and *Tte*UPO).

The GFP11 detection tag proved to be an indispensable asset to distinguish secretion from activity[35]. Between different UPOs, the variation in fluorescence could be further pronounced based on

different accessibilities of the split-GFP-tag. This tendency was shown for *Aae*UPO* where the TwinStrep-GFP11 tag (59 amino acids) yielded a 4-fold increased signal intensity relative to the shorter GFP11 tag (27 amino acids). In some cases, like *Cgl*UPO (Fig. 2A), the fluorescence response greatly differed from the activity depending on the employed signal peptide. This observation might be explained by different cleavage patterns at the N-terminus depending on the respective signal peptide leading to slightly altered overall folds and structures and hence activities of the mature enzyme. In the case of *Cgl*UPO, this hypothesis was strengthened by the occurrence of multiple SDS-Gel bands after enzymatic deglycosylation suggesting multiple patterns of signal peptide cleavages (Supplementary Fig. 8).

In the case of *Aae*UPO* the substrate entrance is among other motifs substantially shaped by the C-terminal helix/loop region, which also contains a crucial stabilising C278-C319 disulfide linkage[58]. Therefore, the attachment of a protein tag to the C-terminus, which we performed in all setups based on the modular design, might be detrimental to the activity. Indeed, using an *Aae*UPO* construct without the C-terminal tag resulted in an improved UPO activity by approx. 40% (Supplementary Fig. 4). The successful placement of an N-terminal tag rather than a C-terminal-modification, however, is challenging. This difficulty is due to the varying cleavage patterns at the N-terminus of the secreted proteins (see above) and also an extra Golden Gate module would be required (signal peptide - N-tag - gene) to preserve the compatibility of the system. Testing an N-terminal Strep II tag resulted in a nearly complete loss of activity (−95%) within the 96 well screening setup (Supplementary Fig. 4). Although a decreased activity is detrimental for the discovery and engineering of enzymes, the observed loss (−40%) when attaching a C-terminal tag is still tolerable, as the split-GFP signal would even in case of a loss of activity provide the signal of successful secretion and allow further enzyme characterisation. To overcome possible limitations, we constructed an additional C-terminal tag module (pAGM9121_TEV-His-GFP11), which includes a TEV protease cleavage site located in front of the His_6-GFP11 detection and purification tag, thereby enabling the removal of the C-terminal appendix after purification and prior to activity measurements.

The utilisation of the GFP11-tag also led to the discovery and verification by subsequent peptide analysis of the peroxygenases *Gma*UPO and *Mwe*UPO—even though no enzymatic activity could be determined. Repeated shake flask production attempts in *S. cerevisiae* and *P. pastoris* did not lead to any specific absorption spectra (native and CO differential spectra) or activities. This points towards occurring problems such as extremely low secretion rates, incorrect haem incorporation or protein misfolding.

The adaptation to episomal plasmid expression in *P. pastoris* proved that the entire modular signal peptide shuffling system could be readily transferred to another yeast organism. The applicability in *P. pastoris* furthermore paves the way towards future-directed evolution enterprises entirely performed in *P. pastoris*, further streamlining the workflow from gene discovery to construct identification and large-scale protein production. In comparison to the *S. cerevisiae*-based episomal system, the *P. pastoris*-based episomal plasmid expression of *Mth*UPO retained 60% of the activity (Fig. 4E). However, there is still plenty of potential for *P. pastoris* production optimisation utilising different promoters, carbon sources, induction and co-feeding strategies[50,62]. A substantial synergistic influence of the promoter−signal peptide combination was observed, as underlined for *Tte*UPO in Fig. 4D. This observation represents an aspect that should be further investigated in detail, for example, by expanding the modular system, including an additional shuffling module for a set of promoters to achieve simultaneous signal peptide and promoter shuffling. Production of *Mth*UPO utilising the integrative plasmid led to substantially improved production when compared to the episomal counterpart. Nevertheless, the obtained interclonal variation within the integrative system is substantial, rendering the episomal plasmid expression more suitable for performing reliable high-throughput endeavours (Fig. 4C). In the case of *Mth*UPO, we observed a diminished split-GFP response compared to the respective *S. cerevisiae* construct, which might be explained through differing glycosylation patterns, as observed by SDS Gel analysis (Supplementary Figs. 8 and 14). We hypothesise that this heterogeneous glycosylation pattern might mask the C-terminal protein tag within a proportion of enzymes in a greater extent than *S. cerevisiae*, thus impeding successful interaction with the GFP 1–10 counter fragment for GFP reconstruction. Also, the removal of the C-terminal tag by endogenous proteases is a possible scenario.

UPOs have been reported to be homologously produced over the course of several weeks in bioreactors to yield *Aae*UPO, *Cra*UPO (*Coprinus radians*), *Cgl*UPO and *Mro*UPO in production titres of 9 mg/L[2], 19 mg/L[63], 40 mg/L[9] and 445 mg/L[41], respectively. Besides the time-consuming production, only the wild-type enzyme can be produced, and the overall recovery of pure protein after several purification steps are reported to be below 20%[2,41,63]. Using a heterologous yeast expression system in a shake flask format, the highest reported production titres are obtained with the engineered *Aae*UPO* in *S. cerevisiae* and *P. pastoris* (each 8 mg/L)[10,15].

The newly discovered peroxygenases *Tte*UPO displayed unprecedented UPO expression titres in S. cerevisiae of 17 mg/L in *S. cerevisiae*. *Mth*UPO revealed a good production titre in *S. cerevisiae* (5 mg/L) and after transfer to *P. pastoris* the overall highest expression yields in *P. pastoris* with 24 mg/L could be achieved. Additionally, we could acquire high recovery of highly pure protein after one-step TwinStrep-based affinity purification[38,40] (Supplementary Figs. 8 and 14).

To compare the results of the bioconversion setups to literature-derived data, in all experiments the secretion variant *Aae*UPO* was included, which exhibits comparable catalytic properties to the fungal wild-type enzyme *Aae*UPO and therefore allows the comparison with previously obtained analysis in the literature[10]. As there are no data available for the homologously expressed *Mth*UPO, *Tte*UPO and *Cgl*UPO, this setup was the best way to allow comparative analysis of enzymatic performances in previous heterologous and homologous setups.

The relevance of expanding the set of recombinant UPOs is reflected by the fact that *Cgl*UPO, *Mth*UPO and *Tte*UPO displayed a different substrate specificity when compared to *Aae*UPO* (Fig. 3). When testing the conversion of naphthalene within our reaction setup, *Aae*UPO* showed the highest activities with TONs of 2950, yielding 92% of 1-naphthol and generally low overoxidation to para-naphthoquinone. This 1-naphthol to naphthoquinone product distribution is in accordance with previously obtained data[60]. *Cgl*UPO, *Tte*UPO and *Mth*UPO showed lowered TONs between 720 and 1260 for 1-naphthol formation and an elevated ratio of 49–55% of the naphthoquinone product. This work on *Mth*UPO catalysed naphthalene oxidation was recently expanded by performing protein engineering and testing a range of naphthalene derivatives[64].

The epoxidation experiments of styrene using *Aae*UPO* led in our setup to 4580 TONs and a 2% *ee*.

Previously reported data for *Aae*UPO revealed up to 10000 TONs with 4.6% *ee* when using a light-driven in situ hydrogen peroxide formation[65]. Combining *Aae*UPO* and tert-butylhydroperoxide added via a syringe pump setup resulted in 3200 TONs and 12% *ee*[66]. Within our setup, when using *Cgl*UPO similar epoxidation activities with TONs of 4120 were obtained.

Most interestingly, *Mro*UPO, *Cgl*UPO and *Mth*UPO exhibited substantially higher enantioselectivities of 30, 44 and 45 % *ee*, respectively, than the reported data for *Aae*UPO and *Aae*UPO*. This indicates differences in the active site geometry of the diverse UPO set and therefore provides an interesting point for further protein engineering towards higher stereoselectivities of styrene epoxidations.

For the benzylic hydroxylation of the homologous phenylalkane row, ranging from phenylethane to phenylpentane *Aae*UPO* displayed the highest activities and selectivities using phenylethane (4540, 98% *ee*) and phenylpropane (1670, 99% *ee*) as substrates, but only traces of product for benzylic hydroxylation of phenylbutane and phenylpentane. Wild-type *Aae*UPO was previously reported to achieve TONs of 10600 and 7100 for phenylethane and phenylpropane, respectively, and excellent selectivities (>99% *ee*) in both cases[67]. Additionally, verifying the observed negative tendency of decreasing benzylic hydroxylation activity for increasing alkyl chain length *Aae*UPO* in combination with an enzymatic cascade for the in situ production of hydrogen peroxide even led up to 294700 TONs[1].

The enzymes *Cgl*UPO, *Mth*UPO and *Tte*UPO showed an opposite selectivity when compared to *Aae*UPO* regarding benzylic hydroxylation and displayed the lowest activities for phenylethane and highest for phenylbutane and –pentane conversion, respectively. *Tte*UPO furthermore catalyses the formation of the opposite alcohol enantiomer (S) compared to the other enzymes for the conversion of phenylpropane to phenylbutane.

Good activities and excellent enantioselectivities could also be achieved when challenging the enzymes for the benzylic hydroxylation of *N*-phthalimide protected phenethylamine in case of the enzymes *Cgl*UPO and *Mth*UPO. This observation is vastly different from *Aae*UPO*, displaying no product formation and no known enantioselective conversion of substrates of similar structure.

The high UPO production yields in *P. pastoris* of *Mth*UPO enabled a proof-of-concept approach to yield the chiral alcohol product on the preparative scale with a challenging phenethylamine derivative and yielded 57% yield and 98.6% *ee*. The direct benzylic hydroxylation of phenethylamine compounds was previously reported for copper-dependent dopamine b-hydroxylases (DbH), but not on a preparative scale[68,69]. As DbHs suffer from difficult expression, their engineering towards similar substrates and higher activities is currently hampered.

To allow other researchers to harness the modular yeast system, we deposited all relevant plasmids (signal peptides, protein tags and expression plasmids) as a kit with the non-profit plasmid repository Addgene called *Yeast Secrete and Detect* Kit (Kit # 1000000166). The herein developed overall workflow for functional UPO secretion and detection can be performed within a minimal period of 15 days (Supplementary Fig. 2). Within this period, beneficial episomal constructs are identified in a 96-well high throughput system, exploiting activity measurements and protein quantification by the split-GFP assay[34,35]. Identified constructs can then be directly used for upscaling to shake flasks, one-step affinity target protein purification and subsequent bioconversion testing.

In summary, the obtained data of this study proves that the built workflow starting from a putative UPO gene, followed by identification of suitable expression constructs via signal peptide shuffling in combination with high-throughput screening in *S. cerevisiae* as well as *P. pastoris* and subsequent production upscaling can lead to highly enantioselective preparative product formations of pharmaceutically valuable building blocks.

In the future, this workflow could be applied to other UPO genes or generally genes of interest, which are suitable for production in yeast, especially for proteins that might require efficient post-translational modifications such as glycosylation and disulfide linkage. Besides target protein secretion, the constructed expression plasmids also allow for intracellular production when no signal peptide is attached. During the submission and revision process of this publication, two papers demonstrated the engineering of *Mth*UPO using the herein developed *S. cerevisiae* setup[64,70] and one publication expanded the episomal *P. pastoris* system to a range of promoters and new UPO enzymes[71].

Methods

Chemicals. Solvents were used as received without further purification. Ethyl acetate and acetone were utilised in GC ultra-grade (≥99.9%) from Carl Roth (Karlsruhe, DE). Acetonitrile was purchased from Merck (Darmstadt, DE) in gradient grade for LC (≥ 99.9%). Deuterated solvents for NMR spectroscopy were purchased from Deutero (Kastellaun, DE). All further reaction chemicals were purchased either from Sigma-Aldrich (Hamburg, DE), TCI Chemicals (Tokyo, JP), Merck (Darmstadt, DE), abcr (Karlsruhe, DE) or Fluka Chemika (Buchs, CH) and used as received.

Enzymes and cultivation media. For cultivation of *E. coli* cells terrific broth (TB) media from Carl Roth (Karlsruhe, DE) was used. For cultivation of *S. cerevisiae* cells D-Galactose, Peptone and Synthetic Complete Mixture (Kaiser) Drop-Out (-URA) were purchased from Formedium (Hunstanton, GB). Yeast nitrogen base (without amino acids) and Yeast extract were purchased from Carl Roth (Karlsruhe, DE). For *P. pastoris* cultivation methanol (99.9% Chromasolv purity grade) purchased from Honeywell Chemicals (Seelze, DE) was used as additional carbon source. PNGaseF and BsaI were purchased from New England Biolabs (Ipswich, US). BbsI and FastDigest AscI were purchased from ThermoFisherScientific (Waltham, US) and T4 DNA Ligase from Promega (Madison, US).

Oligonucleotides and gene parts. All oligonucleotides were purchased in the lowest purification grade "desalted" and minimal quantity at Eurofins Genomics (Ebersberg, DE). The *Pichia pastoris* CAT1 promoter was purchased as a gene part from Twist Bioscience (San Francisco, US). The genes of the *Aae*UPO variant *Aae*UPO*, *Gma*UPO, *Mwe*UPO and the sfGFP 1-10 gene were purchased as plasmid-cloned genes from Eurofins Genomics (Ebersberg, DE). The genes of *Cgl*UPO, *Mth*UPO and *Tte*UPO were retrieved as codon-optimised (*S. cerevisiae* codon usage) gene strands from Eurofins Genomics.

Expression plasmid construction for *S. cerevisiae*. A Level 1 Golden Gate-based shuttle expression plasmid was constructed using a pAGT572 plasmid as backbone structure[72], which can be propagated in *E.coli* and *S. cerevisiae*. It enables antibiotic selection (Ampicillin resistance) and yeast auxotrophy selection (URA3 marker). To enable expression of a target gene a Gal 1.3 Promoter—a truncated, modified version of the widespread GAL1 Promoter is integrated upstream and a strong DIT1 terminator downstream of the cloning acceptor site. As placeholder for a target gene sequence a lacZ cassette (approx. 600 bp) is integrated, which enables β-galactosidase-based blue/white selection of transformants based on the conversion of X-Gal. Upon digestion with BsaI the lacZ cassette is released, and a fitting open reading can be integrated in frame (e.g. Signal Peptide-Gene-C-terminal Tag) into the plasmid, thereby reconstituting a fully functional expression plasmid. The constructed expression plasmid was coined pAGT572_Nemo_2.0. Using the pAGT572 plasmid backbone and the GAL1 Promoter as units a second expression plasmid coined pAGT572_Nemo was constructed that follows the same functionality and principle but exhibits the original GAL1 promoter.

Expression plasmid construction for *Pichia pastoris*. Two level 1 Golden Gate-based shuttle expression plasmids were constructed, which can be propagated in *E. coli* (Amp^R) as well as *P. pastoris* (Hygromycin B^R). To enable episomal plasmid propagation in *P. pastoris*, the plasmids were equipped with a previously described functional ARS sequence[47,73], which was PCR amplified from *Kluveromyces lactis* genomic DNA. The plasmids exhibit the strong constitutive GAP promoter (pPAP001) or the strong methanol inducible promoter CAT1 (pPAP002), both in combination with a strong GAP terminator (tGAP). As placeholder for a target gene sequence, a lacZ cassette is used. For the stable integration of transcription units into the *P. pastoris* genome, a third universal integrative plasmid (pPAP003) was designed. A shuttle plasmid was constructed, which can be propagated in *E. coli* ($Kanamycin^R$) as well as *P. pastoris* ($Hygromycin^R$). As placeholder for a target transcription unit a lacZ cassette is integrated. Upon digestion with BbsI the lacZ cassette is released and a fitting transcription unit (Promoter- ORF- Terminator) can be integrated (derived from respective pPAP001 and pPAP002 episomal plasmids as donors) into the plasmid, thereby reconstituting a fully functional integration plasmid. Several parts (GAP promoter, GAP terminator, AOX integration marker and Hygromycin B resistance marker) of the constructed plasmids

were PCR amplified and derived from a previously introduced Golden Gate based *P. pastoris* assembly system, coined GoldenPiCS[44].

Golden Gate cloning of Level 0 standard parts.
All genetic parts were cloned as individual Level 0 standard modules into the universal Level 0 acceptor plasmid pAGM9121 (Spectinomycin^R). Therefore, three functional units were pre-defined: (a) signal peptide (contains start codon); (b) gene (lacking start and stop codon) and (c) C-terminal Protein-tag (contains stop codon). 4 bp sticky overhangs that are released upon Type II s enzyme treatment (BsaI and BbsI) and guide subsequently a correct reassembly were chosen accordingly to the nomenclature of gene assembly as described within the ModularCloning (MoClo) system[33]. An overview of the reassembly concept is provided in Supplementary Fig. 1. For the cloning of the individual modules suitable oligonucleotides were designed to allow for cloning into pAGM9121. Primers followed a general scheme (Supplementary Fig. 1). Fragments were amplified by PCR from a suitable template sequence or generated by hybridisation of two complementary oligonucleotides. PCR products were analysed as small aliquot (5 µL) by agarose gel electrophoresis for occurrence of the expected size and the remaining sample subsequently recovered and purified using a NucleoSpin® Gel and PCR Clean-up Kit (Macherey-Nagel, Düren, DE). Golden Gate reactions were performed in a total volume of 15 µL. The final reaction volume contained 1-fold concentrated T4 ligase buffer (Promega, Madison, US). Prepared reaction mixtures containing ligase buffer, acceptor plasmid (20 fmol) and the corresponding insert (20 fmol) was adjusted to 13.5 µL with ddH_2O. In a final step, the corresponding enzymes were quickly added. First, a volume of 0.5 µL of the respective restriction enzyme BbsI (5 units/µL) was added and then 1 µL (1–3 units/µL) of T4 ligase. Golden Gate reactions were performed for 3 h (37 °C) and concluded by an additional enzyme inactivation step (80 °C; 20 min). The whole Golden Gate reaction volume was used to transform chemically competent *E. coli* DH10B cells. After heat shock transformation and recovery, the mixture was plated in different quantities on selective LB Agar plates (50 µg × mL^{-1} X-Gal; 100 µg × mL^{-1} Spectinomycin; 150 µM IPTG). Based on the occurrence of the lacZ selection marker one can easily distinguish between white colonies (recombined plasmid) and empty plasmid (blue). In general, the described protocol led to several thousand recombinant colonies with a nearly absolute proportion (>99%) of recombined, white colonies. Single colonies were checked for correct insert sizes by means of colony PCR (pAGM9121 sequencing primer; Supplementary Table 1). Positively identified clones were inoculated into 4 mL of TB-Medium (100 µg × mL^{-1} Spectinomycin) and corresponding plasmid DNA prepared (NucleoSpin Plasmid Kit (Macherey-Nagel, Düren, DE)). After verification of the correct, intended insert sequence by Sanger Sequencing (Eurofins Genomics, Ebersbach, DE) respective plasmids were included for further use within the modular Golden Gate cloning approaches.

Golden Gate cloning of expression plasmids. The expression plasmids (*S. cerevisiae*: pAGT572_Nemo and pAGT572_Nemo 2.0; *P. pastoris*: pPAP001 and pPAP002) were used as respective acceptor plasmid for the assembly of the individual tripartite open reading frames (**5′** Signal Peptide-Gene-C-terminal Tag **3′**). The individual parts were thereby derived as parts from standard level 0 plasmids (pAGM9121 backbone), which can be released from the pAGM9121 backbone upon BsaI restriction digest. Golden Gate reactions were performed in a total volume of 15 µL. The final reaction volume contained 1-fold concentrated T4 ligase buffer. Prepared reaction mixtures containing ligase buffer, the acceptor plasmid (20 fmol) and the corresponding inserts as level 0 modules (Signal Peptide, Gene, C-terminal Tag) were added to 20 fmol each and the overall volume adjusted to 13.5 µL with ddH_2O. In the case of a signal peptide shuffling approach 17 different pAGM9121- Signal Peptide combinations were added in equimolar ratios (1.2 fmol each). In a final step, the corresponding enzymes were quickly added. First, a volume of 0.5 µL of the restriction enzyme BsaI (10 units/µL) was added and then 1 µL (1–3 units/µL) of T4 ligase. Golden Gate reactions were performed using a temperature cycling program (50x passes) between 37 °C (2 min) and 16 °C (5 min) and concluded by an additional enzyme inactivation step (80 °C; 20 min). The whole Golden Gate reaction volume was used to transform chemically competent *E. coli* DH10B cells. After heat shock transformation and recovery the mixture (approx. 320 µL) was split into two fractions, 50 µL were plated on selective LB Agar plates (+ X-Gal; 100 µg × mL^{-1} Ampicillin; + IPTG) and the remaining volume used to directly inoculate 4 mL TB Medium (+ Amp) to preserve the genetic diversity of the shuffling library. The following day the success of the Golden Gate reaction was evaluated based on the performed blue/white screening, discriminating the empty plasmid (lacZ; blue) from recombined, white colonies. In general, the described protocol for ORF assembly and signal peptide shuffling as special case led to several hundred recombinant colonies with a high proportion (>90%) of recombined, white colonies. In the case of single defined, "unshuffled" constructs single colonies were checked for correct insert sizes by means of colony PCR (using respective plasmid sequencing primer). Positively identified clones were inoculated into 4 mL of TB-Medium (+Amp) and corresponding plasmid DNA prepared (NucleoSpin Plasmid Kit (Macherey-Nagel, Düren, DE)). In the case of shuffled signal peptide constructs, plasmid DNA was prepared as a library by direct inoculation of the transformation mixture into the liquid culture and subsequent DNA isolation (see above).

Plasmid transformation into *S. cerevisiae*. Respective single plasmids or plasmid mixtures (pAGT572_Nemo or pAGT572_Nemo 2.0 backbone) were used to transform chemically competent *S. cerevisiae* cells (INVSc1 strain) by polyethylene glycol/lithium acetate transformation. INVSc1 cells were prepared and stored at −80 °C in transformation buffer (15% (v/v) glycerol; 100 mM lithium acetate; 500 µM EDTA; 5 mM Tris-HCl pH 7.4) as 60 µL aliquots until usage. For transformation, an amount of 100 ng of the plasmid preparation was added to 10 µL of lachssperm DNA (10 mg/mL; Sigma Aldrich, Hamburg, DE) and mixed. This mixture was then added to a thawed aliquot of INVSc1 cells on ice. 600 µL of transformation buffer (40% (v/v) polyethylene glycol 4000; 100 mM lithium acetate; 1 mM EDTA; 10 mM Tris-HCl pH 7.4) were added and the cells incubated under rigid shaking (30 °C; 850 rpm) for 30 min. Afterwards, 70 µL of pure DMSO was added and the cells incubated for a further 15 min at 42 °C without shaking. Finally, the cells were precipitated by short centrifugation, the supernatant discarded, and the cell pellet resuspended in 350 µL sterile ddH_2O. Different volumes were plated on Synthetic Complement (SC) Drop Out plates supplemented with 2% (w/v) glucose as carbon source and lacking Uracil as an auxotrophic selection marker. Plates were incubated for at least 48 h at 30 °C till clearly background distinguishable white colonies appeared.

Plasmid transformation into *P. pastoris*. Respective single plasmids or plasmid mixtures (pPAP001 or pPAP002 backbone) were used to transform *P. pastoris* cells (X-33 strain) by means of electroporation. Electrocompetent X-33 cells were prepared according to a condensed transformation protocol for *P. pastoris*[74]. Cells were stored in BEDS solution (10 mM bicine-NaOH pH 8.3, 3% (v/v) ethylene glycol, 5% (v/v) DMSO and 1 M sorbitol) as 60 µL aliquots (−80 °C) till further use. For the transformation of episomal plasmids 20 ng of the circular plasmid were added to one aliquot of thawed competent X-33 cells. The cell-plasmid mix was transferred to an electroporation cuvette (2 mm gap) and cooled for 10 min on ice prior to the transformation. Electroporation was performed using a Micropulser Device (Bio-Rad, Hercules, US) and using manual implemented, standardised settings (1.5 kV, 1 pulse) for all transformation setups, leading to a general pulse interval of 5.4–5.7 ms. Immediately after electroporation cells were recovered in 1 mL of ice-cold YPD-Sorbitol solution (10 g/L peptone, 5 g/L yeast extract, 500 mM sorbitol), transferred to a new reaction tube and incubated for one hour under rigid shaking (30 °C, 900 rpm) in a Thermomix device (Eppendorf, Hamburg, DE). After incubation, cells were precipitated by centrifugation (5.700 rpm, 5 min). The supernatant was discarded, and the cells resuspended in 200 µL of fresh YPD medium. 100 µL of the suspension was then plated on selective YPD Agar plates supplemented with 150 µg/mL Hygromycin B. Plates were incubated at 30 °C for at least 48 h till clearly visible colonies appeared. In general, the described setup led to the occurrence of several hundred colonies per plate. For the transformation of integrative plasmids (pPAP003 backbone) the setup was slightly modified as linearised plasmid is used for transformation. Therefore, previously prepared circular plasmid DNA was digested with AscI (Isoschizomer: SgsI). 2.5 µg of the respective plasmid DNA were mixed with 3 µL of 10x fold FastDigest Buffer, the volume adjusted to 29.5 µL using ddH_2O and in the last step, 0.5 µL of FastDigest SgsI added. Digestion was performed overnight (16 h, 37 °C) and terminated by an enzyme inactivation step (20 min, 80 °C). Linearised plasmid DNA was then subsequently prepared according to the manufacturer instruction using a Nucleospin® Gel and PCR clean up Kit (Macherey-Nagel, Düren, DE). The transformation of *P. pastoris* was performed in a congruent manner as described before, except for using 100 ng linearised plasmid for transformation, since the overall transformation efficiency is substantially reduced in comparison to the transformation of the circular, episomal plasmid.

Microtiter plate cultivation of *S. cerevisiae*. For peroxygenase production in microtiter plate format specialised 96 half-deep well plates were utilised. The model type CR1496c was purchased from EnzyScreen (Heemstede, NL) and plates were covered with fitting CR1396b Sandwich cover for cultivation. Plates and covers were flushed before every experiment thoroughly with 70% ethanol and air-dried under a sterile bench until usage. In each cavity, 220 µL of minimal expression medium were filled and inoculated with single, clearly separated yeast colonies using sterile toothpicks. The minimal selective expression medium (1x concentrated Synthetic complement Drop out stock solution lacking uracil; 2% (w/v) galactose; 71 mM potassium phosphate buffer pH 6.0; 3.2 mM magnesium sulfate; 3.3% (v/v) ethanol; 50 mg/L haemoglobin; 25 µg/L chloramphenicol) was freshly prepared out of sterile stock solutions immediately before each experiment, mixed and added to the cavities. After inoculation of the wells the plates were covered, mounted on CR1800 cover clamps (EnzyScreen) and incubated in a Minitron shaking incubator (Infors, Bottmingen, SUI) for 72 h (30 °C; 230 rpm). After cultivation, the cells were separated from the peroxygenase containing supernatant by centrifugation (3400 rpm; 50 min; 4 °C).

Microtiter plate cultivation expression in *P. pastoris*. General experimental setup as before with *S. cerevisiae*. Each cavity was filled with 220 µL of buffered complex medium (BM) and inoculated with single, clearly separated yeast colonies using sterile toothpicks. Basic BM (20 g/L peptone; 10 g/L yeast extract; 100 mM potassium phosphate buffer pH 6.0; 1x YNB (3.4 g/L yeast nitrogen base without

amino acids; 10 g/L ammonium sulfate); 400 µg/L biotin; 3.2 mM magnesium sulfate; 25 µg/L chloramphenicol; 50 mg/L haemoglobin; 150 µg/L Hygromycin B) was freshly prepared out of sterile stock solutions immediately before each experiment, mixed and added to the cavities. Depending on the type of utilised promoter (pPAP001: P_{GAP} and pPAP002: P_{CAT1}), the BM medium was supplemented with different carbon sources for cultivation and induction, respectively. pPAP001 constructs were cultivated utilising 1.5% (w/v) of glycerol or glucose as sole carbon source. In the case of the methanol inducible CAT1 promoter, a mixed feed strategy was employed combining 0.5% (w/v) of glycerol with 1.5% (v/v) methanol. Cultivation and centrifugation was as described before for *S. cerevisiae*.

Shake flask cultivation *S. cerevisiae*. A preculture of 50 mL of SC Drop out selection media (+ 2% (w/v) raffinose and 25 µg/L chloramphenicol) was inoculated with one single colony derived from a selection plate (SC Drop; -Uracil) and grown for 48 h (30 °C; 160 rpm; 80% humidity). This incubation typically resulted in a final OD_{600nm} of approx. 12 to 13. The main expression culture was inoculated with a starting optical density of 0.3. For large-scale peroxygenase production rich non-selective expression medium (20 g/L peptone; 10 g/L yeast extract; 2% (w/v) galactose; 71 mM potassium phosphate buffer pH 6.0; 3.2 mM magnesium sulfate; 3.3% (v/v) ethanol; 25 µg/L chloramphenicol) was utilised. Cultivation was performed in 2.5 L Ultra yield flasks (Thomson Instrument, Oceanside, US) in a final culture volume of 500 mL per flask after sealing the flask with breathable Aeraseal tape (Sigma Aldrich, Hamburg, DE) allowing for gas exchange. The main cultures were incubated for further 72 h (25 °C; 110 rpm; 80 % humidity). After cultivation, the cells were separated from the peroxygenase containing supernatant by centrifugation (4300 rpm; 35 min; 4 °C).

Shake flask cultivation *P. pastoris*. For the large-scale protein production using shake flasks genomically integrated single constructs (pPAP 003 backbone; integration into chromosomal 3′ region of *P. pastoris* AOX1 gene) were chosen. These constructs were previously identified by screening at least 4 different colonies per individual construct within an MTP screening setup and choosing a respective production strain based on a high as possible, clearly distinguishable NBD conversion in comparison to the background control (pPAP003 empty plasmid).

Precultures were prepared in 50 mL YPD medium (+ 25 µg/L chloramphenicol) and cultivated for 48 h (30 °C; 160 rpm; 80% humidity), typically resulting in a final OD_{600nm} of approx. 17 to 19. The main expression culture was inoculated with a starting optical density of 0.3. For large-scale peroxygenase production BM-based expression media (20 g/L peptone; 10 g/L yeast extract; 100 mM potassium phosphate buffer pH 6.0; 1x YNB (3.4 g/L yeast nitrogen base without amino acids; 10 g/L ammonium sulfate); 400 µg/L biotin; 3.2 mM magnesium sulfate; 25 µg/L chloramphenicol) was utilised. In the case of constitutively expressing GAP constructs 2% (w/v) Glucose was added (BMG media) as a carbon source for Pichia growth. In the case of the methanol inducible CAT1 promoter a two-phase feeding was applied, firstly inoculating the cells into BM medium (see above) supplemented with 0.5% (w/v) glycerol as carbon source. 24 h and 48 h after inoculation 0.8% (v/v) of methanol were added as an inducer of the CAT1 promoter. Cultivation and final centrifugation were performed as described for *S. cerevisiae*.

Supernatant ultrafiltration and protein purification. The supernatant was concentrated approx. 20-fold by means of ultrafiltration. Therefore, a Sartocon Slice 200 membrane holder (Sartorius, Göttingen, DE) was equipped with a Sartocon Slice 200 ECO Hydrosart Membrane (10 kDa nominal cut-off; Sartorius) within a self-made flow setup. The flow system for ultrafiltration was operated by an EasyLoad peristaltic pump (VWR International, Darmstadt, DE). Firstly, the cleared supernatant (1 L) was concentrated approx. 10-fold to a volume of 100 mL and 900 mL of purification binding buffer (100 mM Tris-HCl pH 8.0, 150 mM NaCl) were added as a buffer exchange step. This sample was then concentrated to a final volume of 50 mL. Protein purification was implemented utilising the C-terminal attached double Strep II Tag (WSHPQFEK), coined TwinStrep® (Iba Lifesciences, Göttingen, DE). As column material, Strep-Tactin®XT Superflow® columns (1 mL or 5 mL; Iba Lifesciences) were chosen and the flow system operated by an EasyLoad peristaltic pump (VWR). In a first step, the column was equilibrated with 5 column volumes (CVs) binding buffer. The concentrated sample (50 mL) was filter sterilised (0.2 µm syringe filter) and applied to the column with an approximate flow rate of 1 mL/min. After application, the column was washed with 7 CVs binding buffer. Elution was performed based on binding competition with biotin, therefore approx. 2 CV of elution buffer (100 mM Tris-HCl pH 8.0, 150 mM NaCl; 50 mM biotin) were applied to the column. The pooled elution fraction was then dialysed overnight (4 °C) against 5 L of storage buffer (100 mM potassium phosphate pH 7.0) using ZelluTrans dialysis tubing (6–8 kDa nominal cut-off; Carl Roth) and the recovered, dialysed sample stored at 4 °C till further use.

Plasmid preparation of episomal plasmids from yeast. Yeast plasmids of identified clones were recovered by means of digestive Zymolase cell treatment and alkaline cell lysis. Therefore, clones were inoculated and cultivated for 48 h (30 °C; 250 rpm) in 4 mL of selection medium, in case of *S. cerevisiae* SC Drop out medium (-Uracil; 2 % (w/v) Glucose) was used and in the case of *P. pastoris* single colonies were inoculated into 4 mL of YPD (+ 150 µg/mL Hygromycin) to preserve the selection pressure. After cultivation cells were pelleted by centrifugation and 1 mL of washing buffer (10 mM EDTA NaOH; pH 8.0) added and the pellet resuspended by light vortexing. Cells were subsequently pelleted (5000 × *g*; 10 min) and the supernatant discarded. Afterwards, cells were resuspended in 600 µL of Sorbitol Buffer (1.2 M sorbitol, 10 mM $CaCl_2$, 100 mM Tris-HCl pH 7.5, 35 mM β-mercaptoethanol) and 200 units of Zymolase (Sigma Aldrich, Hamburg, DE) added followed by an incubation step for 45 min (30 °C; 800 rpm) for cell wall digestion. After incubation cells were pelleted by centrifugation (2000 × *g*; 10 min) the supernatant discarded, and the plasmid preparation started with an alkaline lysis step following the manufacturer's instructions (NucleoSpin Plasmid Kit, Macherey Nagel). In the final step, yeast-derived episomal plasmids were eluted in 25 µL elution buffer (5 mM Tris-HCl pH 8.5), and the whole eluate used to transform one aliquot of *E. coli* DH10B (transformation as described above), plating the whole transformation mix on a selective LB-Agar plate (Amp^R). On the following day, single colonies were picked, inoculated into 4 mL of TB medium (+Amp), plasmid prepared and sent for Sanger Sequencing (Eurofins Genomics) to elucidate the respective sequence of the open reading frame.

Thermostability measurements. Thermostability measurements of the purified enzymes were performed by Differential Scanning fluorimetry (DSF) on a Prometheus NT.48 nanoDSF instrument (NanoTemper Technologies GmbH, München, DE) in storage buffer (100 mM Tris-HCl pH 7.0). Approximately 10 µL of sample volume were loaded into a Prometheus NT.48 High Sensitivity Capillary (NanoTemper Technologies GmbH). Protein unfolding was subsequently monitored by following the ratio of intrinsic protein tyrosine and tryptophan fluorescence at 350 nm to 330 nm over time, increasing the temperature from 20 °C to 95 °C with a heating ramp of 1 °C per minute. The melting temperature corresponds to the maximum of the first derivative of the 350/330 nm ratio. All measurements were performed at least in triplicates.

Split-GFP assay. Protein normalisation was performed employing the principle of a split GFP normalisation assay as described by Santos-Aberturas et al.[35] with slight modifications. The GFP fluorescence complementation fragment sfGFP 1–10 was cloned into the Golden Mutagenesis plasmid pAGM22082_cRed[32] for T7 promoter controlled expression in *E. coli* (BL 21 DE3 strain). The sfGFP 1–10 fragment was prepared as an inclusion body preparation according to the previous reports[35]. For measurement, a 96 well Nunc MaxiSorp Fluorescence plate (ThermoFisherScientific, Waltham, US) was blocked (25 min, light shaking) with 180 µL of BSA blocking buffer (100 mM Tris-HCl pH 7.4, 100 mM NaCl, 10% (v/v) glycerol, 0.5% (w/v) BSA). The blocking solution was discarded and 20 µL of the yeast media supernatant (*S. cerevisiae* or *P. pastoris*) derived from the peroxygenase expression plates added. A 10 mL aliquot of the sfGFP 1–10 complementation fragment (storage: −80 °C) was quickly thawed in a water bath and diluted 1x fold into ice-cold TNG buffer (100 mM Tris-HCl pH 7.4, 100 mM NaCl, 10% (v/v) glycerol) and 180 µL of this screening solution added to each well. Immediate fluorescence values (GFP fluorophore: excitation wavelength: 485 nm; emission wavelength: 535 nm; top read mode) were measured using a 96 well plate fluorescence reader Spark 10 M (TECAN, Grödig, AT), setting an empty plasmid control well as 10% of the overall signal intensity (well calculated gain). After storage of the plate for at least one up to three nights (at 4 °C) final fluorescence values were measured in a comparable manner. Protein quantities were then normalised based on the relative fluorescence increase of each respective well (differential values) and in comparison, to the empty plasmid backbone.

DMP assay. The use of 2,6-Dimethoxyphenol (DMP) as a suitable microtiter plate substrate for the measurement of peroxygenase catalysed conversion to the colorimetric product coerulignone has been described before[60]. The described conditions have been adapted with slight modifications. In brief, 20 µL of peroxygenase containing supernatant were transferred to a transparent polypropylene 96-well screening plate (Greiner Bio-One, Kremsmünster, AT) and 180 µL of screening solution (final: 100 mM potassium phosphate pH 6.0; 3 mM 2,6-Dimethoxyphenol; 1 mM hydrogen peroxide) added. Absorption values (λ: 469 nm) of each well were immediately measured after addition in a kinetic mode (measurement interval: 30 s) over a duration of 5 min utilising the 96-well microtiter plate reader SpectraMax M5 (Molecular Devices, San José, US). Slope values of absorption increase corresponding to coerulignone formation were obtained, paying special attention to the linearity of the observed slope to obtain reliable relative DMP conversion values for comparison of the respective wells.

NBD assay. The use of 5-nitro-1,3-benzodioxole (NBD) as a suitable microtiter plate substrate for the measurement of peroxygenase catalysed conversion to the colorimetric product 4-Nitrocatechol has been described before[37,75]. Screening as described above for DMP but adding 180 µL of screening solution (final: 100 mM potassium phosphate pH 6.0; 1 mM NBD; 1 mM hydrogen peroxide; 12% (v/v) acetonitrile). Absorption values (λ: 425 nm) of each well were immediately measured after addition in a kinetic mode (measurement interval: 30 s) over a duration of 5 min. Slope values of absorption increase corresponding to 4-nitrocatechol formation were obtained, paying special attention to the linearity of the observed

slope to obtain reliable relative NBD conversion values for comparison of the respective wells.

Resting-state absorption and haem CO complex measurements. The pooled and dialysed elution fractions (100 mM potassium phosphate pH 7.0) were subsequently used to record absorption spectra of the respective enzymes (*Mro*UPO, *Cgl*UPO, *Mth*UPO, *Tte*UPO) in their native, resting state (ferric iron; Fe^{3+}). For all measurements, a QS High precision Quartz Cell cuvette (Hellma Analytics, Müllheim, DE) with a path length of 10 mm was used. Spectra were recorded on a Biospectrometer Basic device (Eppendorf, Hamburg, DE) in the spectral range from 250 to 600 nm (interval: 1 nm) and subtracting the utilised storage buffer (100 mM potassium phosphate pH 7.0) as previous blank measurement. Haem carbon dioxide spectra (CO assay) were recorded after reducing the haem iron to its ferrous form (Fe^{2+}). Therefore, a spatula tip of sodium dithionite as the reducing agent was added to 1 mL of a respective enzyme fraction (see above) and mixed thoroughly till complete dissolution. This sample was immediately flushed with a constant carbon dioxide flow for 2 min (approx. 1 bubble/sec) to obtain the thiolate-haem carbon dioxide complex. The sample was immediately transferred to a cuvette and absorption measured as described above. The CO assay was also employed for the measurement of peroxygenase concentrations in the concentrated *P. pastoris* supernatant obtained after ultrafiltration. In this case, the supernatant was 10-fold diluted with potassium phosphate buffer (100 mM, pH 7.0). A spatula tip of sodium dithionite was then added to 2 mL of the diluted supernatant sample. After dividing the respective sample into two parts of 1 mL, one part was treated with carbon monoxide for 2 min as described above, and the CO untreated sample is used as a blank reference. Absorption measurements were performed by UV/Vis spectroscopy using a JASCO V-770 Spectrophotometer (JASCO Deutschland GmbH, Pfungstadt). The CO absorption maximum was measured at 444 nm, and a reference absorption wavelength was measured at 490 nm. For calculation, an extinction coefficient of 91,000 M^{-1} cm^{-1} was used, which appears to be generally valid for most haem-thiolate enzymes according to literature[76]. The enzyme concentration in the supernatant was then calculated using the formula:

$$c[\mu M] = \text{dilution factor} \times \frac{A_{444nm} - A_{490nm}}{0.091 \mu M^{-1} cm^{-1}}$$

pH range of NBD conversion. pH dependency of NBD conversion of the respective enzymes was investigated using different buffer system in the range between pH 2.0 and 11.0 (even numbers only). Each buffer was prepared as a 100 mM stock solution, potassium phosphate buffer was used for the pH values 2.0, 7.0 and 8.0. Sodium citrate was used in the range of pH 3.0 to 6.0 and Tris-HCl was used in the range of pH 8.0 to 11.0. Purified enzyme solutions (100 mM potassium phosphate pH 7.0) were diluted 10 to 20x fold in ddH_20 prior to the measurements leading to weakly buffered solutions as screening samples. The NBD assay was then performed as described before, mixing 20 µL of the enzyme dilution with 180 µL screening solution (87 mM corresponding buffer pH x; 500 µM NBD; 1 mM H_2O_2). All samples were measured as three biological replicates. Due to the strong pH-dependency of the molar extinction coefficient of the corresponding detected product 4-nitrocatechol a normalisation was performed. Therefore, the product 4-Nitrocatechol was prepared as 10 mM stock solution dissolved in acetonitrile and diluted into 990 µl of the corresponding screening buffer (final concentration: 10 µM) and after 5 min an absorption spectrum in the interval of 400 to 600 nm (Biospectrometer Basic device) recorded. Calculation of the correction factor of the respective samples (pH 2.0 to pH 11.0) was then performed regarding the utilised measurement wavelength of 425 nm. Finally, in consideration of the obtained pH correction factor, individual activity values derived from the respective measured absorption values were calculated.

Protein concentration determination and purification yield. Protein concentrations of the respective protein samples were determined after dialysis of the elution fractions (storage buffer: 100 mM potassium phosphate pH 7.0). In this regard, the colorimetric BCA assay was utilised, employing a Pierce™ BCA Protein Assay Kit (ThermoFisherScientific, Waltham, US) following the instructions of the manufacturer. Samples were measured in biological triplicates (25 µL of a previously diluted sample) and concentrations calculated based on a previously performed calibration curve using BSA (0–1000 µg/mL) as reference protein. To determine the overall yield of enzyme production per litre of culture volume, the determined concentration in the elution fraction was extrapolated to the overall NBD activity of the sample after ultrafiltration (column load). This calculation is performed since NBD is a highly specific substrate for peroxygenase activity, comparable background samples processed in a similar manner but using empty plasmid controls did not show any measurable conversion of NBD. Samples of every purification step (load, flow-through, wash and elution fraction) were collected, and NBD conversion rates of the respective fractions measured immediately after purification. In the case of non-complete binding of the enzyme fraction (remaining NBD activity in flow-through fraction) this remaining non-bound enzyme amount was taken into consideration for calculation for the overall volumetric production yield. The via BCA assay determined protein concentration of the elution fraction was extrapolated to the activity of the respective non-bond fraction, assuming a constant specific enzyme activity for NBD conversion and considering the volumes of the respective fractions, leading to an approximate enzyme titre per litre.

SDS gel analysis and PNGaseF treatment. Obtained elution fractions of the respective UPO enzymes were analysed for the apparent molecular weight and overall purity after the performed one step TwinStrep purification by means of SDS PAGE. Therefore, samples of the column load (after ultrafiltration; see above), elution fractions after dialysis and deglycosylated elution fraction samples were analysed on self-casted SDS PAGE (10 or 12% of acrylamide) utilising a Bio-Rad (Hercules, US) Mini-Protean® Gel electrophoresis system. A defined PageRuler Prestained Protein Ladder (ThermoFisherScientific, Waltham, US) was included, covering a MW range between 10 and 180 kDa. Proteins were visualised using a colloidal Coomassie G-250 staining solution. To obtain N-type deglycosylated protein samples, elution fractions were enzymatically treated with Peptide-N-Glycosidase F (PNGaseF) from *Flavobacterium meningosepticum*, which is capable of cleaving Asparagine linked high mannose type glycan structures as typically occurring in *P. pastoris* and *S. cerevisiae* derived glycosylation patterns. Therefore, 45 µL of a respective elution fraction was mixed with 5 µL of denaturing Buffer (final 0.5% SDS; 40 mM DTT) and denatured for 10 min (100 °C). After a cooling step to room temperature 6 µL of NP-40 solution (final: 1 %) and 6 µL of GlycoBuffer2 (500 mM sodium phosphate; pH 7.5) were added and the solution thoroughly mixed. Finally, 1 µL of PNGaseF (New England Biolabs, Ipswich, US) was added and the sample incubated under light shaking (37 °C) for 3 h. After incubation, the sample was prepared for further analysis by adding 5x fold SDS sample buffer and subsequent SDS PAGE analysis executed as described before. In the case of native deglycosylation, 90 µL of enzyme sample were mixed with 10 µL of GlycoBuffer2 (500 mM Sodium Phosphate; pH 7.5) and 1 µL of PNGaseF added. The mixture was incubated at 37 °C in a thermal PCR cycler (24 or 48 h) and subsequently analysed for UPO activity in comparison with an equally treated sample (without PNGAseF addition) by means of the NBD assay (see above).

Protein identification by MS. Protein samples after protein purification (in 100 mM Tris-HCl pH 8.0, 150 mM NaCl; 50 mM biotin) were enzymatically digested with trypsin and desalted according to ref. [77]. The resulting peptides were separated using C18 reverse-phase chemistry employing a pre-column (EASY column SC001, length 2 cm, ID 100 µm, particle size 5 µm) in line with an EASY column SC200 with a length of 10 cm, an inner diameter (ID) of 75 µm and a particle size of 3 µm on an EASY-nLC II (all from Thermo Fisher Scientific). Peptides were eluted into a Nanospray Flex ion source (Thermo Fisher Scientific) with a 60 min gradient increasing from 5% to 40% acetonitrile in ddH_2O with a flow rate of 300 nL/min and electrosprayed into an Orbitrap Velos Pro mass spectrometer (Thermo Fisher Scientific). The source voltage was set to 1.9 kV, the S Lens RF level to 50%. The delta multipole offset was -7.00. The AGC target value was set to 1e06 and the maximum injection time (max IT) to 500 ms in the Orbitrap. The parameters were set to 3e04 and 50 ms in the LTQ with an isolation width of 2 Da for precursor isolation and MS/MS scanning. Peptides were analysed using a Top 10 DDA scan strategy employing HCD fragmentation with stepped collision energies (normalised collision energy 40, 3 collision energy steps, width 15). MS/MS spectra were used to search the TAIR10 database (ftp://ftp.arabidopsis.org, 35394 sequences, 14486974 residues) amended with target protein sequences with the Mascot software v.2.5 linked to Proteome Discoverer v.2.1. The enzyme specificity was set to trypsin, and two missed cleavages were tolerated. Carbamidomethylation of cysteine was set as a fixed modification and oxidation of methionine. Searches were performed with enzyme specificity set to trypsin and semi-trypsin to identify truncated protein N-termini. The precursor tolerance was set to 7 ppm, and the product ion mass tolerance was set to 0.8 Da. A decoy database search was performed to determine the peptide spectral match (PSM) and peptide identification false discovery rates (FDR). PSM, peptide and protein identifications surpassing respective FDR thresholds of $q < 0.01$ were accepted.

UPO bioconversions for subsequent GC-MS and chiral HPLC analytics. For the tested hydroxylation (naphthalene, phenylethane, -propane, -butane and -pentane) and epoxidation (styrene) reactions, purified UPOs enzyme samples (stored in 100 mM potassium phosphate; pH 7.0) produced in *S. cerevisiae* were used. Respective reactions (total volume: 400 µL) were performed as biological triplicates in 100 mM potassium phosphate (pH 7.0) containing 100 nM of UPO, 1 mM of the respective substrate and 500 µM H_2O_2. The substrate was prior dissolved in pure acetone (20 mM stock solution) yielding a 5% (v/v) co-solvent ratio in the final reaction mixture. Reactions were performed for 60 min (25 °C, 850 rpm) and subsequently quenched by the addition of 400 µl ethyl acetate (internal standard: 1 mM ethyl benzoate). Extraction was accomplished by 30 s of vigorous vortexing, followed by brief centrifugation (1 min, 8400 rpm). The organic layer was then utilised for respective GC-MS measurements as described in Supplementary Table 7. In the case of the hydroxylation reaction of *N*-phthaloyl-phenylethyl amine, purified UPOs enzyme samples (stored in 100 mM potassium phosphate, pH 7.0.) previously produced in *P. pastoris* were used. Reactions (total volume: 500 µL) were performed as biological triplicates in 100 mM potassium phosphate (pH 7.0) containing 100 nM of the respective UPO, 250 µM of the substrate *N*-phthaloyl-phenylethyl amine and 250 µM H_2O_2. The substrate was prior dissolved in pure

acetone (5 mM stock solution) yielding a 5% (v/v) co-solvent ratio in the final reaction mixture. Reactions were performed for 60 min (30 °C, 850 rpm) and subsequently quenched by the addition of 650 µL ethyl acetate (internal standard: 1 mM ethyl benzoate). Extraction was accomplished by 30 s of vigorous vortexing, followed by brief centrifugation (1 min, 8400 rpm). 200 µL of the resulting organic layer were utilised for GC-MS measurements. The remaining organic solvent was evaporated using a mild nitrogen stream, the precipitate resolved in 200 µL isopropanol and utilised for chiral HPLC measurements. For the larger scale hydroxylation reaction of *N*-phthaloyl-phenylethyl amine with *Cgl*UPO general procedures were followed as described above with some slight alterations. In contrast to the previous small-scale reaction (500 µL), within this approach, ten reactions (each total volume: 1 mL) were performed in parallel in 100 mM potassium phosphate (pH 7.0) containing 250 nM *Cgl*UPO, 250 µM substrate and 250 µM H_2O_2. Reactions were performed for 60 min (30 °C, 850 rpm) and subsequently quenched by the addition of 1 mL ethyl acetate to each reaction vial. Extraction was accomplished by 30 s of vigorous vortexing, followed by brief centrifugation (1 min, 8400 rpm). The organic layers of all samples were combined, and the solvent was gradually evaporated using a mild nitrogen stream. The precipitate was then resolved in 200 µL isopropanol and utilised for chiral HPLC measurements (Supplementary Figs. 18–20).

Achiral gas chromatography–mass spectrometry (GC–MS). Measurements were performed on a Shimadzu GCMS-QP2010 Ultra instrument (Shimadzu, Kyoto, JP) using a SH-Rxi-5Sil MS column (30 m x 0.25 mm, 0.25 µm film, Shimadzu, Kyoto, JP) or OPTIMA 5MS Accent column (25 m x 0.20 mm, 0.20 µm film, Macherey-Nagel, Düren, DE) and helium as carrier gas. 1 µl of each sample was injected splitless with an injection temperature of 280 °C. The split/splitless uniliner inlets (3.5 mm, 5.0 × 95 mm for Shimadzu GCs, deactivated wool) from Restek (Bad Homburg, DE) were utilised and regenerated if needed by CS-Chromatography (Langerwehe, DE). The temperature program was adjusted, as shown in Supplementary Table 7. The interface temperature was set to 290 °C. Ionisation was obtained by electron impact with a voltage of 70 V, and the temperature of the ion source was 250 °C. The MS is equipped with dual-stage turbomolecular pumps and a quadrupole enabling a selected ion monitoring acquisition mode (SIM mode). Calibration and quantification were implemented in SIM mode with the corresponding *m/z* traces, as shown in Supplementary Table 7. The detector voltage of the secondary electron multiplier was adjusted in relation to the tuning results with perfluorotributylamine. The GC–MS parameter was controlled with GCMS Real Time Analysis, and for data evaluation, GCMS Postrun Analysis (GCMSsolution Version 4.45, Shimadzu, Kyoto, JP) was used.

Chiral gas chromatography–mass spectrometry (GC–MS). Measurements were performed on a Shimadzu GCMS-QP2020 NX instrument (Shimadzu, Kyoto, JP) with a Lipodex E column (25 m x 0.25 mm, Macherey-Nagel, Düren, DE) and helium as carrier gas. 1 µl of each sample was injected splitless with an OPTIC-4 (Shimadzu, Kyoto, JP) injector utilising a temperature profile in the liner (35 °C, 1 °C/s to 220 °C hold 115 s). The column temperature program was adjusted as shown in Supplementary Table 7. The interface temperature was set to 200 °C. Ionisation was obtained by electron impact with a voltage of 70 V, and the temperature of the ion source was 250 °C. The MS is equipped with dual stage turbomolecular pumps and a quadrupole enabling a selected ion monitoring acquisition mode (SIM mode). Calibration and quantification were implemented in SIM mode with the corresponding *m/z* traces, as shown in Supplementary Table 7. The detector voltage of the secondary electron multiplier was adjusted in relation to the tuning results with perfluorotributylamine. The GC–MS parameters were controlled with GCMS Real Time Analysis, and for data evaluation GCMS Postrun Analysis (GCMSsolution Version 4.45, Shimadzu, Kyoto, JP) was used.

GC–MS calibration curves. For product quantification, calibration curves were created as depicted in Supplementary Fig. 15. The quantification was achieved in Scan mode (*N*-(2-hydroxy-2-phenylethyl) phthalimide) or SIM mode (all other substrates) whereby each concentration data point was measured as triplicates and correlated to an internal standard (IS). The final product concentration was adjusted in 100 mM potassium phosphate buffer (pH 7.0) with the corresponding stock solutions in acetone yielding to 5% (v/v) final co-solvent proportion in the buffer system. Extraction was achieved adding 650 µL (*N*-(2-hydroxy-2-phenylethyl)phthalimide) or 400 µL (all other substrates) of ethyl acetate (containing 1 mM of the internal standard) and vortexing for 30 s, followed by brief centrifugation (1 min, 8400 rpm). The organic layer was utilised for GC–MS measurements applying the corresponding temperature program as listed in Supplementary Table 7. For enantiomeric product identification corresponding R-enantiomer standards were utilised (Supplementary Fig. 16).

Preparative work

N-*Phthaloyl-phenylethyl amine*. Phthalic anhydride (0.59 g, 4.0 mmol), phenylethyl amine (0.51 mL, 4.0 mmol) were dissolved in dichloromethane (40 mL) at room temperature. Molecular sieves (4 Å pore diameter) and triethylamine (2.0 mL, 14.5 mmol) were added, and the reaction mixture was refluxed for 36 h. After the reaction was completed (TLC control) the mixture was filtered, and the solvent was evaporated under reduced pressure. The residue was dissolved in ethyl acetate, washed with sodium bicarbonate solution and water and dried over sodium sulphate. After filtration, the product was obtained under reduced pressure to yield 0.31 g (80%) as an orange solid. No further purification was necessary.

^{1}H-NMR (400 MHz, $CDCl_3$): δ 7.83 (dd, *J* 5.4, 3.1 Hz, 2H), 7.70 (dd, *J* 5.5, 3.0 Hz, 2H), 7.32 – 7.17 (m, 5H), 3.96 – 3.90 (m, 2H), 3.02 – 2.95 (m, 2H) ppm;

^{13}C-NMR (100 MHz, $CDCl_3$): δ 168.15, 137.99, 133.88, 132.06, 128.83, 128.53, 126.62, 123.19, 39.27, 34.60 ppm;

MS (ESI, MeOH): m/z 274.1 ($[M + Na]^+$), calcd: 251.09.

*(R,S)-2-*N*-Phthaloyl-1-phenylethanol*. Phthalic anhydride (0.30 g, 2.0 mmol) and 2-amino-1-phenylethanol (0.27 g, 2.0 mmol) were placed into a microwave vessel under stirring (magnetic). The vessel was heated to 150 °C for 30 min in the microwave reactor. After cooling to room temperature, the product was washed with HCl (1 M, 20 mL) and recrystallised from dichloromethane/n-hexane to yield 0.47 g (89%) as colourless crystals.

^{1}H-NMR (400 MHz, $CDCl_3$): δ 7.82 (dd, *J* 5.4, 3.1 Hz, 2H), 7.70 (dd, *J* 5.5, 3.0 Hz, 2H), 7.48 – 7.40 (m, 2H), 7.39 – 7.27 (m, 3H), 5.06 (dt, *J* 8.6, 4.2 Hz, 1H), 4.07 – 3.85 (m, 2H), 3.03 (d, *J* 5.0 Hz, 1H) ppm;

^{13}C-NMR (100 MHz, $CDCl_3$): δ 168.69, 141.02, 134.06, 131.81, 128.53, 128.03, 125.83, 123.39, 72.47, 45.67 ppm;

MS (ESI, MeOH): m/z 268.1 ($[M + H]^+$), 290.0 ($[M + Na]^+$), calcd: 267.09.

*(S)-(+)-2-*N*-Phthaloyl-1-phenylethanol (chemical conversion)*. Phthalic anhydride (0.30 g, 2.0 mmol) and (S)-(+)-2-amino-1-phenylethanol (0.27 g, 2.0 mmol) were placed into a microwave vessel under stirring (magnetic). The vessel was heated to 150 °C for 30 min in the microwave reactor. After cooling to room temperature, the product was washed with HCl (1 M, 20 mL) and recrystallised from dichloromethane/n-hexane to yield 0.44 g (82%) as colourless crystals.

^{1}H-NMR (400 MHz, $CDCl_3$): δ 7.85 (dd, *J* 5.5, 3.1 Hz, 2H), 7.73 (dd, *J* 5.5, 3.1 Hz, 2H), 7.50 – 7.27 (m, 5H), 5.19 – 4.96 (m, 1H), 4.10 – 3.86 (m, 2H), 2.97 – 2.78 (m, 1H) ppm;

^{13}C-NMR (100 MHz, $CDCl_3$): δ 168.75, 141.05, 134.13, 131.88, 128.60, 128.11, 125.86, 123.46, 72.68, 45.76 ppm;

MS (ESI, MeOH): m/z 289.9 ($[M + Na]^+$), calcd: 267.09;

$$[a]_{20}^{D} + 23.9(c0.75, CHCl_3).$$

*(S)-(+)-2-*N*-Phthaloyl-1-phenylethanol (enzymatic conversion)*. *N*-Phthaloyl-phenylethyl amine (15.8 mg, 62.9 µmol) was dissolved in acetone (15 mL) and poured into a solution of potassium phosphate buffer (100 mM, 263 mL, pH 7.0), hydrogen peroxide (210 µM, 3.2 mL) and *Mth*UPO (250 nM, 15 mL). The solution (total: 300 mL) was stirred at 30 °C for 1 h. Afterwards the mixture was extracted using ethyl acetate (3 × 60 ml). The organic phase was washed with brine, dried with sodium sulphate, filtered and concentrated under reduced pressure. The crude product was purified by column chromatography on silica gel using dichloromethane/ethyl acetate with 1% formic acid (1/5 → 1/1) obtaining 9.70 mg (57 %) (S)-(+)-2-*N*-Phthaloyl-1-phenylethanol as a pale-yellow solid.

^{1}H-NMR (400 MHz, $CDCl_3$): *δ* 7.86 (dd, *J* 5.5, 3.1 Hz, 2H), 7.73 (dd, *J* 5.5, 3.0 Hz, 2H), 7.49 – 7.43 (m, 2H), 7.42 – 7.27 (m, 3H), 5.08 (dd, *J* 8.7, 3.6 Hz, 1H), 4.11 – 3.86 (m, 2H), 2.83 (s, 1H) ppm;

^{13}C-NMR (100 MHz, $CDCl_3$): *δ* 168.76, 141.04, 134.14, 131.89, 128.62, 128.13, 125.86, 123.48, 72.72, 45.77 ppm;

MS (ESI, MeOH): *m/z* 289.9 ($[M + Na]^+$), calcd: 267.09;

$$[a]_{20}^{D} + 21.0(c1.55, CHCl_3).$$

N-*Phthaloyl-2-oxo-phenylethyl amine*. (R,S)-*N*-Phthaloyl-phenylethanol (0.18 g, 0.67 mmol) was dissolved in dimethyl sulfoxide (6 mL) at room temperature. Under ice cooling, acetic anhydride (1.2 mL) was added, and the reaction mixture was stirred for 16 h at room temperature. After the reaction was completed (TLC control) the mixture was quenched with ethyl acetate (20 mL), and the mixture was washed with sodium perchlorate solution (6 %), sodium thiosulfate solution (10 %) and brine and dried over sodium sulfate. After filtration, the product was obtained under reduced pressure to yield 0.15 g (84 %) as a colourless solid. No further purification was necessary.

^{1}H-NMR (400 MHz, $CDCl_3$): *δ* 8.06 – 7.98 (m, 2H), 7.91 (dd, *J* = 5.5, 3.0 Hz, 2H), 7.76 (dd, *J* = 5.5, 3.0 Hz, 2H), 7.69 – 7.48 (m, 3H), 5.14 (s, 2H) ppm;

^{13}C-NMR (100 MHz, $CDCl_3$): *δ* 190.94, 167.88, 134.43, 134.11, 134.02, 132.25, 128.89, 128.14, 123.55, 44.19 ppm;

MS (ESI, MeOH): *m/z* 288.1 ($[M + Na]^+$), calcd: 265.07.

Column and analytic thin layer chromatography. All solvents for column chromatography were purchased from Merck Millipore (Darmstadt, DE) and distilled prior to use. Column chromatography was carried out using Merck silica gel 60 (40–63 µm). For analytic thin layer chromatography, Merck TLC silica gel 60 F254 aluminium sheets were used. Compounds were visualised by using UV light (254/366 nm).

Nuclear magnetic resonance (NMR). NMR spectra were recorded using a 400 MHz Agilent DD2 400 NMR spectrometer at 25 °C. The chemical shifts of 1H NMR spectra are referenced on the signal of the internal standard tetramethylsilane ($\delta = 0.000$ ppm). Chemical shifts of 13 C NMR spectra are referenced on the solvent residual signals of $CDCl_3$ ($\delta = 77.000$ ppm).

Electrospray ionisation mass spectrometry (ESI-MS). ESI mass spectra were recorded on an API3200 Triple Quadrupole mass spectrometer (AB Sciex) equipped with an electrospray ion source (positive spray voltage 5.5 kV, negative spray voltage 4.5 kV, source heater temperature 400 °C).

Specific optical rotation. Specific optical rotations of compounds were recorded on a P-2000 Digital Polarimeter (JASCO, Pfungstadt, DE) utilising a wavelength of 589 nm.

Chiral HPLC. HPLC chromatograms were recorded on an Agilent High Performance LC (Agilent Technologies, Waldbronn, DE). The used chiral column material was Chiralpak AS-H HPLC (Daicel, Tokyo, JP) (25 cm × 4.6 mm). Substances were dissolved in HPLC-grade isopropanol prior to analysis, and a sample volume of 5 µL injected. The eluent (20% isopropanol, 80% *n*-hexane) was used in a flow rate of 1 mL/min with the runtime of 30 min at 30 °C.

Microwave reactions. Microwave reactions were carried out using an Initiator + device (Biotage, Düsseldorf, DE).

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The authors declare that the data supporting the findings of this study are available within the paper and its Supplementary Information files. Source data is provided as Supplementary Data 1.

Received: 28 September 2020; Accepted: 31 March 2021;
Published online: 12 May 2021

References

1. Ni, Y. et al. Peroxygenase-catalyzed oxyfunctionalization reactions promoted by the complete oxidation of methanol. *Angew. Chem. Int Ed. Engl.* **55**, 798–801 (2016).
2. Ullrich, R., Nuske, J., Scheibner, K., Spantzel, J. & Hofrichter, M. Novel haloperoxidase from the agaric basidiomycete *Agrocybe aegerita* oxidizes aryl alcohols and aldehydes. *Appl. Environ. Microbiol.* **70**, 4575–4581 (2004).
3. Zhang, W. et al. Selective aerobic oxidation reactions using a combination of photocatalytic water oxidation and enzymatic oxyfunctionalisations. *Nat. Catal.* **1**, 55–62 (2018).
4. Wang, Y., Lan, D., Durrani, R. & Hollmann, F. Peroxygenases en route to becoming dream catalysts. What are the opportunities and challenges? *Curr. Opin. Chem. Biol.* **37**, 1–9 (2017).
5. Wang, Y., Lan, D., Durrani, R. & Hollmann, F. Peroxygenases en route to becoming dream catalysts. What are the opportunities and challenges? *Curr. Opin. Chem. Biol.* **37**, 1–9 (2017).
6. Hobisch, M. et al. Recent developments in the use of peroxygenases—exploring their high potential in selective oxyfunctionalisations. *Biotechnol. Adv.* https://doi.org/10.1016/j.biotechadv.2020.107615 (2020).
7. Sigmund, M.-C. & Poelarends, G. J. Current state and future perspectives of engineered and artificial peroxygenases for the oxyfunctionalization of organic molecules. *Nat. Catal.* **3**, 690–702 (2020).
8. Ullrich, R. et al. Side chain removal from corticosteroids by unspecific peroxygenase. *J. Inorg. Biochem.* **183**, 84–93 (2018).
9. Kiebist, J. et al. A peroxygenase from *Chaetomium globosum* catalyzes the selective oxygenation of testosterone. *Chembiochem* **18**, 563–569 (2017).
10. Molina-Espeja, P. et al. Directed evolution of unspecific peroxygenase from *Agrocybe aegerita*. *Appl. Environ. Microbiol.* **80**, 3496–3507 (2014).
11. Martin-Diaz, J., Paret, C., Garcia-Ruiz, E., Molina-Espeja, P. & Alcalde, M. Shuffling the neutral drift of unspecific peroxygenase in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* https://doi.org/10.1128/AEM.00808-18 (2018).
12. Molina-Espeja, P. et al. Synthesis of 1-Naphthol by a natural peroxygenase engineered by directed evolution. *Chembiochem* **17**, 341–349 (2016).
13. Gomez de Santos, P. et al. Selective synthesis of the human drug metabolite 5′-hydroxypropranolol by an evolved self-sufficient peroxygenase. *ACS Catal.* **8**, 4789–4799 (2018).
14. Gomez De Santos, P. et al. Benchmarking of laboratory evolved unspecific peroxygenases for the synthesis of human drug metabolites. *Tetrahedron* **75**, 1827–1831 (2019).
15. Molina-Espeja, P., Ma, S., Mate, D. M., Ludwig, R. & Alcalde, M. Tandem-yeast expression system for engineering and producing unspecific peroxygenase. *Enzym. Micro. Technol.* **73-74**, 29–33 (2015).
16. Rapoport, T. A., Jungnickel, B. & Kutay, U. Protein transport across the eukaryotic endoplasmic reticulum and bacterial inner membranes. *Annu. Rev. Biochem.* **65**, 271–303 (1996).
17. Williams, E. J. B., Pal, C. & Hurst, L. D. The molecular evolution of signal peptides. *Gene* **253**, 313–322 (2000).
18. Rakestraw, J. A., Sazinsky, S. L., Piatesi, A., Antipov, E. & Wittrup, K. D. Directed evolution of a secretory leader for the improved expression of heterologous proteins and full-length antibodies in *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* **103**, 1192–1201 (2009).
19. Mateljak, I., Tron, T. & Alcalde, M. Evolved alpha-factor prepro-leaders for directed laccase evolution in *Saccharomyces cerevisiae*. *Micro. Biotechnol.* **10**, 1830–1836 (2017).
20. Vina-Gonzalez, J., Elbl, K., Ponte, X., Valero, F. & Alcalde, M. Functional expression of aryl-alcohol oxidase in *Saccharomyces cerevisiae* and *Pichia pastoris* by directed evolution. *Biotechnol. Bioeng.* **115**, 1666–1674 (2018).
21. Vina-Gonzalez, J., Gonzalez-Perez, D., Ferreira, P., Martinez, A. T. & Alcalde, M. Focused directed evolution of aryl-alcohol oxidase in *Saccharomyces cerevisiae* by using chimeric signal peptides. *Appl. Environ. Microbiol.* **81**, 6451–6462 (2015).
22. González-Benjumea, A. et al. Fatty acid epoxidation by *Collariella virescens* peroxygenase and heme-channel variants. *Catal. Sci. Technol.* **10**, 717–725 (2020).
23. Linde, D. et al. Two new unspecific peroxygenases from heterologous expression of fungal genes in *Escherichia coli*. *Appl. Environ. Microbiol.* **86**, e02899–02819 (2020).
24. Engler, C., Gruetzner, R., Kandzia, R. & Marillonnet, S. Golden Gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. *PLoS One* **4**, e5553, (2009).
25. Engler, C., Kandzia, R. & Marillonnet, S. A one pot, one step, precision cloning method with high throughput capability. *PLoS One* **3**, e3647 (2008).
26. Werner, S., Engler, C., Weber, E., Gruetzner, R. & Marillonnet, S. Fast track assembly of multigene constructs using Golden Gate cloning and the MoClo system. *Bioeng. Bugs* **3**, 38–43 (2012).
27. Sarrion-Perdigones, A. et al. GoldenBraid: an iterative cloning system for standardized assembly of reusable genetic modules. *PLoS One* **6**, e21622 (2011).
28. Andreou, A. I. & Nakayama, N. Mobius assembly: a versatile Golden-Gate framework towards universal DNA assembly. *PLOS ONE* **13**, e0189892 (2018).
29. Pollak, B. et al. Loop assembly: a simple and open system for recursive fabrication of DNA circuits. *N. Phytol.* **222**, 628–640 (2019).
30. van Dolleweerd, C. J. et al. MIDAS: a modular DNA assembly system for synthetic biology. *ACS Synth. Biol.* **7**, 1018–1029 (2018).
31. Potapov, V. et al. Comprehensive profiling of four base overhang ligation fidelity by T4 DNA ligase and application to DNA assembly. *ACS Synth. Biol.* **7**, 2665–2674 (2018).
32. Püllmann, P. et al. Golden mutagenesis: an efficient multi-site-saturation mutagenesis approach by Golden Gate cloning with automated primer design. *Sci. Rep.* **9**, 10932 (2019).
33. Weber, E., Engler, C., Gruetzner, R., Werner, S. & Marillonnet, S. A modular cloning system for standardized assembly of multigene constructs. *PLoS One* **6**, e16765 (2011).
34. Cabantous, S. & Waldo, G. S. In vivo and in vitro protein solubility assays using split GFP. *Nat. Methods* **3**, 845–854 (2006).
35. Santos-Aberturas, J., Dorr, M., Waldo, G. S. & Bornscheuer, U. T. In-depth high-throughput screening of protein engineering libraries by split-GFP direct crude cell extract data normalization. *Chem. Biol.* **22**, 1406–1414 (2015).
36. Gonzalez-Perez, D., Molina-Espeja, P., Garcia-Ruiz, E. & Alcalde, M. Mutagenic organized recombination process by homologous IN vivo grouping (MORPHING) for directed enzyme evolution. *PLoS One* **9**, e90919 (2014).
37. Poraj-Kobielska, M., Kinne, M., Ullrich, R., Scheibner, K. & Hofrichter, M. A spectrophotometric assay for the detection of fungal peroxygenases. *Anal. Biochem.* **421**, 327–329 (2012).
38. Schmidt, T. G. M. & Skerra, A. The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nat. Protoc.* **2**, 1528–1535 (2007).
39. Hochuli, E., Bannwarth, W., Döbeli, H., Gentz, R. & Stüber, D. Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. *Nat. Biotechnol.* **6**, 1321–1325 (1988).
40. Schmidt, T. G. M. et al. Development of the Twin-Strep-tag® and its application for purification of recombinant proteins from cell culture supernatants. *Protein Expr. Purif.* **92**, 54–61 (2013).

41. Gröbe, G. et al. High-yield production of aromatic peroxygenase by the agaric fungus *Marasmius rotula*. *AMB Express* **1**, 31–31 (2011).
42. Babot, E. D., del Rio, J. C., Kalum, L., Martinez, A. T. & Gutierrez, A. Oxyfunctionalization of aliphatic compounds by a recombinant peroxygenase from *Coprinopsis cinerea*. *Biotechnol. Bioeng.* **110**, 2323–2332 (2013).
43. Berka, R. M. et al. Comparative genomic analysis of the thermophilic biomass-degrading fungi *Myceliophthora thermophila* and *Thielavia terrestris*. *Nat. Biotechn.* **29**, 922–927 (2011).
44. Prielhofer, R. et al. GoldenPiCS: a Golden Gate-derived modular cloning system for applied synthetic biology in the yeast *Pichia pastoris*. *BMC Syst. Biol.* **11**, 123 (2017).
45. Obst, U., Lu, T. K. & Sieber, V. A modular toolkit for generating *Pichia pastoris* secretion libraries. *ACS Synth. Biol.* **6**, 1016–1025 (2017).
46. Ahmad, M., Hirz, M., Pichler, H. & Schwab, H. Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. *Appl. Microbiol. Biotechnol.* **98**, 5301–5317 (2014).
47. Camattari, A. et al. Characterization of a panARS-based episomal vector in the methylotrophic yeast *Pichia pastoris* for recombinant protein production and synthetic biology applications. *Microb. Cell Factories* **15**, 139 (2016).
48. Yang, J. et al. Hygromycin-resistance vectors for gene expression in *Pichia pastoris*. *Yeast* **31**, 115–125 (2014).
49. Gu, Y. et al. Construction of a series of episomal plasmids and their application in the development of an efficient CRISPR/Cas9 system in *Pichia pastoris*. *World J. Microbiol. Biotechnol.* **35**, 79 (2019).
50. Vogl, T. et al. A toolbox of diverse promoters related to methanol utilization: functionally verified parts for heterologous pathway expression in *Pichia pastoris*. *ACS Synth. Biol.* **5**, 172–186 (2016).
51. Clare, J. J., Rayment, F. B., Ballantine, S. P., Sreekrishna, K. & Romanos, M. A. High-level expression of tetanus toxin fragment C in *Pichia Pastoris* strains containing multiple tandem integrations of the gene. *Nat. Biotechnol.* **9**, 455–460 (1991).
52. Clare, J. J. et al. Production of mouse epidermal growth factor in yeast: high-level secretion using *Pichia pastoris* strains containing multiple gene copies. *Gene* **105**, 205–212 (1991).
53. Ahmad, M., Hirz, M., Pichler, H. & Schwab, H. Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. *Appl. Microbiol. Biotechnol.* **98**, 5301–5317 (2014).
54. Schäfer, B. *Naturstoffe der chemischen Industrie* (Elsevier, 2007).
55. Knorrscheidt, A. et al. Identification of novel unspecific peroxygenase chimeras and unusual YfeX axial heme ligand by a versatile high-throughput GC–MS approach. *ChemCatChem* **12**, 4788–4795 (2020).
56. Delic, M. et al. The secretory pathway: exploring yeast diversity. *FEMS Microbiol. Rev.* **37**, 872–914 (2013).
57. Vieira Gomes, A. M., Souza Carmo, T., Silva Carvalho, L., Mendonça Bahia, F. & Parachin, N. S. Comparison of yeasts as hosts for recombinant protein production. *Microorganisms* **6**, 38 (2018).
58. Ramirez-Escudero, M. et al. Structural insights into the substrate promiscuity of a laboratory-evolved peroxygenase. *ACS Chem. Biol.* **13**, 3259–3268 (2018).
59. Piontek, K. et al. Structural basis of substrate conversion in a new aromatic peroxygenase: cytochrome P450 functionality with benefits. *J. Biol. Chem.* **288**, 34767–34776 (2013).
60. Molina-Espeja, P. et al. Synthesis of 1-Naphthol by a natural peroxygenase engineered by directed evolution. *ChemBioChem* **17**, 341–349 (2016).
61. Wu, Z. et al. Signal peptides generated by attention-based neural networks. *ACS Synth. Biol.* **9**, 2154–2161 (2020).
62. Vogl, T. et al. Engineered bidirectional promoters enable rapid multi-gene co-expression optimization. *Nat. Commun.* **9**, 3589 (2018).
63. Anh, D. H. et al. The coprophilous mushroom *Coprinus radians* secretes a haloperoxidase that catalyzes aromatic peroxygenation. *Appl. Environ. Microbiol.* **73**, 5477–5485 (2007).
64. Knorrscheidt, A. et al. Accessing Chemo- and Regioselective Benzylic and Aromatic Oxidations by Protein Engineering of an Unspecific Peroxygenase. *ChemRxiv* https://doi.org/10.26434/chemrxiv.13265618.v1 (2020).
65. Churakova, E. et al. Specific photobiocatalytic oxyfunctionalization reactions. *Angew. Chem. Int. Ed.* **50**, 10716–10719 (2011).
66. Rauch, M. C. R. et al. Peroxygenase-catalysed epoxidation of styrene derivatives in Neat Reaction Media. *ChemCatChem* **11**, 4519–4523 (2019).
67. Kluge, M., Ullrich, R., Scheibner, K. & Hofrichter, M. Stereoselective benzylic hydroxylation of alkylbenzenes and epoxidation of styrene derivatives catalyzed by the peroxygenase of *Agrocybe aegerita*. *Green. Chem.* **14**, 440–446 (2012).
68. Yamaguchi, S., Suzuki, A., Togawa, M., Nishibori, M. & Yahiro, H. Selective oxidation of thioanisole with hydrogen peroxide using copper complexes encapsulated in zeolite: formation of a thermally stable and reactive copper hydroperoxo species. *ACS Catal.* **8**, 2645–2650 (2018).
69. May, S. W. & Phillips. Robert, S. Asymmetric sulfoxidation by dopamine. beta.-hydroxylase, an oxygenase heretofore considered specific for methylene hydroxylation. *J. Am. Chem. Soc.* **102**, 5981–5983 (1980).
70. Knorrscheidt, A. et al. Simultaneous Screening of Multiple Substrates with an Unspecific Peroxygenase Enabled Modified Alkane and Alkene Oxyfunctionalisations. *Cat. Sci. Technol.* (2021), in press.
71. Püllmann, P. & Weissenborn, M. J. Improving the heterologous production of fungal peroxygenases through an episomal Pichia pastoris promoter and signal peptide shuffling system. *bioRxiv* https://doi.org/10.1101/2020.12.23.424034 (2020).
72. Scheler, U. et al. Elucidation of the biosynthesis of carnosic acid and its reconstitution in yeast. *Nat. Commun.* **7**, 12942 (2016).
73. Liachko, I. & Dunham, M. J. An autonomously replicating sequence for use in a wide range of budding yeasts. *FEMS Yeast Res.* **14**, 364–367 (2014).
74. Lin-Cereghino, J. et al. Condensed protocol for competent cell preparation and transformation of the methylotrophic yeast *Pichia pastoris*. *Biotechniques* **38**, 44–48 (2005).
75. Molina-Espeja, P. et al. Directed evolution of unspecific peroxygenase from *Agrocybe aegerita*. *Appl. Environ. Microbiol.* **80**, 3496–3507 (2014).
76. Guengerich, F. P., Martin, M. V., Sohl, C. D. & Cheng, Q. Measurement of cytochrome P450 and NADPH–cytochrome P450 reductase. *Nat. Protoc.* **4**, 1245–1251 (2009).
77. Majovsky, P. et al. Targeted proteomics analysis of protein degradation in plant signaling on an LTQ-Orbitrap mass spectrometer. *J. Proteome Res.* **13**, 4246–4258 (2014).

Acknowledgements

M.J.W and A.K. thank the Bundesministerium für Bildung und Forschung ("Biotechnologie 2020+ Strukturvorhaben: Leibniz Research Cluster", 031A360B) for generous funding. P.P. thanks the Landesgraduiertenförderung Sachsen-Anhalt for a PhD scholarship. J.M. thanks the Friedrich-Naumann-Stiftung for a PhD scholarship. Prof. Jürgen Pleiß (University of Stuttgart), Prof. Dirk Holtmann and Sebastian Bormann (DECHEMA Frankfurt) are kindly acknowledged for sharing putative and described UPO gene sequences. We would like to thank Cătălin Voiniciuc (Leibniz Institute of Plant Biochemistry Halle) for fruitful discussions regarding protein production in *Pichia pastoris* and furthermore providing Pichia plasmid parts and the X33-strain. Prof. Karin Breunig (Martin Luther University Halle-Wittenberg) is kindly acknowledged for generously supplying genomic DNA of *K. lactis*. We would like to especially thank Anja Ehrlich (Leibniz Institute of Plant Biochemistry Halle) for outstanding technical support and patience regarding chiral HPLC analyis of phenethylamine conversions. Furthermore, Prof. Markus Pietzsch and Dr. Franziska Seifert (Martin Luther University Halle-Wittenberg) are kindly acknowledged for discussions and providing access to the DSF device for thermostability measurements. Our sincere gratitude also goes to Prof. Martin Hofrichter and Dr. Harald Kellner (Technical University of Dresden) for discussions and providing the protein sequence of *Mwe*UPO. Dr. Swanhild Lohse (IPB Halle) is acknowledged for the initial construction of the utilised pAGT572 backbone structure of episomal *S. cerevisiae* plasmids.

Author contributions

P.P. and M.J.W. designed the research. P.P. performed all experiments apart from the enzymatic conversions in Fig. 3 (performed by A.K.), Fig. 5 (supported by J.M. and P.R.P. and co-designed by B.W.) and the protein identification by MS (performed by W.H.). P.P. and S.M. designed the modular Golden Gate yeast system and M.A. developed the underlying 96-well *S. cerevisiae* expression system. P.P. and M. J. W. wrote the manuscript. All authors contributed to the proofreading of the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

Evolved *Aae*UPO* enzyme used in the current study is protected by CSIC patent WO/2017/081355 (licensed in exclusivity to EvoEnzyme S.L). M.A. is co-founder and advisor of EvoEnzyme S.L. The authors declare no further competing interest.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42003-021-02076-3.

Correspondence and requests for materials should be addressed to M.J.W.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

5 | Chapter III

This chapter has been published as:

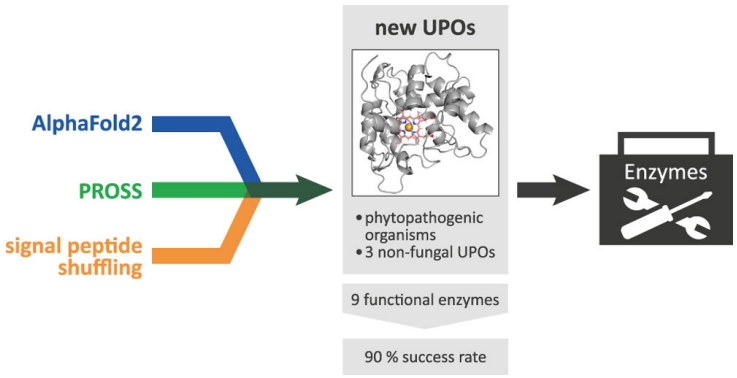
Functionally Diverse Peroxygenases by AlphaFold2, Design, and Signal Peptide Shuffling

by: **Judith Münch***, Niklas Dietz*, Shiran Barber-Zucker, Franziska Seifert, Susanne Matschi, Pascal Püllmann, Sarel J. Fleishman, and Martin J. Weissenborn

*shared first authorship

in: *ACS Catalysis* **2024**, *14* (7), 4738–4738; doi: https://doi.org/10.1021/acscatal.4c00883

Supplementary information can be assessed free of charge at the article's landing page.



This chapter presents a methodology for expressing and secreting novel unspecific peroxygenases (UPOs), characterized by unknown structures and initially lacking secretion. The workflow consists of the application of an algorithm called PROSS to increase protein stability to AlphaFold2 model structures of 10 unique UPOs, followed by a signal peptide shuffling to facilitate heterologous production. Screening of the enzyme variants revealed nine functionally active UPOs, all produced in the heterologous host *P. pastoris*, among them three UPOs from oomycetes, which are the first non-fungal UPOs to be experimentally characterized. This workflow shows a successful combination of computational (AlphaFold2 and PROSS) and laboratory (signal peptide shuffling) methods to enable enzyme engineering. Additionally, the high accuracy and reliability demonstrated by the new modeling and design methods underscore their potential to considerably expand the enzyme pool available for both research and practical applications.

ACS Catalysis

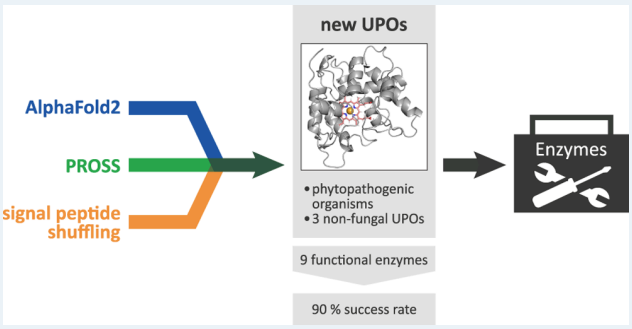
pubs.acs.org/acscatalysis
Research Article

Functionally Diverse Peroxygenases by AlphaFold2, Design, and Signal Peptide Shuffling

Judith Münch, Niklas Dietz, Shiran Barber-Zucker, Franziska Seifert, Susanne Matschi, Pascal Püllmann, Sarel J. Fleishman,* and Martin J. Weissenborn*

Cite This: ACS Catal. 2024, 14, 4738–4748
Read Online

ABSTRACT: Unspecific peroxygenases (UPOs) are fungal enzymes that attract significant attention for their ability to perform versatile oxyfunctionalization reactions using H_2O_2. Unlike other oxygenases, UPOs do not require additional reductive equivalents or electron transfer chains that complicate basic and applied research. Nevertheless, UPOs generally exhibit low to no heterologous production levels and only four UPO structures have been determined to date by crystallography limiting their usefulness and obstructing research. To overcome this bottleneck, we implemented a workflow that applies PROSS stability design to AlphaFold2 model structures of 10 unique and diverse UPOs followed by a signal peptide shuffling to enable heterologous production. Nine UPOs were functionally produced in *Pichia pastoris*, including the recalcitrant *Cci*UPO and three UPOs derived from oomycetes—the first nonfungal UPOs to be experimentally characterized. We conclude that the high accuracy and reliability of new modeling and design workflows dramatically expand the pool of enzymes for basic and applied research.



KEYWORDS: *unspecific peroxygenase, yeast, Pichia pastoris, enzyme design, heterologous expression, protein stability*

INTRODUCTION

Unspecific peroxygenases (UPOs) are secreted, fungal enzymes that have attracted great interest in recent years.[1−3] They perform versatile oxyfunctionalization reactions like hydroxylations and epoxidations on a broad substrate scope.[4] In contrast to P450 monooxygenases, UPOs do not rely on molecular oxygen, NAD(P)H, and electron transport chains but solely require hydrogen peroxide as a cosubstrate, in which the oxygen is already prereduced.[5−8] They are divided into two major groups: long- and short-type UPOs (group I and II) of approximately 45 and 29 kDa.

The stability of UPOs, their substrate scope,[4] and turnover numbers (TONs) of up to 900,000 strongly favor them for industrial application.[9] So far, the main limitation to using UPOs has been their poor and difficult heterologous production. Although databases contain sequence information on >4000 putative UPOs,[4,8] only about 50 have been heterologously produced since the discovery of this family in 2004.[10−21] Thus, selecting UPOs for fundamental and applied research is dominated by considerations of heterologous production, limiting the phylogenetic and functional scope of enzymes that have been subjected to research.

As extracellular enzymes, UPOs require an *N*-terminal signal peptide that coordinates cell trafficking and secretion. The importance of the signal peptide choice regarding the quantity of the secreted target protein has been demonstrated in a previous work: An *Aae*UPO variant (hereinafter: PaDa-I), which was evolved for higher heterologous production, exhibits nine mutations compared to the wild type, four of them are within the signal peptide. The mutations in the signal peptide alone resulted in a 27-fold improvement in functional production in *Saccharomyces cerevisiae*.[15] Building on this work, our lab pursued an approach that focused on rapidly testing a signal peptide panel derived from yeast organisms, basidiomycetes, ascomycetes, and animals rather than evolving the natural signal peptide. With that approach, the production level in *S. cerevisiae* of the *Aae*UPO variant PaDa-I could be further doubled compared to the evolved signal peptide.[18] This signal peptide shuffling technique further enabled the production of six other UPOs.[18] Further increase in heterologous production was achieved through a combined shuffling of a promotor and signal peptide library in *Pichia pastoris* (syn. *Komagataella phaffii*).[19]

In addition to poor secretion, low heterologous production could be due to marginal protein stability.[22] Production under

Received: February 7, 2024
Accepted: February 28, 2024
Published: March 13, 2024

https://doi.org/10.1021/acscatal.4c00883
ACS Catal. 2024, 14, 4738−4748

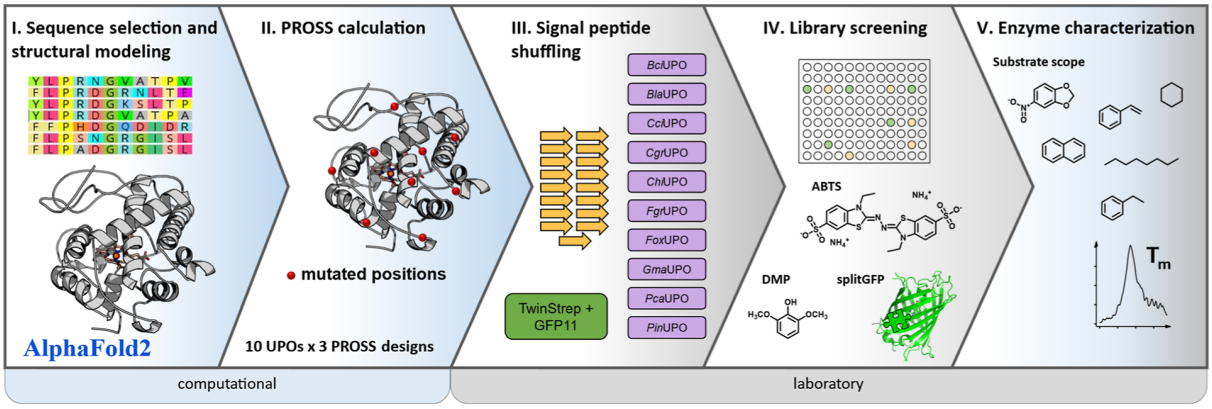


Figure 1. Overview of the different steps of the work protocol.

non-native conditions such as overexpression or the use of heterologous expression systems can lead to improperly folded proteins or aggregation[23] resulting in low protein yields. Increasing native-state stability may consequently lead to improved yields. Campaigns to enhance stability and production levels by introducing beneficial mutations based on directed evolution have been successful.[24,25] However, the high labor intensity of this iterative approach renders it challenging to use when multiple enzymes are targeted. Furthermore, these approaches require detectable starting levels of secretion and activity in the relevant host, but some UPOs exhibit none. To address such challenges in heterologous production, in recent years, several algorithms have been developed to design stabilizing mutations.[26−29] Among them is the Protein Repair One-Stop Shop (PROSS) algorithm.[23,30] PROSS combines phylogenetic analysis with Rosetta atomistic calculations to design multipoint mutants with a favorable native-state energy. In dozens of previous studies, stability increases were accompanied by gains in functional production levels after a single design calculation and experimental screening of 3−5 constructs per protein target.[31,32] These studies include, among others, the successful design of challenging oxygenases, such as high-redox potential laccases and versatile peroxidases (VPs).[31,32]

Until recently, however, PROSS was limited only to the small fraction of enzymes for which crystallographic structures are available. With only four experimentally determined UPO structures, this class of enzymes is not amenable to atomistic design calculations.[17,33−35] In a previous study dedicated to improve the functional expression of VPs using trRosetta, a legacy AI-based structure predictor, and PROSS, three of 11 enzymes that could not be functionally expressed in yeast before were functionally produced and characterized.[31]

The dramatic recent improvement in AI-based ab initio structure prediction methods such as AlphaFold2 has enabled predictions that are almost as accurate as those obtained from X-ray crystallography.[36,37] These groundbreaking developments could enable the use of AlphaFold2 to provide structures for PROSS designs and empower researchers to unleash the full potential of improving protein production and stability directly from sequence with no recourse to experimental data. In the current study, we combine the much more accurate AlphaFold2 predictor with PROSS and signal peptide shuffling to 10 UPOs, achieving unprecedented levels of success with nine enzymes exhibiting functional heterologous expression compared with only one of the 10 wild-type enzymes exhibiting limited functional expression (Figure 1).

RESULTS AND DISCUSSION

Sequence Selection, Design Calculations, and Library Construction. We chose eight short-type UPOs derived from phytopathogens (Table 1), as UPOs from phytopathogenic fungi may be involved in coping with plant defensive compounds.[38,39] Gaining access to those enzymes could prove to be a crucial step in understanding their role in the natural environment. Three of these UPOs originate from nonfungal oomycetes, which are more closely related to algae than fungi, further expanding the scope and novelty of our study as these would be the first functionally characterized UPOs from nonfungal origin.[40,41] In addition to these short-type UPOs, we subjected two challenging long-type UPOs derived from basidiomycetes to prove the generality of the protocol (Figure 2). In previous work, we used signal peptide shuffling to successfully produce an UPO derived from *Galerina marginata* in *S. cerevisiae*; yet, the recombinant enzyme did not exhibit detectable activity.[18] Another UPO derived from *G. marginata*, which exhibits 94% sequence identity to the one we selected, was shown recently (after the start of our work) to be functionally expressed in *P. pastoris*.[14] Throughout this work, we use the name *Gma*UPO referring to the first published *Gma*UPO.[18] *Cci*UPO has so far only been produced in *Aspergillus oryzae*[16,42] and not in any fast-growing microbe but holds substantial scientific interest for its relatively high activity levels.[11,43−47]

Table 1. Overview of the UPOs Studied in This Work

enzyme	original organism	type	division/class
*Bci*UPO	*Botrytis cinerea*[a]	short-type	Ascomycota
*Bla*UPO	*Bremia lactucae*[a]	short-type	Oomycota
*Cci*UPO	*Coprinopsis cinerea*	long-type	Basidiomycota
*Cgr*UPO	*Colletotrichum graminicola*[a]	short-type	Ascomycota
*Chi*UPO	*Colletotrichum higginsianum*[a]	short-type	Ascomycota
*Fgr*UPO	*Fusarium graminearum*[a]	short-type	Ascomycota
*Fox*UPO	*Fusarium oxysporum*[a]	short-type	Ascomycota
*Gma*UPO	*Galerina marginata*	long-type	Basidiomycota
*Pca*UPO	*Phytophthora cactorum*[a]	short-type	Oomycota
*Pin*UPO	*Phytophthora infestans*[a]	short-type	Oomycota

[a]Phytopathogenic organism.

https://doi.org/10.1021/acscatal.4c00883
ACS Catal. 2024, 14, 4738−4748

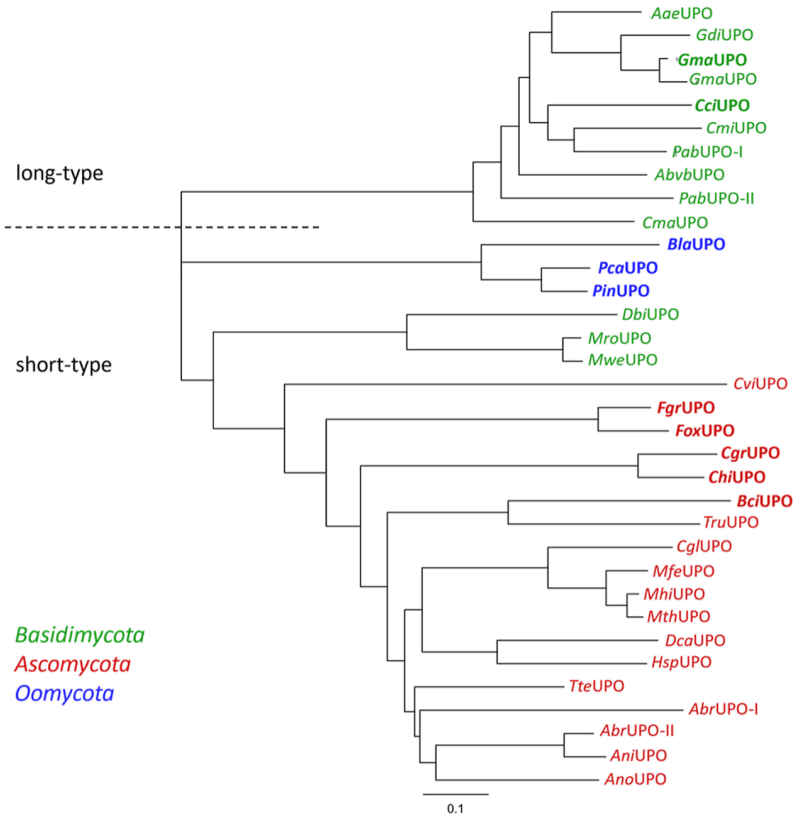


Figure 2. Neighbor-joining phylogenetic tree of selected previously known and heterologously produced UPOs (before September 2023) and our new UPO targets using Jukes−Cantor genetic distances. Basidiomycota (green), Ascomycota (red), and Oomycota (blue) (Table S3). UPOs in bold font were examined in this work. The dotted lines separate UPO sequences of group I and II (short and long UPOs). Generated with Geneious Prime 2023.1.2 (Biomatters Inc., Auckland, New Zealand).

The ten selected UPO sequences were modeled using AlphaFold2 (AF2; October 2021 release) and relaxed via the AF2 suite using AMBER. All models show high predicted reliability (average pLDDT score >90% for the best model). The model with the highest pLDDT score out of five calculated models for each target was subjected to the PROSS stability design.

PROSS combines phylogenetic analysis and Rosetta atomistic calculations to restrict design choices at each position according to their likelihood to occur in evolution and to stabilize the native state. In the final step, Rosetta combinatorial design is used to compute up to nine designs per starting structure with varying numbers of stabilizing mutations. As AF2 does not model ligands, we used the previously determined structures of *Mro*UPO (pdb:5FUK) and PaDa-I (pdb:5OXU) as templates for determining the positions involved in the essential heme and magnesium (Mg^{2+})[48] binding. All AF2 models aligned well with the reference crystal structures (long UPOs root-mean-square deviation (RMSD) to 5OXU: <0.5 Å; short UPOs RMSD to 5FUK: <1.1 Å). Furthermore, the amino acids that coordinate heme and Mg^{2+} exhibit similar constellations as those in the reference crystallographic structures (Figure S18). These amino acids were not allowed to mutate or change conformations during design calculations. The design was further restricted in the substrate tunnel, putative *N*-glycosylation sites, regions that structurally vary between the five calculated AF2 models, regions with low AF2-predicted accuracy (pLDDT <90% or 5 Å from these residues), and segments that are aligned to only a few sequence homologues. Finally, three designs for each enzyme were selected with different mutational loads of low (7−13 mutations, on average 4% of the sequence), medium (14−22 mutations, on average 7% of the sequence), and high (21−34 mutations, on average 11% of the sequence) for experimental testing (Table S1).

Seventeen different signal peptides[18] were fused to each of the 30 enzyme variants (three PROSS designs for each of the 10 UPO enzymes) leading to 510 different designs that were analyzed. All signal peptides and utilized vectors are available via Addgene (Yeast Secrete and Detect Kit #1000000166).

Functional Expression of Stabilized UPOs. Each variant was analyzed for its activity toward the two peroxidase substrates: (i) ABTS (2,2′-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid)) and (ii) DMP (2,6-dimethoxyphenol). Additionally, secretion was detected via a split-GFP assay[49] (Figures S1−S10).

PROSS designs of nine of the 10 UPO targets showed detectable secretion via the split-GFP assay (Figure 3). Of these nine, seven were active toward at least one of the substrates under screening conditions with five displaying activity toward ABTS and seven toward DMP (Figure 3). Solely, *Fox*UPO exhibited no secretion or activity.

Secretion and activity greatly varied among the PROSS variants. For *Cgr*UPO and *Fgr*UPO, the highest secretion levels were seen in the PROSS designs with the lowest mutational load. For *Bci*UPO, *Cci*UPO, *Chi*UPO, and *Pin*UPO, the highest secretion levels were achieved with a medium mutational load. *Bla*UPO, *Gma*UPO, and *Pca*UPO showed

https://doi.org/10.1021/acscatal.4c00883
ACS Catal. 2024, 14, 4738−4748

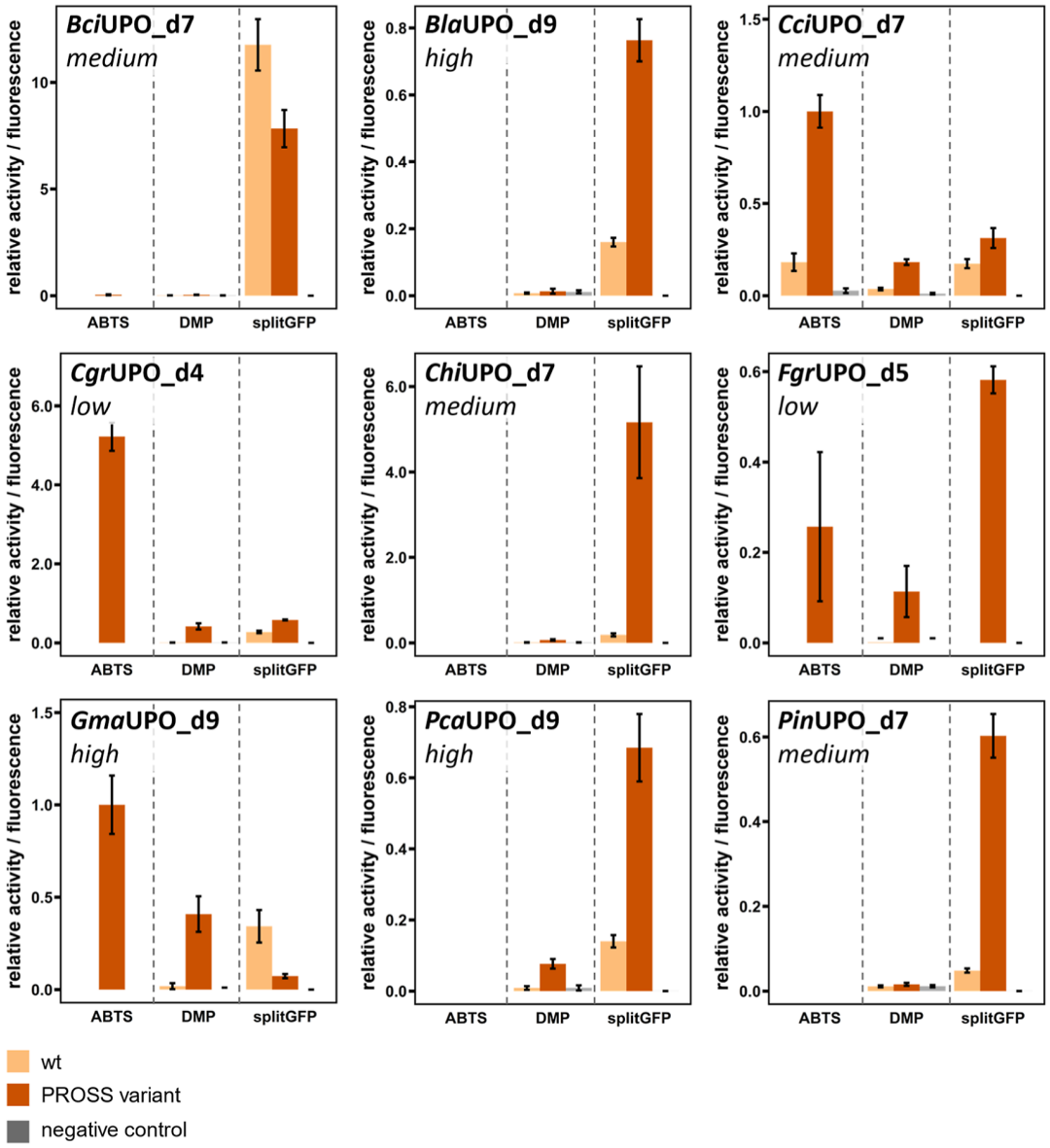


Figure 3. Comparison of the best PROSS design (orange) of each UPO with its corresponding wild type (yellow), expressed with the same signal peptide under screening conditions. Empty vector as the negative control (gray). Display of all three screening assays ABTS, DMP, and split-GFP. For each assay, the values are standardized to the positive control (*Mth*UPO, not displayed). Only the ABTS assays of *Cci*UPO_d7 and *Gma*UPO_d9 were standardized to the PROSS variant. The mutational load of the chosen PROSS variant (low, medium, or high) is given in *italics*. Data are mean ± SD. Measurements were performed in general with eight replicates (n = 8), n = 6 in the case of *Cgr*UPO_d4, and n = 12 in the case of *Chi*UPO_d7.

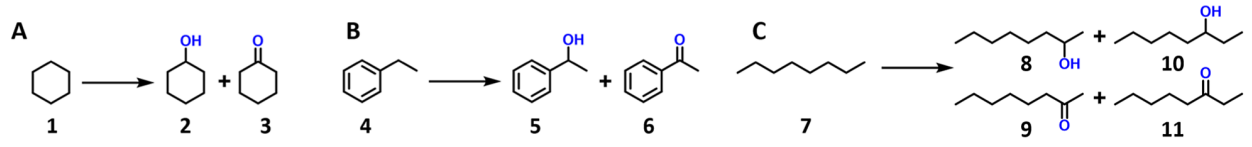
the highest secretion levels with the design carrying the highest number of mutations (Tables S1 and S2). Thus, as observed in previous studies, the number of mutations introduced during design is only weakly correlated with the observed improvement in production levels.[50]

We also observed a huge variation in secreted enzyme levels, depending on the signal peptide. This variation can be demonstrated, for instance, by *Chi*UPO secretion relative to the positive control (*Mth*UPO). The best *Chi*UPO PROSS variant (*Chi*UPO_d7, 19 mutations, medium mutational load) led to three different signal peptide-dependent secretion levels: (i) within a range of 100−300% of the fluorescence signal compared to the positive control, (ii) between 20 and 60% of the positive control signal, and (iii) below 10% of the positive control fluorescence signal (Figure S1). This observation further emphasizes the enormous potential of signal peptide shuffling.[18,19]

The best-performing signal peptide varied among the different UPOs. During the screening of *Gma*UPO and *Cci*UPO variants, we confirmed that long-type UPO signal peptides were generally more prevalent for these long-type UPOs.[18,19] *Gma*UPO showed the highest secretion level and activity with its natural signal peptide. In contrast, short-type UPOs did not show a preference for any signal peptide group, as previously observed.[18,19] To reduce the screening effort of future studies of long-type UPOs, our results recommend using a reduced library containing only long-type UPO signal peptides. The diversity in signal peptides and differing mutational loads for the different PROSS variants indicates the strength of the combined approach, as neither the most

https://doi.org/10.1021/acscatal.4c00883
ACS Catal. 2024, 14, 4738−4748

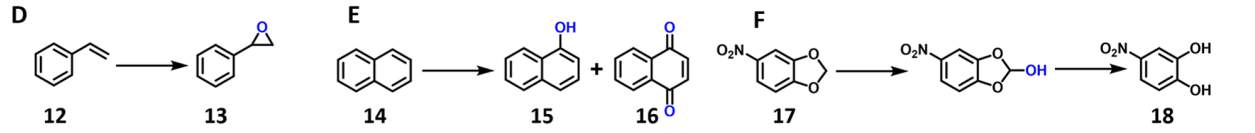
Table 2. Selection of Oxyfunctionalization Products Obtained with PROSS-Stabilized UPO Designs[a]



Enzyme	TON	
	2	3
*Cci*UPO_d7	1793±353	136±14
*Gma*UPO_d9	1452±85	223±25

Enzyme	TON	
	5	6
*Bci*UPO_d7	6±1	3±0.1
*Cci*UPO_d7	198±19	2418±180
*Cgr*UPO_d4	38±1	n.d.
*Gma*UPO_d9	334±21	2284±195

Enzyme		TON	
*Cci*UPO_d7	-OH	24±9 (**8**)	506±33 (**10**)
	=O	430±27 (**9**)	56±1 (**11**)
*Gma*UPO_d9	-OH	4±1 (**8**)	306±11 (**10**)
	=O	331±43 (**9**)	92±11 (**11**)



Enzyme	TON
*Bci*UPO_d7	63±9
*Bla*UPO_d9	46±12
*Cci*UPO_d7	3320±67
*Cgr*UPO_d4	56±8
*Gma*UPO_d9	3747±378
*Pca*UPO_d9	179±13
*Pin*UPO_d7	176±6

Enzyme	TON	
	15	16
*Bci*UPO_d7	39±9	17±3
*Bla*UPO_d9	12±2	n.d.
*Cci*UPO_d7	67±7	184±7
*Cgr*UPO_d4	57±1	n.d.
*Gma*UPO_d9	98±13	167±5
*Pca*UPO_d9	n.d.	46±2

Enzyme	TON[c]
*Bci*UPO_d7	230±12
*Cci*UPO_d7	725±16
*Cgr*UPO_d4	23±1
*Gma*UPO_d9	699±14
*Pca*UPO_d9	272± 6[b]
*Pin*UPO_d7	270±7 [b]

[a]TONs are based on GC-MS measurements if not stated otherwise and comparison with product standards. TON data are mean ± SD of measurements performed in triplicates. Reaction conditions are as follows: 4 mM substrate, 1.5 mM H_2O_2 (slow addition over the course of the reaction), 5% acetone (v/v), 250 nM enzyme, 0.1 mM potassium phosphate buffer (pH 7.0), 2 h reaction time, 30 °C. [b]Reaction time overnight. [c]TONs are based on absorption measurement. TON data are mean ± SD of measurements made in triplicates.

suitable signal peptide nor the best PROSS design can be determined a priori; yet, the effort required to screen their combinations is more limited than a typical in vitro evolution campaign.

To assess the significance of the PROSS mutations on the enzymes, the best-performing signal peptide for each enzyme was combined with the corresponding wild-type variants. Most wild-type UPOs exhibited some degree of secretion, according to the split-GFP assay. Only wild-type *Gma*UPO and wild-type *Bci*UPO showed increased secretion compared to their respective PROSS designs. All other wild-type UPOs displayed decreased secretion rates compared with the PROSS designs (Figure 3). Upscaling experiments revealed that secretion of wild-type *Bci*UPO was not possible in shake flasks, while the best PROSS design of *Bci*UPO (*Bci*UPO_d7, medium mutational load) was actively secreted under these conditions, suggesting that the production conditions are another critical determinant of functional production.[50] The only wild-type enzyme that showed activity during this work was *Cci*UPO, but its activity level was fivefold lower than the best PROSS design for this enzyme (Figure 3). From these experiments, we conclude that although some of the selected UPOs are producible in their wild-type form, the PROSS designs showed higher production levels, and the design process was essential for obtaining functional UPOs.

Diverse Enzyme Characteristics and Substrate Scope in Designs. For each enzyme, we chose the PROSS design and signal peptide combination that showed the highest activity for DMP or ABTS (Figure 3 and Table S2) for subsequent shake flask expression and enzyme characterization. The occurrence and correct identity of all enzymes were confirmed via LC-MS (Figure S17 and Table S7). The UV absorption spectra showed the characteristic of heme−thiolate proteins with a Soret band maximum around 420 nm for *Bci*UPO_d7 (424 nm), *Cci*UPO_d7 (418 nm), *Cgr*UPO_d4 (420 nm), *Chi*UPO_d7 (410 nm), *Fgr*UPO_d5 (415 nm), *Gma*UPO_d9 (418 nm), and *Pca*UPO_d9 (417 nm) (Figure S13). For *Bla*UPO_d9 and *Pin*UPO_d7, this maximum could not be detected due to challenges during purification and potential problems in heme incorporation (Figure S13).

We analyzed the substrate scope for each enzyme using concentrated supernatant, and TONs were determined with GC-MS. We chose test substrates that demand two-electron peroxygenase activity rather than one-electron peroxidase activity as these are the more challenging reactions (Table 2). Among the substrates were two with activated C−H bonds (NBD [5-nitro-1,3-benzodioxole] and ethylbenzene), styrene as an epoxidation substrate, the aromatic substrate naphthalene, and the two aliphatic substrates, cyclohexane and octane with nonactivated sp^3-carbons. Very low enzyme concentrations (production titer <0.5 mg/L) for *Chi*UPO_d7 and *Fgr*UPO_d5 after shake flask expression impeded determining their substrate scope.

The overall highest activities were demonstrated by the two long-type UPOs, *Cci*UPO_d7 and *Gma*UPO_d9. Styrene was

https://doi.org/10.1021/acscatal.4c00883
ACS Catal. 2024, 14, 4738−4748

Table 3. Comparison of the Transition Temperature (T_m) and Onset of Aggregation (T_{agg}) for Selected PROSS-Optimized UPOs Expressed in Both *P. pastoris* and *S. cerevisiae*

	expression host P. pastoris C-terminal-tag TwinStrep (two additional tryptophan)			expression host S. cerevisiae C-terminal-tag His2 (no additional tryptophan)			
enzyme	fluorescence shift	T_m	T_{agg}[a]	fluorescence shift	T_m	T_{agg}[a]	$\Delta T_m = T_m$ (*Pp*) $- T_m$ (*Sc*) (°C)
*Bci*UPO_d7	blue	44.3 ± 0.2 °C	33.9 ± 0.5 °C	red	59.4 ± 0.2 °C	n.d.	−15.1
*Cci*UPO_d7	blue	58.7 ± 0.3 °C	63.9 ± 0.9 °C	red	46.6 ± 0.3 °C	56.2 ± 0.3 °C	+12.1
*Cgr*UPO_d4	blue	42.8 ± 1.4 °C	37.7 ± 0.1 °C	red	54.3 ± 0.1 °C	n.d.	−11.5
*Chi*UPO_d7	red	62.7 ± 0.1 °C	n.d.	red	54.4 ± 0.1 °C	n.d.	+8.3
*Gma*UPO_d9	blue	58.5 ± 0.2 °C	49.9 ± 0.2 °C	red	46.7 ± 0.3 °C	46.0 ± 0.9 °C	+11.8
*Pca*UPO_d9	blue	50.0 ± 1.6 °C	41.1 ± 1.3 °C	red	52.9 ± 0.1 °C	n.d.	−2.9

[a]Onset of aggregation; n.d.—not detected, data are mean ±SD, and measurements were performed in triplicates.

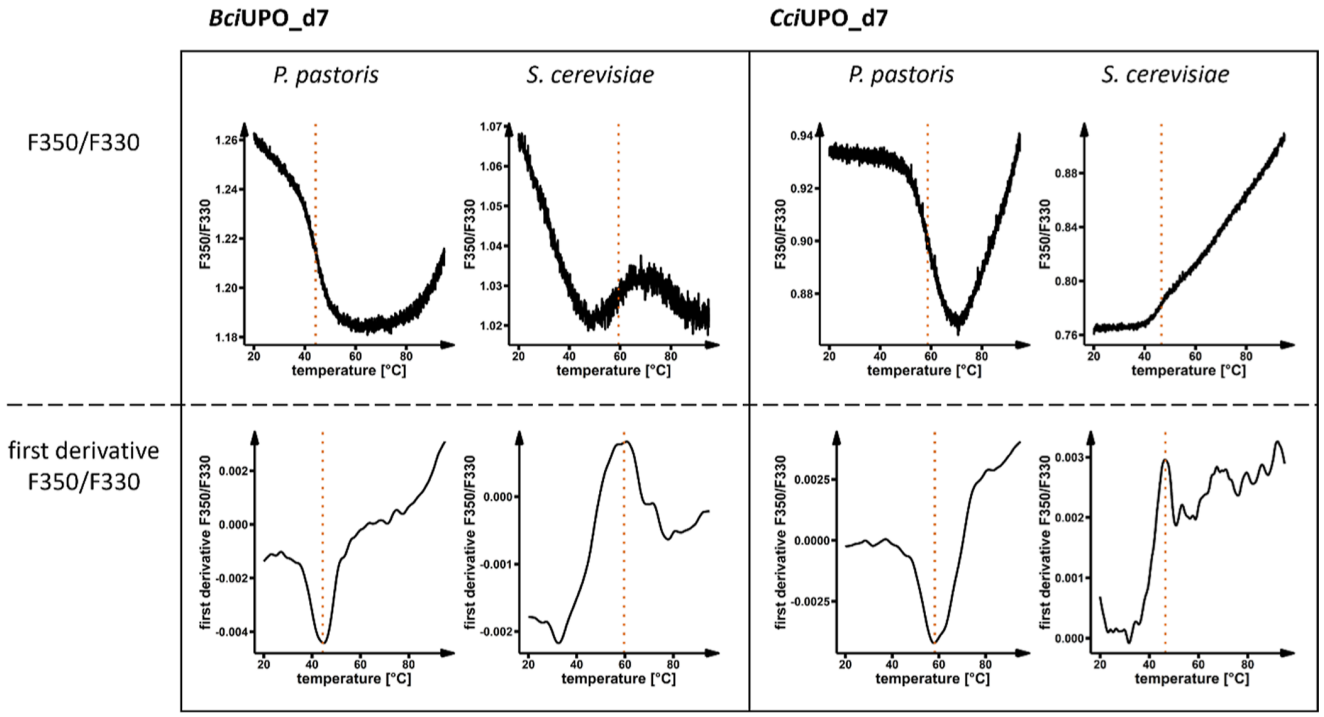


Figure 4. Comparison of fluorescence intensity change during temperature increase for *Bci*UPO_d7 and *Cci*UPO_d7. Both enzymes were expressed in *P. pastoris* (with the *C*-terminal TwinStrep tag) and *S. cerevisiae* (with the *C*-terminal His2 tag). The ratio of fluorescence intensity at 350 nm/330 nm is shown in the upper row, and the first derivative of this fluorescence ratio is shown in the lower row. The inflection point, at which the transition temperature T_m in each case was determined, is marked (orange dotted line).

converted to styrene oxide with more than 3000 TONs by both enzymes, and ethylbenzene was mainly overoxidized to acetophenone with nearly 2500 TONs (Table 2). Both activities are within the range of other UPOs like PaDa-I (4000 TONs for ethylbenzene hydroxylation and 3000 TONs for styrene epoxidation) or *Mth*UPO (500 TONs for ethylbenzene hydroxylation and 1100 TONs for styrene epoxidation).[18,51] Both long-type UPOs also exhibit activities toward nonactivated C−H bonds, shown here by the conversion of cyclohexane to cyclohexanol with more than 1500 TONs and the oxyfunctionalization of octane with total TONs of 730 (*Gma*UPO_d9) and 1010 (*Cci*UPO_d7) (Table 2), which is within the range of *Mth*UPO,[51] but significantly lower compared to 20,000 TONs for PaDa-I.[52] Activity toward styrene epoxidation was also displayed by *Bla*UPO_d9, *Bci*UPO_d7, *Cgr*UPO_d4, *Pca*UPO_d9, and *Pin*UPO_d7, the latter four also convert NBD, but all reactions lead only to low TONs between 20 and 270 (Table 2). These results further demonstrate that *Bla*UPO_d9 and *Pin*UPO_d7 are indeed functionally produced, even though they did not show activity toward ABTS or DMP under screening conditions, leading to a total of nine functionally produced UPOs. Aromatic hydroxylation was detected for *Bci*UPO_d7, *Bla*UPO_d9, *Cci*UPO_d7, *Cgr*UPO_d4, *Gma*UPO_d9, and *Pca*UPO_d9 with naphthalene as the substrate but with low TONs between 12 and 184.

The determined enantiomeric ratios (e.r.) for styrene oxide (**13**) and 1-phenylethanol (**5**) show only small differences between the investigated enzymes with a lowest e.r. of 59:41 (*Cci*UPO, styrene oxide) and a highest e.r. of 85:15 (*Gma*UPO, 1-phenylethanol). All enzymes showed a preference for the *R* enantiomer in the case of both products (Table S8).

In addition to the substrate scope, we also determined the apparent transition temperature (T_m) for all best PROSS designs produced in both *P. pastoris* and *S. cerevisiae*. The

https://doi.org/10.1021/acscatal.4c00883
ACS Catal. 2024, 14, 4738−4748

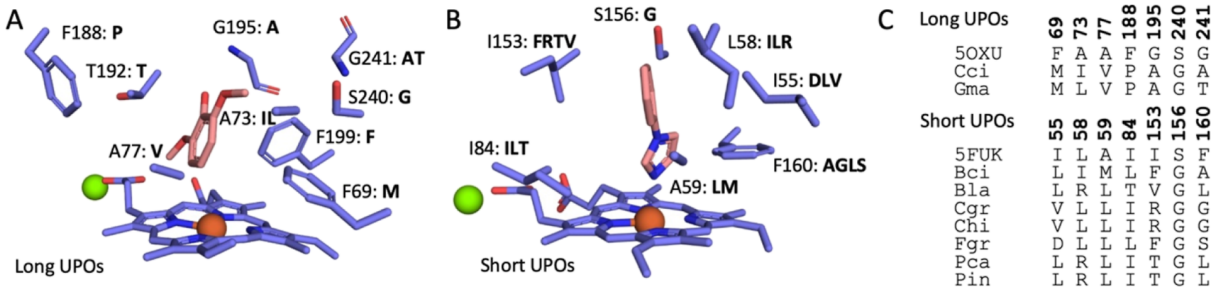


Figure 5. Active-site sequence diversity among successfully designed UPOs. (A) Long UPO active site. PaDa-I (pdb:5OXU) heme and active-site residues are presented in purple sticks (unless glycine, only side chain atoms are presented) and Mg^{2+} is presented as a green sphere. DMP was grafted from pdb:5OY2 and is presented in pink sticks. PaDa-I positions and identities are written next to each presented residue (left side). For each position, the identities of *Cci*UPO and *Gma*UPO in the same position are written on the right side, respectively. (B) Short UPO active site. *Mro*UPO (pdb:5FUK) heme and active-site residues are presented in purple sticks (only side chain atoms are presented) and Mg^{2+} is presented as a green sphere. 1-Phenylimidazole was grafted from pdb:7O1X and is presented in pink sticks. *Mro*UPO positions and identities are written next to each presented residue (except for position 58 that is assigned as leucine, as 5FUK possesses T58L mutation; left side). For each position, the identities of *Bci*UPO, *Bla*UPO, *Cgr*UPO, *Chi*UPO, *Fgr*UPO, *Pca*UPO, and *Pin*UPO in the same position are written on the right side. (C) Identities of each UPO in the presented positions (only positions that are not completely conserved are shown). Long UPOs: PaDa -I, *Cci*UPO, and *Gma*UPO are 63−72% identical to one another. Short UPOs: *Cgr*UPO and *Chi*UPO are 80% identical, *Pca*UPO and *Pin*UPO are 87% identical, *Bla*UPO is 66% identical to both *Pca*UPO and *Pin*UPO, and all other UPO pairs are 25−50% identical to one another.

enzymes expressed in *P. pastoris* contained a *C*-terminal TwinStrep tag for purification while the counterparts from *S. cerevisiae* carried a *C*-terminal His2 tag. Both sets of UPOs were purified via respective affinity chromatography. To investigate the thermal stability, we applied nano differential scanning fluorimetry (nanoDSF). Parallel to the fluorescence measurements, we detected light scattering to determine the potential onset of aggregation. Interestingly, the resulting temperature values for both sets were different depending on the production host and *C*-terminal tag. For *P. pastoris* as the production host, *Chi*UPO_d7 showed the highest apparent transition temperature with 62.7 °C. Both long-type UPOs *Cci*UPO_d7 and *Gma*UPO_d9 showed similar transition temperatures of about 58.6 °C. The lowest transition temperature was measured for *Cgr*UPO_d7 with 42.8 °C (Table 3, Figures 4 and S15A). Comparing the transition temperatures of the recombinant enzymes secreted from *S. cerevisiae*, *Bci*UPO_d7 had the highest T_m with 59.4 °C, whereas the lowest value was found for *Cci*UPO_d4 with 46.6 °C (Table 3, Figures 4 and S15B). The transition and onset temperatures of *Bla*UPO, *Fgr*UPO, and *Pin*UPO could not be determined with satisfying reliability due to very low enzyme concentrations after purification. The data we obtained revealed strong differences in T_m for the individual PROSS designs, e.g., ΔT_m of 15.1 °C for *Bci*UPO_d7 with a higher thermal stability when expressed in *S. cerevisiae* and ΔT_m of 12.2 °C for *Cci*UPO_d7 with a higher thermal stability when expressed in *P. pastoris*. We further noticed a reverse trajectory of the fluorescence intensity change (ratio at 350 nm/330 nm) for *Bci*UPO_d7, *Cci*UPO_d7, *Cgr*UPO_d4, and *Gma*UPO_d9. With increasing temperature, an increase in fluorescence is observed for enzymes expressed in *S. cerevisiae* with the His2-GFP11 tag, whereas a decrease is mainly observed for variants expressed in *P. pastoris* with the TwinStrep-GFP11 tag (Table 3, Figures 4 and S15). Considering our approach, we introduced two additional surface-exposed Trp residues which are present in the TwinStrep-GFP11 tag[53] in contrast to the His2-GFP11 tag. If the majority of Trp residues beforehand are buried in the enzyme, as is the case, e.g., for *Bci*UPO_d7, adding two surface-exposed Trp residues can lead to a different fluorescence behavior as observed in our case. Regarding the positioning of Trp in *Cci*UPO_d7, which has rather surface-exposed Trp residues (Figure S14), the blue shift trajectory, as detected for the construct from *P. pastoris* as the host, can be expected. Hence, the red-shifted progress of the enzyme from *S. cerevisiae* is rather surprising. We assume that the glycosylation pattern (both *O*- and *N*-glycosylation) might play an important role in this context as well as in the differences in protein stability. It is well known that glycosylation can strongly affect protein folding and stability[54,55] as well as intrinsic protein dynamics[56] and therefore activity.[57,58] Previous research also revealed the influence of glycans on the fluorescence behavior of intrinsic Trp fluorescence itself.[59,60] Considering the differences published for glycosylation patterns from *P. pastoris* and *S. cerevisiae*,[61] we therefore hypothesize that these differences affect both enzyme stability and the fluorescence behavior of our PROSS designs. Glycosylation has also been shown to have an influence on the aggregation tendency of proteins.[62−64] The light scattering measurement to detect the temperature for the onset of aggregation supports this finding also for the enzymes investigated in the present study (Table 3 and Figure S16). While only *Cci*UPO_d7 and *Gma*UP_d9 expressed in *S. cerevisiae* show typical aggregation behavior, the aggregation tendency for the UPOs expressed in *P. pastoris* is more pronounced. These data could be due to hyper-glycosylation in *S. cerevisiae*.[65] Overall, it is crucial to acknowledge that both the host organism, along with the resulting glycosylation pattern, and the chosen tag can influence enzyme stability and folding behavior.

Structural Underpinnings of Functional Diversity. Finally, we analyzed the AF2 structure models of the nine UPOs that were functional following the design (Figure 5). The models exhibit only minor backbone differences from the reference crystallographic structure (Figure S18). The enzyme active sites were not allowed to be designed, and any differences in amino acids in them are strictly due to the natural diversity of UPOs. Whereas the long-type UPOs exhibit few amino acid changes in the active site (Figure 5A,C), the short-chain ones exhibit high diversity (Figure 5B,C). For instance, positions 55, 58, and 153 (pdb 5FUK

https://doi.org/10.1021/acscatal.4c00883
ACS Catal. 2024, 14, 4738−4748

numbering) exhibit hydrophobic identities in most of the short UPOs and charged identities in some of the variants. Even among the hydrophobic-to-hydrophobic exchanges, some are predicted to change the sterics of the active site dramatically, as in Val/Phe at position 153 and Gly/Leu at position 160. These large changes in electrostatics and sterics are likely to change the positioning of the substrate relative to that of the heme, leading to the observed functional differences.

CONCLUSIONS

Due to their challenging heterologous production, only a few UPOs are available for engineering campaigns, and their respective crystal structures are rarely determined. The improvement of ab initio structure prediction tools like AF2 provides access to numerous—if not all—UPO structures and thus opens the way to structure-based stability design. We combined AF2 for structure prediction, PROSS for introducing stabilizing mutations, and signal peptide shuffling to increase enzyme secretion. These three methods had not been previously used in conjunction and enabled the production of a highly challenging class of UPOs. Our success rate in functional expression reached 90%, as only *Fox*UPO did not show any activity or secretion. This rate is significantly higher compared to our previous workflow where we combined only the legacy trRosetta and PROSS without a signal peptide shuffling, which resulted in the successful expression of only three of 11 target VPs.[31] PROSS proved to be very beneficial for gaining functionally produced enzymes as (i) secretion levels are increased substantially in most cases for the PROSS design compared to the wild-type enzyme and (ii) eight out of nine wild-type enzymes did not show any activity for the tested screening substrates (ABTS and DMP), while seven PROSS UPOs displayed activity on at least one of the screening substrates (ABTS or DMP). All nine produced PROSS UPOs were active on at least one tested substrate, if not a screening substrate, then during substrate scope analysis with six additional substrates. This high success rate demonstrates a clear path to protein engineering of even challenging enzymes using modern modeling and design software. In fact, our success rate surpasses that of other PROSS applications[31,50] suggesting that either the combination with signal peptide shuffling has a beneficial effect or the AF2-predicted structures offer an advantage over crystal structures. For the latter possibility, we speculate that the high success rate may be due to restricting designs in positions that exhibit low modeling confidence (AF2 plDDT scores <90%) and their vicinity (this functionality is automatically enabled in the PROSS web server when AF2 models are used; https://PROSS.weizmann.ac.il). Crystallographic structure analysis does not provide a comparable way to assign confidence to structures.

Nevertheless, PROSS on its own is insufficient to gain functionally active, secreted UPOs. In our work, we combined 30 PROSS designs of 10 UPOs with 17 different signal peptides and screened them with three different assays to determine the secretion level and activity for all combinations. Our work shows again, in accordance with the previous work,[18,19] that no single most suitable signal peptide was found for all UPOs, and secretion rates differed greatly between different samples during the signal-peptide-shuffling screening.

Successful secretion of UPOs in a fast-growing microbial host is the first essential step on the path to the facile access of customized enzymes for industrial application. Seven of the newly characterized enzymes derive from phytopathogenic organisms. It may be interesting to study their effect on plant material to gain further insights into the natural functions of UPOs and their possible interaction with the plant defense system. Three of the UPOs are derived from oomycetes, extending the range of available UPOs, for the first time, to nonfungal organisms.

It is important to note that design calculations were not applied to the active site. The sequence variations within the active site (Figure 5) are therefore entirely derived from natural UPO diversity. In most cases, the variants introduce multiple simultaneous changes relative to one another. Combinations of mutations within the active site are likely to be epistatic with one another or with mutations and backbone differences outside of the active site.[66−68] Such epistatic relationships are known to slow functional innovation in natural and laboratory evolution.[69] Thus, our structural analysis highlights the major strength of the design process: instead of painstakingly re-engineering the active site and the enzyme backbone, it exploits the natural structural and sequence diversity to expose new substrate specificities and reactivities. In the case of UPOs, thousands of natural sequences are known, only a few dozen of which have been successfully characterized until now. We envision that the computational design will dramatically accelerate the discovery of oxyfunctionalization reactions in this family.

Finally, the workflow we employed can be readily adopted by other laboratories, as AF2 and PROSS are available by web servers, and all required signal peptides and vectors for yeast production and secretion have been previously deposited at Addgene (Yeast Secrete and Detect Kit #1000000166).[18,19] In this study, we could readily identify the heme-binding site by a comparison to two of the known experimental structures of UPOs. If a homologous experimental structure is not known, however, the active site may be identified by using other means. Conservation analysis or prior experimental studies may provide guidance in such cases. Additionally, recent advances enable ab initio structure prediction of the protein with known ligands.[70] These new methods may extend the pipeline that we have demonstrated here even to cases in which no homologous structure has been experimentally determined. Thus, the combination of AI-based structure prediction and atomistic design can provide enzymologists and protein engineers with access to a vast array of functions from previously untappable natural enzymes.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acscatal.4c00883.

Material and methods; gene and protein sequence of all enzymes investigated in this study; initial screening data; GC-MS temperature programs; utilized linker and tags; initial screening data and overview over best PROSS design and signal peptide combination; relative abundance of signal peptides during the initial screening; transition temperature profiles of all purified enzymes; calibration curve for GC-MS measurements; and absorption spectra (PDF)

https://doi.org/10.1021/acscatal.4c00883
ACS Catal. 2024, 14, 4738−4748

AUTHOR INFORMATION

Corresponding Authors

Sarel J. Fleishman − *Department of Biomolecular Sciences, Weizmann Institute of Science, Rehovot 7600001, Israel;* orcid.org/0000-0003-3177-7560; Email: sarel@weizmann.ac.il

Martin J. Weissenborn − *Institute of Chemistry, Martin Luther-University Halle-Wittenberg, Halle (Saale) 06120, Germany; Leibniz Institute of Plant Biochemistry, Halle (Saale) 06120, Germany;* orcid.org/0000-0002-1200-4485; Email: martin.weissenborn@chemie.uni-halle.de

Authors

Judith Münch − *Institute of Chemistry, Martin Luther-University Halle-Wittenberg, Halle (Saale) 06120, Germany; Leibniz Institute of Plant Biochemistry, Halle (Saale) 06120, Germany*

Niklas Dietz − *Institute of Chemistry, Martin Luther-University Halle-Wittenberg, Halle (Saale) 06120, Germany; Leibniz Institute of Plant Biochemistry, Halle (Saale) 06120, Germany*

Shiran Barber-Zucker − *Department of Biomolecular Sciences, Weizmann Institute of Science, Rehovot 7600001, Israel;* Present Address: Scala Biodesign LTD, 50 Dizengoff St., Tel Aviv 6433222, Israel

Franziska Seifert − *Department of Pharmaceutical Technology and Biopharmacy, Institute of Pharmacy, Martin-Luther-University Halle-Wittenberg, Halle (Saale) 06120, Germany*

Susanne Matschi − *Leibniz Institute of Plant Biochemistry, Halle (Saale) 06120, Germany*

Pascal Püllmann − *Leibniz Institute of Plant Biochemistry, Halle (Saale) 06120, Germany;* Present Address: Molecular Design and Engineering, Bayer AG, Aprather Weg 18A, 42113 Wuppertal, Germany

Complete contact information is available at: https://pubs.acs.org/10.1021/acscatal.4c00883

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. Both J.M. and N.D. contributed equally and have the right to list their name first in their CV.

Funding

Friedrich-Naumann-Stiftung für die Freiheit Bundesministerium für Bildung and Forschung (Maßgeschneiderte Inhaltsstoffe 2, 031B0834A), German Research Foundation (DFG, project ID 436494874, TP A05, RTG 2670), German Research Foundation (DFG, project ID 505185500), Israel Science Foundation (ISF, #1844).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

J.M. thanks the Friedrich-Naumann-Stiftung für die Freiheit for a PhD scholarship. M.J.W. and J.M. thank the Bundesministerium für Bildung und Forschung (Maßgeschneiderte Inhaltsstoffe 2, 031B0834A) and German Research Foundation (DFG, project ID 436494874, TP A05, RTG 2670) for their generous funding. N.D. thanks the German Research Foundation (DFG, project ID 505185500) for their support. We would like to thank Domenika Thieme (Leibniz-Institute for Plant Biochemistry) for performing protein identification measurements at the LC-MS after trypsin digestion.

ABBREVIATIONS

AF2, AlphaFold2; ABTS, 2,2′-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid); DMP, 2,6-dimethoxyphenol; GC-MS, gas chromatography-mass spectrometry; NBD, 5-nitro-1,3-benzodioxole; *P. pastoris*, *Pichia pastoris* (*Komagataella phaffii*); PROSS, Protein Repair One-Stop Shop; T_{agg}, onset of aggregation; T_m, transition temperature; TON, turnover number; UPO, unspecific peroxygenase

REFERENCES

(1) Beltrán-Nogal, A.; Sánchez-Moreno, I.; Méndez-Sánchez, D.; Gómez de Santos, P.; Hollmann, F.; Alcalde, M. Surfing the wave of oxyfunctionalization chemistry by engineering fungal unspecific peroxygenases. *Curr. Opin. Struct. Biol.* **2022**, *73*, 102342.

(2) Monterrey, D. T.; Menés-Rubio, A.; Keser, M.; Gonzalez-Perez, D.; Alcalde, M. Unspecific peroxygenases: The pot of gold at the end of the oxyfunctionalization rainbow? *Curr. Opin. Green Sustain. Chem.* **2023**, *41*, 100786.

(3) Kinner, A.; Rosenthal, K.; Lütz, S. Identification and expression of new unspecific peroxygenases-Recent advances, challenges and opportunities. *Front. Bioeng. Biotechnol.* **2021**, *9*, 705630.

(4) Hobisch, M.; Holtmann, D.; Gomez de Santos, P.; Alcalde, M.; Hollmann, F.; Kara, S. Recent developments in the use of peroxygenases-Exploring their high potential in selective oxyfunctionalisations. *Biotechnol. Adv.* **2021**, *51*, 107615.

(5) Li, Z.; Jiang, Y.; Guengerich, F. P.; Ma, L.; Li, S.; Zhang, W. Engineering cytochrome P450 enzyme systems for biomedical and biotechnological applications. *J. Biol. Chem.* **2020**, *295* (3), 833−849.

(6) Guengerich, F. P.; Waterman, M. R.; Egli, M. Recent structural insights into cytochrome P450 function. *TIPS* **2016**, *37* (8), 625−640.

(7) Manikandan, P.; Nagini, S. Cytochrome P450 structure, function and clinical significance: a review. *Curr. Drug Targets* **2018**, *19* (1), 38−54.

(8) Münch, J.; Püllmann, P.; Zhang, W.; Weissenborn, M. J. Enzymatic hydroxylations of sp3-carbons. *ACS Catal.* **2021**, *11* (15), 9168−9203.

(9) Hobisch, M.; De Santis, P.; Serban, S.; Basso, A.; Byström, E.; Kara, S. Peroxygenase-Driven Ethylbenzene Hydroxylation in a Rotating Bed Reactor. *Org. Process Res. Dev.* **2022**, *26* (9), 2761−2765.

(10) Aranda, C.; Municoy, M.; Guallar, V.; Kiebist, J.; Scheibner, K.; Ullrich, R.; del Río, J. C.; Hofrichter, M.; Martínez, A. T.; Gutiérrez, A. Selective synthesis of 4-hydroxyisophorone and 4-ketoisophorone by fungal peroxygenases. *Catal. Sci. Technol.* **2019**, *9* (6), 1398−1405.

(11) Babot, E. D.; del Río, J. C.; Kalum, L.; Martínez, A. T.; Gutiérrez, A. Oxyfunctionalization of aliphatic compounds by a recombinant peroxygenase from Coprinopsis cinerea. *Biotechnol. Bioeng.* **2013**, *110* (9), 2323−2332.

(12) Bormann, S.; Kellner, H.; Hermes, J.; Herzog, R.; Ullrich, R.; Liers, C.; Ulber, R.; Hofrichter, M.; Holtmann, D. Broadening the biocatalytic toolbox—Screening and expression of new unspecific peroxygenases. *Antioxidants* **2022**, *11* (2), 223.

(13) Carro, J.; González-Benjumea, A.; Fernández-Fueyo, E.; Aranda, C.; Guallar, V.; Gutiérrez, A.; Martínez, A. T. Modulating fatty acid epoxidation vs hydroxylation in a fungal peroxygenase. *ACS Catal.* **2019**, *9* (7), 6234−6242.

(14) Ma, Y.; Liang, H.; Zhao, Z.; Wu, B.; Lan, D.; Hollmann, F.; Wang, Y. A Novel Unspecific Peroxygenase from Galatian Marginata for Biocatalytic Oxyfunctionalization Reactions. *J. Mol. Catal.* **2022**, *531*, 112707.

(15) Molina-Espeja, P.; Garcia-Ruiz, E.; Gonzalez-Perez, D.; Ullrich, R.; Hofrichter, M.; Alcalde, M. Directed evolution of unspecific

peroxygenase from Agrocybe aegerita. *Appl. Environ. Microbiol.* **2014**, *80* (11), 3496−3507.
(16) Pecyna, M. J.; Schnorr, K. M.; Ullrich, R.; Scheibner, K.; Kluge, M. G.; Hofrichter, M. Fungal peroxygenases and methods of application. WO 2008/119780 A2, 2008.
(17) Rotilio, L.; Swoboda, A.; Ebner, K.; Rinnofner, C.; Glieder, A.; Kroutil, W.; Mattevi, A. Structural and biochemical studies enlighten the unspecific peroxygenase from Hypoxylon sp. EC38 as an efficient oxidative biocatalyst. *ACS Catal.* **2021**, *11* (18), 11511−11525.
(18) Püllmann, P.; Knorrscheidt, A.; Münch, J.; Palme, P. R.; Hoehenwarter, W.; Marillonnet, S.; Alcalde, M.; Westermann, B.; Weissenborn, M. J. A modular two yeast species secretion system for the production and preparative application of unspecific peroxygenases. *Commun. Biol.* **2021**, *4* (1), 562.
(19) Püllmann, P.; Weissenborn, M. J. Improving the heterologous production of fungal peroxygenases through an episomal Pichia pastoris promoter and signal peptide shuffling system. *ACS Synth. Biol.* **2021**, *10* (6), 1360−1372.
(20) Ebner, K.; Pfeifenberger, L. J.; Rinnofner, C.; Schusterbauer, V.; Glieder, A.; Winkler, M. Discovery and Heterologous Expression of Unspecific Peroxygenases. *Catalysts* **2023**, *13* (1), 206.
(21) König, R.; Kiebist, J.; Kalmbach, J.; Herzog, R.; Schmidtke, K.-U.; Kellner, H.; Ullrich, R.; Jehmlich, N.; Hofrichter, M.; Scheibner, K. Novel unspecific peroxygenase from Truncatella angustata catalyzes the synthesis of bioactive lipid mediators. *Microorganisms* **2022**, *10* (7), 1267.
(22) Magliery, T. J. Protein stability: computation, sequence statistics, and new experimental methods. *Curr. Opin. Struct. Biol.* **2015**, *33*, 161−168.
(23) Goldenzweig, A.; Goldsmith, M.; Hill, S. E.; Gertman, O.; Laurino, P.; Ashani, Y.; Dym, O.; Unger, T.; Albeck, S.; Prilusky, J.; et al. Automated structure-and sequence-based design of proteins for high bacterial expression and stability. *Mol. Cell* **2016**, *63* (2), 337−346.
(24) Akbulut, N.; Tuzlakoğlu Öztürk, M.; Pijning, T.; İşsever Öztürk, S.; Gümüşel, F. Improved activity and thermostability of Bacillus pumilus lipase by directed evolution. *J. Biotechnol.* **2013**, *164* (1), 123−129.
(25) Socha, R. D.; Tokuriki, N. Modulating protein stability-directed evolution strategies for improved protein function. *FEBS J.* **2013**, *280* (22), 5582−5595.
(26) Modarres, H. P.; Mofrad, M.; Sanati-Nezhad, A. Protein thermostability engineering. *RCS Adv.* **2016**, *6* (116), 115252−115270.
(27) Kulshreshtha, S.; Chaudhary, V.; Goswami, G. K.; Mathur, N. Computational approaches for predicting mutant protein stability. *J. Comput. Aided Mol. Des.* **2016**, *30*, 401−412.
(28) Goldenzweig, A.; Fleishman, S. J. Principles of protein stability and their application in computational design. *Annu. Rev. Biochem.* **2018**, *87*, 105−129.
(29) Wijma, H. J.; Fürst, M. J. L. J.; Janssen, D. B. A computational library design protocol for rapid improvement of protein stability: FRESCO. *Protein Eng.* **2018**, *1685*, 69−85.
(30) Weinstein, J. J.; Goldenzweig, A.; Hoch, S.; Fleishman, S. J. PROSS 2: a new server for the design of stable and highly expressed protein variants. *Bioinformatics* **2021**, *37* (1), 123−125.
(31) Barber-Zucker, S.; Mindel, V.; Garcia-Ruiz, E.; Weinstein, J. J.; Alcalde, M.; Fleishman, S. J. Stable and functionally diverse versatile peroxidases designed directly from sequences. *J. Am. Chem. Soc.* **2022**, *144* (8), 3564−3571.
(32) Barber-Zucker, S.; Mateljak, I.; Goldsmith, M.; Kupervaser, M.; Alcalde, M.; Fleishman, S. J. Designed high-redox potential laccases exhibit high functional diversity. *ACS Catal.* **2022**, *12* (21), 13164−13173.
(33) Ramirez-Escudero, M.; Molina-Espeja, P.; Gomez de Santos, P.; Hofrichter, M.; Sanz-Aparicio, J.; Alcalde, M. Structural insights into the substrate promiscuity of a laboratory-evolved peroxygenase. *ACS Chem. Biol.* **2018**, *13* (12), 3259−3268.
(34) Piontek, K.; Strittmatter, E.; Ullrich, R.; Gröbe, G.; Pecyna, M. J.; Kluge, M.; Scheibner, K.; Hofrichter, M.; Plattner, D. A. Structural basis of substrate conversion in a new aromatic peroxygenase: cytochrome P450 functionality with benefits. *J. Biol. Chem.* **2013**, *288* (48), 34767−34776.
(35) Linde, D.; Santillana, E.; Fernández-Fueyo, E.; González-Benjumea, A.; Carro, J.; Gutiérrez, A.; Martínez, A. T.; Romero, A. Structural characterization of two short unspecific peroxygenases: two different dimeric arrangements. *Antioxidants* **2022**, *11* (5), 891.
(36) Kryshtafovych, A.; Schwede, T.; Topf, M.; Fidelis, K.; Moult, J. Critical assessment of methods of protein structure prediction (CASP)—Round XIV. *Proteins.* **2021**, *89* (12), 1607−1617.
(37) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583−589.
(38) Karich, A.; Ullrich, R.; Scheibner, K.; Hofrichter, M. Fungal unspecific peroxygenases oxidize the majority of organic EPA priority pollutants. *Front. Microbiol.* **2017**, *8*, 1463.
(39) Westrick, N. M.; Smith, D. L.; Kabbage, M. Disarming the host: detoxification of plant defense compounds during fungal necrotrophy. *Front. Plant Sci.* **2021**, *12*, 651716.
(40) Beakes, G. W.; Glockling, S. L.; Sekimoto, S. The evolutionary phylogeny of the oomycete "fungi. *Protoplasma* **2012**, *249*, 3−19.
(41) Cavalier-Smith, T.; Chao, E. E. Phylogeny and megasystematics of phagotrophic heterokonts (kingdom Chromista). *J. Mol. Evol.* **2006**, *62*, 388−420.
(42) Aranda, C.; Ullrich, R.; Kiebist, J.; Scheibner, K.; del Río, J. C.; Hofrichter, M.; Martínez, A. T.; Gutiérrez, A. Selective synthesis of the resveratrol analogue 4, 4′-dihydroxy-trans-stilbene and stilbenoids modification by fungal peroxygenases. *Catal. Sci. Technol.* **2018**, *8* (9), 2394−2401.
(43) Karich, A.; Scheibner, K.; Ullrich, R.; Hofrichter, M. Exploring the catalase activity of unspecific peroxygenases and the mechanism of peroxide-dependent heme destruction. *J. Mol. Catal. Chem.* **2016**, *134*, 238−246.
(44) Aranda, C.; Olmedo, A.; Kiebist, J.; Scheibner, K.; del Río, J. C.; Martínez, A. T.; Gutiérrez, A. Selective epoxidation of fatty acids and fatty acid methyl esters by fungal peroxygenases. *ChemCatChem* **2018**, *10* (18), 3964−3968.
(45) Knorrscheidt, A.; Püllmann, P.; Schell, E.; Homann, D.; Freier, E.; Weissenborn, M. J. Identification of novel unspecific peroxygenase chimeras and unusual YfeX axial heme ligand by a versatile high-throughput GC-MS approach. *ChemCatChem* **2020**, *12* (19), 4788−4795.
(46) Olmedo, A.; Ullrich, R.; Hofrichter, M.; del Río, J. C.; Martínez, Á. T.; Gutiérrez, A. Novel Fatty Acid Chain-Shortening by Fungal Peroxygenases Yielding 2C-Shorter Dicarboxylic Acids. *Antioxidants* **2022**, *11* (4), 744.
(47) Babot, E. D.; Aranda, C.; Kiebist, J.; Scheibner, K.; Ullrich, R.; Hofrichter, M.; Martínez, A. T.; Gutiérrez, A. Enzymatic epoxidation of long-chain terminal alkenes by fungal peroxygenases. *Antioxidants* **2022**, *11* (3), 522.
(48) Ullrich, R.; Poraj-Kobielska, M.; Scholze, S.; Halbout, C.; Sandvoss, M.; Pecyna, M. J.; Scheibner, K.; Hofrichter, M. Side chain removal from corticosteroids by unspecific peroxygenase. *J. Inorg. Biochem.* **2018**, *183*, 84−93.
(49) Santos-Aberturas, J.; Dörr, M.; Waldo, G. S.; Bornscheuer, U. T. In-depth high-throughput screening of protein engineering libraries by split-GFP direct crude cell extract data normalization. *Chem. Biol.* **2015**, *22* (10), 1406−1414.
(50) Peleg, Y.; Vincentelli, R.; Collins, B. M.; Chen, K.-E.; Livingstone, E. K.; Weeratunga, S.; Leneva, N.; Guo, Q.; Remans, K.; Perez, K.; et al. Community-wide experimental evaluation of the PROSS stability-design method. *J. Mol. Biol.* **2021**, *433* (13), 166964.
(51) Knorrscheidt, A.; Soler, J.; Hünecke, N.; Püllmann, P.; Garcia-Borràs, M.; Weissenborn, M. J. Simultaneous screening of multiple substrates with an unspecific peroxygenase enabled modified alkane

https://doi.org/10.1021/acscatal.4c00883
ACS Catal. 2024, 14, 4738−4748

and alkene oxyfunctionalisations. *Catal. Sci. Technol.* **2021**, *11* (18), 6058−6064.
(52) Hilberath, T.; van Oosten, R.; Victoria, J.; Brasselet, H.; Alcalde, M.; Woodley, J. M.; Hollmann, F. Toward Kilogram-Scale Peroxygenase-Catalyzed Oxyfunctionalization of Cyclohexane. *Org. Process Res. Dev.* **2023**, *27*, 1384−1389.
(53) Schmidt, T. G.; Batz, L.; Bonet, L.; Carl, U.; Holzapfel, G.; Kiem, K.; Matulewicz, K.; Niermeier, D.; Schuchardt, I.; Stanar, K. Development of the Twin-Strep-tag® and its application for purification of recombinant proteins from cell culture supernatants. *Protein Expr. Purif.* **2013**, *92* (1), 54−61.
(54) Shental-Bechor, D.; Levy, Y. Effect of glycosylation on protein folding: a close look at thermodynamic stabilization. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105* (24), 8256−8261.
(55) Powell, L. M.; Pain, R. H. Effects of glycosylation on the folding and stability of human, recombinant and cleaved α1-antitrypsin. *J. Mol. Biol.* **1992**, *224* (1), 241−252.
(56) Lee, H. S.; Qi, Y.; Im, W. Effects of N-glycosylation on protein conformation and dynamics: Protein Data Bank analysis and molecular dynamics simulation study. *Sci. Rep.* **2015**, *5* (1), 8926.
(57) Chung, D.; Sarai, N. S.; Knott, B. C.; Hengge, N.; Russell, J. F.; Yarbrough, J. M.; Brunecky, R.; Young, J.; Supekar, N.; Vander Wall, T.; et al. Glycosylation is vital for industrial performance of hyperactive cellulases. *ACS Sustain. Chem. Eng.* **2019**, *7* (5), 4792−4800.
(58) Pol-Fachin, L.; Siebert, M.; Verli, H.; Saraiva-Pereira, M. L. Glycosylation is crucial for a proper catalytic site organization in human glucocerebrosidase. *Glycoconj. J.* **2016**, *33* (2), 237−244.
(59) Sun, F.; Zong, W.; Liu, R.; Chai, J.; Liu, Y. Micro-environmental influences on the fluorescence of tryptophan. *SAA* **2010**, *76* (2), 142−145.
(60) Zhu, B. C.; Laine, R. A.; Barkley, M. D. Intrinsic tryptophan fluorescence measurements suggest that polylactosaminyl glycosylation affects the protein conformation of the gelatin-binding domain from human placental fibronectin. *Eur. J. Biochem.* **1990**, *189* (3), 509−516.
(61) Tran, A.-M.; Nguyen, T.-T.; Nguyen, C.-T.; Huynh-Thi, X.-M.; Nguyen, C.-T.; Trinh, M.-T.; Tran, L.-T.; Cartwright, S. P.; Bill, R. M.; Tran-Van, H. Pichia pastoris versus Saccharomyces cerevisiae: a case study on the recombinant production of human granulocyte-macrophage colony-stimulating factor. *BMC Res. Notes* **2017**, *10*, 148.
(62) Kwon, K.-S.; Yu, M.-H. Effect of glycosylation on the stability of α1-antitrypsin toward urea denaturation and thermal deactivation. *Biochim. Biophys. Acta* **1997**, *1335* (3), 265−272.
(63) Høiberg-Nielsen, R.; Westh, P.; Skov, L.; Arleth, L. Interrelationship of steric stabilization and self-crowding of a glycosylated protein. *Biophys. J.* **2009**, *97* (5), 1445−1453.
(64) Kayser, V.; Chennamsetty, N.; Voynov, V.; Forrer, K.; Helk, B.; Trout, B. L. Glycosylation influences on the aggregation propensity of therapeutic monoclonal antibodies. *Biotechnol. J.* **2011**, *6* (1), 38−44.
(65) Vieira Gomes, A. M.; Souza Carmo, T.; Silva Carvalho, L.; Mendonça Bahia, F.; Parachin, N. S. Comparison of yeasts as hosts for recombinant protein production. *Microorganisms* **2018**, *6* (2), 38.
(66) Miton, C. M.; Tokuriki, N. How mutational epistasis impairs predictability in protein evolution and design. *Protein Sci.* **2016**, *25* (7), 1260−1272.
(67) Lipsh-Sokolik, R.; Khersonsky, O.; Schröder, S. P.; de Boer, C.; Hoch, S.-Y.; Davies, G. J.; Overkleeft, H. S.; Fleishman, S. J. Combinatorial assembly and design of enzymes. *Science* **2023**, *379* (6628), 195−201.
(68) Weinstein, J. Y.; Martí-Gómez, C.; Lipsh-Sokolik, R.; Hoch, S. Y.; Liebermann, D.; Nevo, R.; Weissman, H.; Petrovich-Kopitman, E.; Margulies, D.; Ivankov, D.; et al. Designed active-site library reveals thousands of functional GFP variants. *Nat. Commun.* **2023**, *14* (1), 2890.
(69) Breen, M. S.; Kemena, C.; Vlasov, P. K.; Notredame, C.; Kondrashov, F. A. Epistasis as the primary factor in molecular evolution. *Nature* **2012**, *490* (7421), 535−538.
(70) Krishna, R.; Wang, J.; Ahern, W.; Sturmfels, P.; Venkatesh, P.; Kalvet, I.; Lee, G. R.; Morey-Burrows, F. S.; Anishchenko, I.; Humphreys, I. R.; et al. Generalized biomolecular modeling and design with RoseTTAFold All-Atom. *bioRxiv* **2023**, 2023.10.09.561603.

https://doi.org/10.1021/acscatal.4c00883
ACS Catal. 2024, 14, 4738−4748

6 | Chapter IV

This chapter has been published as:

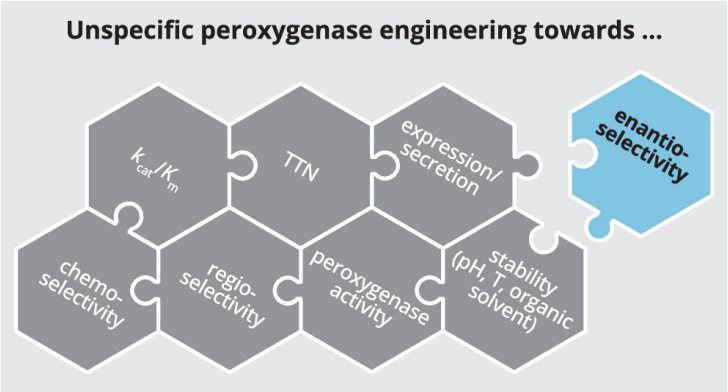
Computational-Aided Engineering of a Selective Unspecific Peroxygenase toward Enantiodivergent β-Ionone Hydroxylation

by: **Judith Münch***, Jordi Soler*, Nicole Hünecke, Dominik Homann, Marc Garcia-Borràs, and Martin J. Weissenborn

*shared first authorship

in: *ACS Catalysis* **2023**, *13* (13), 8963-8972; doi: https://doi.org/10.1021/acscatal.3c00702

Supplementary information can be assessed free of charge at the article's landing page.



This chapter describes an engineering campaign with the goal to evolve a native unspecific peroxygenases from poor toward high enantioselectivity, targeted for a specific substrate. The selected enzyme *Mth*UPO demonstrates good starting-activity and regioselectivity for the model substrate β-ionone, which is converted to 4-hydroxy-β-ionone. Computational-modelling guided the construction of a small smart library, designed to alter the substrate positioning within the active site to favor near-attack conformations for either pro-(*R*) or pro-(*S*) C4-H orientations. The product screening was performed using the MISER GC-MS method, which only screens for activity and not enantioselectivity. The most active enzyme variants were subsequently screened for improved enantioselectivity. Within two rounds of directed evolution, two enzyme variants were found exhibiting up to 17-fold activity increase, 99.6 % regioselectivity and enantiomeric ratios of 96.6:3.4 (*R*) and 0.3:99.7 (*S*).

ACS Catalysis

pubs.acs.org/acscatalysis

Research Article

Computational-Aided Engineering of a Selective Unspecific Peroxygenase toward Enantiodivergent β-Ionone Hydroxylation

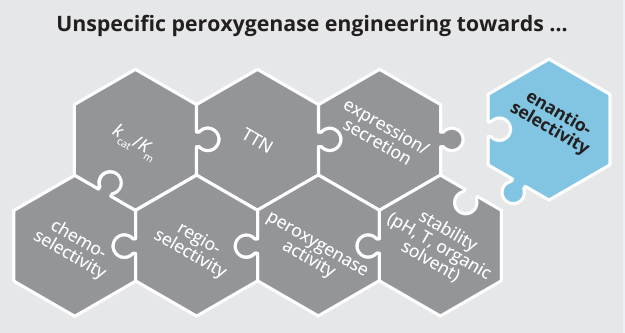
Judith Münch,[#] Jordi Soler,[#] Nicole Hünecke, Dominik Homann, Marc Garcia-Borràs,* and Martin J. Weissenborn*

Cite This: *ACS Catal.* 2023, 13, 8963–8972

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Unspecific peroxygenases (UPOs) perform oxyfunctionalizations for a wide range of substrates utilizing H_2O_2 without the need for further reductive equivalents or electron transfer chains. Tailoring these promising enzymes toward industrial application was intensely pursued in the last decade with engineering campaigns addressing the heterologous expression, activity, stability, and improvements in chemo- and regioselectivity. One hitherto missing integral part was the targeted engineering of enantioselectivity for specific substrates with poor starting enantioselectivity. In this work, we present the engineering of the short-type *Mth*UPO toward the enantiodivergent hydroxylation of the terpene model substrate, β-ionone. Guided by computational modeling, we designed a small smart library and screened it with a GC−MS setup. After two rounds of iterative protein evolution, the activity increased up to 17-fold and reached a regioselectivity of up to 99.6% for the 4-hydroxy-β-ionone. Enantiodivergent variants were identified with enantiomeric ratios of 96.6:3.4 (*R*) and 0.3:99.7 (*S*), respectively.



KEYWORDS: *directed evolution, unspecific peroxygenase, terpenes, hydroxylation, β-ionone, computational-guided protein engineering*

INTRODUCTION

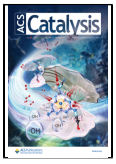
Oxyfunctionalyzation reactions are of tremendous importance in the field of synthetic chemistry, as they give access to new synthetic strategies, especially in late-stage functionalizations of complex molecules.[1] Nature offers an abundance of enzymes for catalyzing oxyfunctionalization reactions.[2,3] The most prominent and well-known class of enzymes in this field are the cytochrome P450 monooxygenases (P450s). P450s display a huge versatility of substrate-binding pockets and, thus, a tremendous substrate scope while maintaining high selectivity.[3,4] Two decades ago, a new enzyme class with similar properties emerged: the unspecific peroxygenases (UPOs) with fungal origin.[5−7] UPOs are inherently stable, secreted enzymes that catalyze reactions outside the regulated cell environment.[6] In contrast to P450s, they do not utilize molecular oxygen but hydrogen peroxide, in which the oxygen is already pre-reduced. This averts the necessity of cofactors such as NAD(P)H and a complex electron transport chain. Limited access to protocols for UPO production was formerly a huge drawback. Recent work gave access to the heterologous production of UPOs in fast-growing host organisms using protein engineering,[7] signal peptide,[8] and promoter shuffling.[9]

Engineering efforts at the mature protein led to UPO variants with improved peroxygenase/peroxidase ratios;[10] increases in activity (TON and k_{cat}/K_m);[11] improvements in thermo-,[12] pH-,[13] and solvent-stabilities;[12] and shifts in chemo- and regioselectivities.[10,14−16] Contrary to what their name suggests, UPOs often inherently exhibit excellent chemo-, regio-, and enantioselectivities for a range of substrates.[17] However, the engineering of UPOs for the oxyfunctionalization of specific substrates with initially poor enantioselectivities in a targeted manner remained elusive. A recent untargeted approach using the FuncLib algorithm led to impressive shifts in enantioselectivity.[18]

Our objective was to address the existing limitations and demonstrate the successful engineering of a UPO toward the enantioselective conversion of a selected substrate with low inherent enantioselectivity.

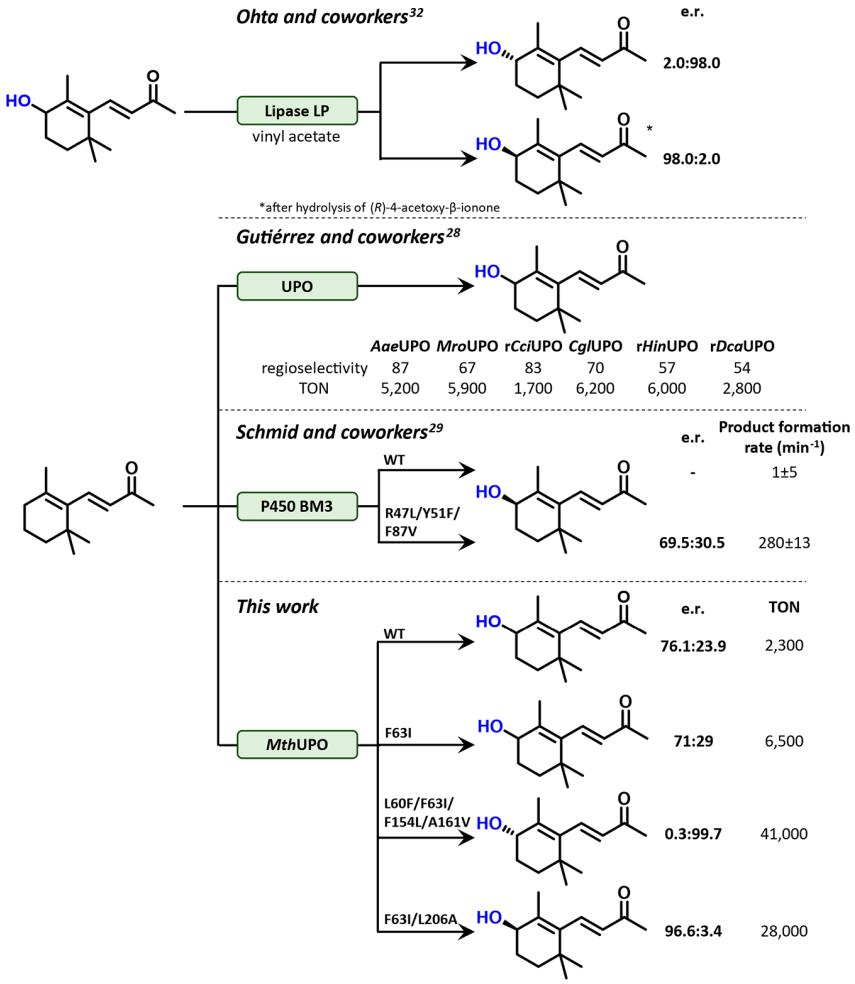
Particularly challenging targets for enantioselectivity engineering are terpenes and terpenoids. They are the largest group of natural products with approximately 80,000 different reported structures.[19] They occur mainly as secondary plant metabolites and display outstanding pharmacological bio-

Received: February 14, 2023
Revised: May 17, 2023
Published: June 21, 2023



https://doi.org/10.1021/acscatal.3c00702
ACS Catal. 2023, 13, 8963−8972

Scheme 1. Enzymatic Hydroxylation of β-Ionone



activity, making them attractive for medical and chemical industries. Their great structural diversity offers many opportunities for different regio- and enantioselective oxyfunctionalization reactions.

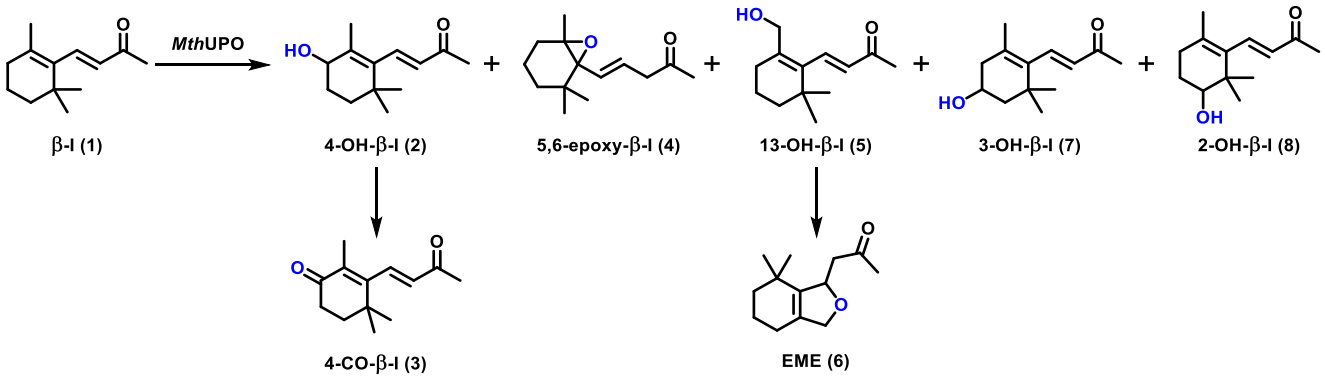
There have been several impressive engineering efforts of P450s toward terpenes and steroids. The engineering of $P450_{BM3}$ was demonstrated toward the enantiodivergent hydroxylation of five different steroid substrates with selectivity values between e.r. of 14:86 and 0:100. The variants showing this significant improvement in enantioselectivity were obtained through a directed evolution campaign with three rounds of CASTing. The library design was, hereby, based on molecular dynamics (MD) simulations and mutability landscaping.[20] Another work reported the engineering of $P450_{BM3}$ toward the regio- and stereoselective hydroxylation of the diterpenoid β-cembrenediol. The engineering campaign was based on the insertion of rational mutants and three sequential rounds of site saturation mutagenesis. The obtained variants were able to hydroxylate the C9 and C10 position, respectively, with regioselectivities up to 100% and a diastereometric ratio of 89:11 for C9 and 74:26 for C10.[21]

One interesting group of terpenoids are ionones—also known as rose ketones. They are highly valued compounds in the fragrance industry; the annual production of β-ionone amounts to 4000−8000 tons.[22] They further function as building blocks for the synthesis of many carotenoids and retinol (vitamin A).[23] Numerous pharmacological effects are described for ionones and their derivatives including anticancer, chemopreventive, cancer-promoting, melanogenesis, anti-inflammatory, and antimicrobial actions.[24] 3-hydroxy-β-ionone displays anti-cancer properties inhibiting progression and inducing apoptosis of SCC15 cells.[25] 5,6-epoxy-β-ionone inhibits the tumor-promoting agent 12-O-tetradecanoylphorbol-13-acetate (TPA) even more effectively compared to parent compound β-ionone.[26] 4-Hydroxy-β-ionone derivatives proved to be potent inhibitors of prostate cancer cell proliferation (LNCaP, MDA-PCa-2b, C4-2B, and 22Rv1) and full antagonists of the wild-type androgen receptor (AR) and various clinically important mutated ARs.[27]

Initial studies on the enzymatic oxyfunctionalization of β-ionone have been carried out (Scheme 1). The hydroxylation of α- and β-ionone performed by several UPOs led to a diverse range of hydroxylation and epoxidation products.[28] The hydroxylation of β-ionone has also been pursued using various P450s.[29−31] First, P450 engineering efforts showed an up to 280-fold increased product formation rate toward α- and β-ionone hydroxylations, however, enhancing the enantioselectivity proved challenging.[29] Enantioselective 4-hydroxy-β-ionone formation has been achieved both through enzymatic kinetic resolution[32] and by CYP2B6, which was recombinantly expressed in *Trichoplusia ni* cells.[31]

https://doi.org/10.1021/acscatal.3c00702
ACS Catal. 2023, 13, 8963−8972

Scheme 2. β-Ionone (1) Oxygenation by wt *Mth*UPO Showing the Different Oxygenated Derivatives[a]



[a]the regioselective distribution of the products for different enzyme variants can be found in Table S9.

With the maturation of computational modeling methods for the study of biocatalysts, new avenues are opening up for the generation of ever smaller and more intelligent protein library designs.[33]

In the present work, we engineered new UPO variants to enantioselectively access C4 hydroxylated stereoisomers of β-ionone. We applied a computational-aided engineering approach based on substrate-bound (restrained-) MD simulations to explore near-attack conformations (NACs) of the selective hydroxylation and characterized relevant binding modes of the model substrate β-ionone. This led to the identification of relevant residues for the substrate positioning and, hence, the design of a small smart library to alter the active site pocket of *Mth*UPO. In this way, we could direct the selectivity of the oxyfunctionalization toward enantioselective *R*/*S* C4 hydroxylation. The screening was performed by the previously developed multiple injection in a single experimental run (MISER) GC−MS method[14,16,34] focusing on activity increase. In the MISER setup, 96 samples are injected into the GC within one experimental run. Product quantifications are performed solely in the MS via different *m*/*z* ratios without the need for substrate/product separation, allowing an injection frequency of up to 30 s and, hence, a GC analysis of one microtiter plate within 48 min. The best variants were rescreened with a chiral GC−MS to determine enantioselectivities. This enabled the engineering of two highly active and enantiodivergent *Mth*UPO variants for (*R*/*S*)-4-hydroxy-β-ionone formation in two rounds of computationally guided enzyme evolution.

RESULTS AND DISCUSSION

Identification of the β-Ionone Hydroxylating UPO *Mth*UPO. We commenced with a pre-screening of five UPO enzymes to determine starting activities and selectivities, thereby, focusing on hydroxylation over epoxidation reactions (Table S4). We selected *Mth*UPO as an enzyme, as it displayed a regioselectivity of 88% for the main hydroxylation at the C4 position yielding 4-hydroxy-β-ionone (2), and a turnover number (TON) of approximately 2300. Further side products were β-ionone-5,6-epoxide (5,6-epoxy-β-I, 4), 7,11-epoxymegastigma-5(6)-en-9-one (EME, 6), 2-hydroxy-β-ionone (2-OH-β-I, 8), and 3-hydroxy-β-ionone (3-OH-β-I, 7, Scheme 2). Wild-type (wt) *Mth*UPO reached an enantiomeric ratio of 76:24 with a preference for (*R*)-4-hydroxy-β-ionone (2).

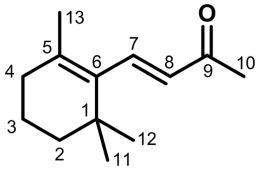
Development of an Engineering Strategy Based on Computational Modeling. The screening of different variants is often the most time-consuming step within a directed evolution campaign. Refined methods for creating small and smart mutant libraries have been developed to address this issue, including the combinatorial active-site saturation test (CAST),[35,36] iterative saturation mutagenesis (ISM),[37] or focused rational iterative site-specific mutagenesis (FRISM).[38] Inspired by these strategies, we utilized a computational-aided directed evolution strategy. This approach combined *in silico* analysis and experimental screening using the MISER GC−MS method.[14,16,34] The final aim was to engineer enantiodivergent enzyme variants that allow the selective formation of (*R*/*S*)-4-hydroxy-β-ionone products.

An accurate computational protein model was generated to reproduce the experimental observations for wt *Mth*UPO-catalyzed oxyfunctionalization of β-ionone. We combined density functional theory (DFT) calculations and MD simulations to characterize its reactivity pattern. DFT calculations were carried out using a truncated model, which includes the heme pyrrole core, a methyl thiolate to mimic the Cys axial ligation, and the β-ionone substrate (see Supporting Information for details). These data served to explore the intrinsic reactivity of β-ionone toward heme compound I. The transition states (TSs) were modeled for epoxidations and C−H activations of all non-equivalent C−H positions. The C−H activation was modeled as a hydrogen-atom transfer (HAT), which corresponds to the rate-limiting step of the hydroxylation reaction. As expected, calculations indicated that the epoxidation at cyclic double bond is energetically more favorable than the HAT C−H activation at the vinylic C4 position by only a few kcal·mol^{-1} (Figures 1 and S2). Other inactivated C−H positions exhibit higher HAT barriers.

To assess the accessible catalytically relevant binding modes of the substrate when bound in the active site and to study the catalyst control exerted by the enzyme considering its conformational landscape, we performed MD simulations with the β-ionone bound in wt *Mth*UPO's active site. The MD simulations in this step were performed using a previously generated computational model of wt *Mth*UPO.[14] Optimal geometric parameters required for C−H activation via HAT and epoxidation were taken from DFT model-optimized TSs (Figures S3 and S4).

MD simulations with the bound substrate indicated that β-ionone can explore NACs that could effectively lead to C−H activation via HAT at C4, C2, C10, C11, and C12 positions (Figures S3 and S4). The epoxidation conformations, however, barely explored NACs.

https://doi.org/10.1021/acscatal.3c00702
ACS Catal. 2023, 13, 8963−8972



Structure	ΔG‡
TS C5 epoxidation (q)	18.5
TS C8 epoxidation (d)	20.1
TS C2-HAT (q)	24.2
TS C3-HAT (q)	26.3
TS C4-HAT (d)	19.4
TS C10-HAT (q)	25.8
TS C11/12-HAT (q)	30.9
TS C13-HAT (d)	20.3

Figure 1. Intrinsic reactivity of β-ionone explored using a truncated computational model ("*theozyme*", i.e., theoretical enzyme). DFT calculated C−H activation via HAT and epoxidation TSs for all non-equivalent positions of β-ionone. Epoxidation of the cyclic double bond and α,β-alkene and hydroxylations at vinylic C4 (cyclic) and C13 (methyl) positions are intrinsically the most favored oxidation of β-ionone in the absence of any catalyst control. See Figure S2 for additional details. Energies are given in kcal·mol^{-1}.

Even though the C−O bond formation at C5 leading to epoxide formation is energetically preferred over the HAT at the C4 position (DFT optimized lowest TS, Figure 1), we experimentally observed only scarce amounts of epoxide product formation (Table S9).

This is in agreement with the fact that the C−O bond formation TS requires a very tight approach of the substrate to the catalytic Fe=O moiety (TS C5 epoxidation, Figure 2). This required NAC for epoxidation is sterically hindered in the active site of *Mth*UPO and cannot be efficiently explored by β-ionone—as revealed by MD simulations. On the other hand, the NAC required for C4 HAT is easily accessible for β-ionone (TS C4-HAT, Figures 2 and S5). This leads to a hampered epoxide formation while binding modes that allow the energetically most favored hydroxylation at position C4 are easily explored.

Taking all this together, computations described the catalyst control exerted by the wt *Mth*UPO on the selectivity oxyfunctionalization of β-ionone, reproducing the experimental observations, and provided a sound model to serve as the starting point for our semi-rational engineering strategy.

We attempted to improve our starting activity before focusing on shifting the enantioselectivity by screening an in-house *Mth*UPO library, which was obtained during a previous enzyme engineering campaign toward NBD hydroxylation (Figure S6). The most active variant F63I was selected as the parent variant for the engineering campaign displaying a 4.5-fold higher product formation rate than wt *Mth*UPO with similar regio- and enantioselectivities.

To restrict the accessible binding modes of β-ionone and, thus, control the stereoselectivity of the C4 hydroxylation, we aimed to design focused libraries based on computational predictions. To identify positions for the mutagenesis, MD simulations were performed with *Mth*UPO F63I in which β-ionone is forced to explore NACs for an effective C4−H HAT, mimicking DFT model TSs geometries (Figure 3A). Geometric restraints were included during the MD trajectories (restrained-MD simulations) in which the pro-*R* or pro-*S* C4−H positions—in independent trajectories—were forced to be at short distance to the Fe-oxo catalytic species (Figure 3A, see Supporting Information for details). These restrained-MD simulations provided structural descriptions of the arrangement of the active site that is required for accommodating β-ionone in a pro-R and pro-S NAC, respectively (Figure 3B). It is crucial to acknowledge that *Mth*UPO F63I already demonstrated the ability to catalyze C4−H activations, although the efficiency was limited, suggesting the presence

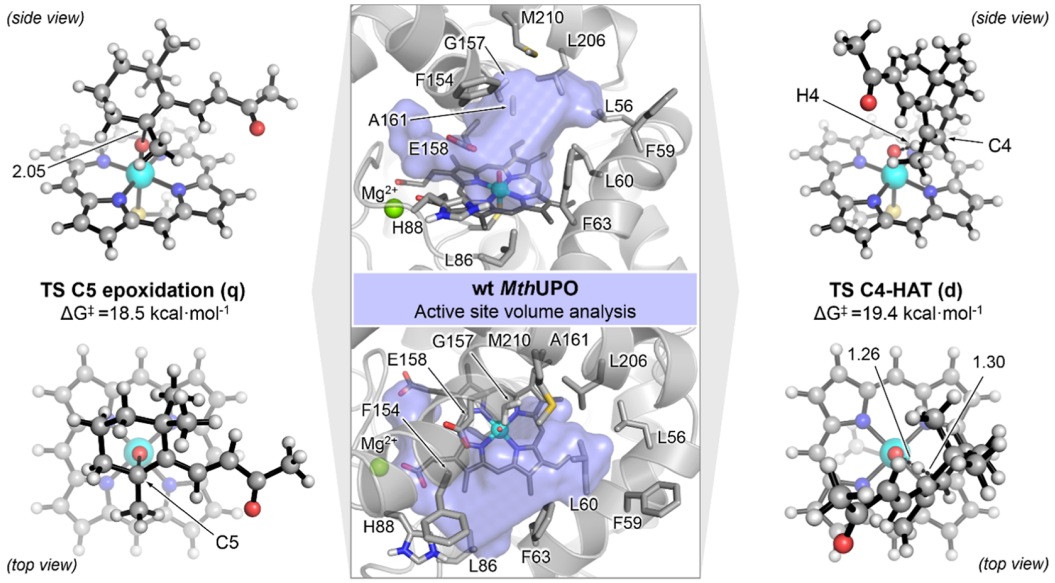


Figure 2. (Center) Characterized active site cavity of wt *Mth*UPO from *holo* state MD simulations. Representative structure of the most populated cluster (estimated form backbone RMSD analysis) obtained from five independent replicas of 1000 ns MD trajectories each (total of 5000 ns of accumulated simulation time).[14] DFT optimized, lowest in energy, TS geometries for: (Right) C4−H hydrogen atom transfer; and (Left) C5−C6 epoxidation. See Figures 1 and S2 for additional details. Energies are given in kcal·mol^{-1}, and key distances in Å. β-Ionone could easily bind in the available space in the active site in a catalytically relevant binding mode that resembles the NAC required for C4−H activation, but not for C5−C6 epoxidation. This is due to steric requirements and the shape of the active site cavity.

https://doi.org/10.1021/acscatal.3c00702
ACS Catal. 2023, 13, 8963−8972

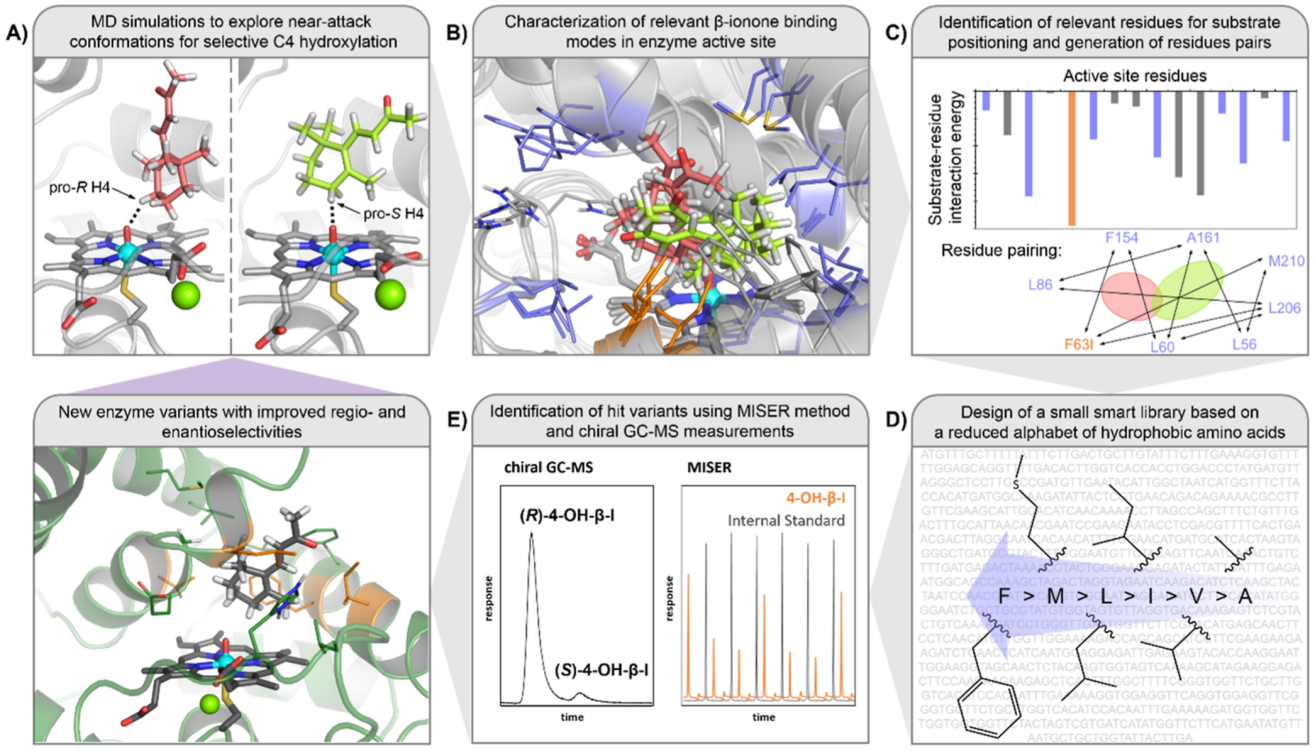


Figure 3. Schematic description of the computational-aided protein engineering protocol followed in this work. Red and green circles represent *β*-I substrate when bound in the active site cavity within different conformations. Complete descriptions of position selection for rounds 1 and 2 of protein engineering can be found in Figures S8 and S10, respectively.

of these incipient catalytically relevant binding modes that have yet to be fully exploited.

These simulations further revealed which active site amino acids exhibit strong interactions with the substrate, thus, affecting its binding mode. We reasoned that selecting positions as mutagenic spots that show strong steric interactions with the substrate during these restrained-MD simulations would give access to new variants with reshaped active sites. These new active sites are expected to better accommodate *β*-ionone in desired reactive conformations due to complementarity while destabilizing other alternative binding modes, thus inducing a shift in both regio- and enantioselectivity (Figure 3C). Following this reasoning, a total of eight positions (L56, L60, F63I, L86, F154, A161, L206, and M210) were initially selected based on restrained-MD analysis (Figure S8) and estimated substrate−residue interactions (estimated from MM-GBSA calculations, Figure S7). Amino acid positions that were determined to have a key role in proper substrate positioning (G157) or in H_2O_2 activation (H88 and E158) were not considered for mutagenesis.

Because targeted active site reshaping for enantioselectivity shifts might involve the reorientation of the substrate in the catalytic pocket, we hypothesized that mutating pairs of residues at the same time might facilitate this substrate reorientation due to cooperative effects.[39] Cooperativity, i.e., epistasis, between residue pairs can occur through direct interaction between residues but often also between distal indirect interactions.[40] These potential cooperative non-additive effects would be considered by including amino acid pairs for double saturation mutagenesis instead of saturation of single positions at a time.[36,41] We, hence, designed a set of 10 residue-pairs of proximal and—in terms of active site arrangement—also distal amino acid pairs combining previously selected active site residues (Figures 3C, a complete description of position selection can be found in Figure S8).

The *Mth*UPO active site is mainly composed of hydrophobic amino acids. The hydrophobic character was kept by employing a reduced amino acid alphabet consisting of Ala, Leu, Ile, Val, and Phe. For position M210, Met was also included (Figure 3D). This also reduces the number of possible variants for each double-saturation screening down to 25 (30 if position M210 is included). The screening is performed with a 3.5-fold oversampling leading to a library coverage of 97%. Inclusion of position M210 results in a 2.9-fold oversampling and 94% coverage.[42]

The screening was carried out in 96-well microtiter plates using the multiple injection in a single experimental run (MISER) (Figure 3E) setup,[14,16] allowing the determination of enzyme variant's activity but not its enantioselectivity. We reasoned that reshaping the active site results in an increased activity and simultaneously affects their enantioselectivity, as demonstrated multiple times before.[43]

Development of Enantiodivergent *Mth*UPO Variants in Two Rounds of Directed Evolution. In the first round, ten different double saturation pairs were screened (Table S5), adding up to more than 900 variants. The best variants showed increased activities of 2.0-fold under screening conditions compared to the parental variant *Mth*UPO F63I (Figure S15). With this first focused protein library, already impressive enantioselectivity improvements were detected (Table 1).

The best for the (*R*)-enantiomer formation showed an e.r. of 96.6:3.4 and a 4.3-fold improved activity of 28,000 TONs (*Mth*UPO F63I, L206A, coined R1A, Figure 4). In the same relatively small library, variants with inverted enantioselectivity

https://doi.org/10.1021/acscatal.3c00702
ACS Catal. 2023, 13, 8963−8972

Table 1. Enantioselectivity after One Round of Protein Engineering

entry	new mutations	e.r.
*Mth*UPO F63I (parent)		70.7:29.3
1	L206A	96.6:3.4
2	I63L, L206A	95.2:4.8
3	L56I, M210F	85.2:14.8
4	L206F	82.0:18.0
5	I63L, L206F	81.3:18.8
6	L206V	80.5:19.5
7	L56I, M210	79.5:20.5
8	L56V, A161I	29.4:70.6
9	I63V, F154I	12.0:88.0
10	L60F, A161I	11.5:88.5
11	I63V, F154L	7.3:92.7
12	L60F, A161V	3.3:96.7

could also be identified. The best showed an e.r. of 3.3:96.7 for the *S*-enantiomer and 20,000 TONs (*Mth*UPO L60F, F63I A161V, coined R1B).

Both variants were selected as parental variants for a second round of directed evolution (Figure S21).

Building upon the methodology established in the first iteration, we performed similar computational modeling. MD simulations, in this case without any restrain on β-ionone, were carried out to characterize and analyze the variants R1A and R1B (Figures S9 and S11). These simulations were used to design pairs of residues for library construction and subsequent screening, following the above-described protocol (Figures 3, S10, and S12).

In the second round, ten different double saturation pairs were screened with the parental variant R1A and eight different double saturation pairs with the parental variant R1B (Tables S6 and S7). The best variants based on R1A showed only a marginal increase in activity and no improved enantioselectivity. Based on these observations, we concluded that the evolutionary pathway had come to a local minimum and identified the best variant for (*R*)-4-hydroxy-β-ionone formation as *Mth*UPO F63I, L206A (R1A). This variant showed a high TON of 28,000 and an enantiomeric ratio of 96.6:3.4.

The best variant based on the *S*-selective variant R1B, however, displayed a substantial improvement. *Mth*UPO L60F, F63I, A161V, and F154L (R2B) displayed a 2.0-fold activity improvement to a TON of 41,000 and an e.r. of 99.7:0.3 (Figures 4 and S15, Table 2). The regioselective abundance of the main product also increased during the enzyme engineering campaign from 88% (*Mth*UPO wt expressed in *Saccharomyces cerevisiae*) to 99.2% (R1A) and 99.6% (R2B) (Table S9).

Table 2. Enantioselectivity after Round 2B of Protein Engineering

entry	new mutations	e.r.
*Mth*UPO L60F, F63I, A161V (parent)		3.3:96.7
1	F154L	0.3:99.7
2	F154I	0.5:99.5
3	F154V	0.4:99.6
4	F154I, M210L	0.5:99.5
5	F154L, M210L	0.4:99.6

With the best engineered variant in hand, we repeated the enzymatic reactions in a scale-up experiment to give access to larger amounts of the hydroxylated product and demonstrate its utility. We were able to obtain 34 mg of (*S*)-4-hydroxy-β-ionone (55%)-utilizing variant, R2B, as a catalyst. We further obtained 51.1 mg (*R*)-4-acetoxy-β-ionone (37.0%) after hydroxylation of the substrate through variant R1A and subsequent derivatization to determine the specific optical rotation.

Both enzyme variants were purified and their transition temperature (T_M) determined (Table S13). Both variants showed a decreased transition temperature from initial 58.1 °C (wt MthUPO) to 53.7 °C (R1A) and 52.9 °C (R2B). To gain insights into the kinetics of the variants, the apparent catalytic parameters K_m and k_{cat} were determined for the substrate NBD (5-nitro-1,3-benzodioxole) and the co-substrate H_2O_2 (Table S14 and Figure S22). Both variants show an improved apparent k_{cat}/K_m (1.9−4-fold) for both substrates relative to the wildtype (NBD: 1.9×10^4 M^{-1} s^{-1}, H_2O_2: 1.3×10^4 M^{-1} s^{-1}). Especially, variant R2B revealed a strong improvement of the apparent k_{cat} toward both NBD (33.8 s^{-1}) and H_2O_2 (67.1 s^{-1}) in comparison with wt *Mth*UPO (k_{cat} 7.1 s^{-1}). This is consistence with our previous work, demonstrating the beneficial influence of the mutations L60F and F63I for the NBD conversion.[14]

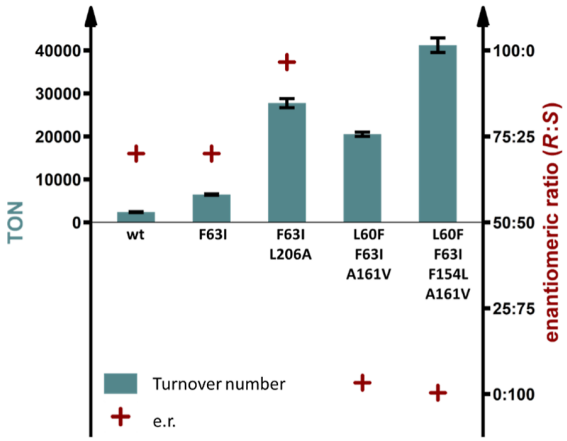


Figure 4. Bar chart showing TON within 1 h of allylic hydroxylation of β-ionone by different *Mth*UPO variants. Turnover data are mean ± s.d. of measurements in triplicates. TON (teal bars) determined by GC−MS, and enantiomeric excess (red cross) determined by chiral GC−MS (Figure S20).

Molecular Basis of the Highly Enantioselective Variants R1A and R2B. Finally, we performed computational modeling to rationalize the molecular basis for the enantiodivergence exhibited by R1A and R2B engineered variants. MD simulations with β-ionone bound in the active site of R1A and R2B—without including any geometric restraints during simulations—were carried out. MD simulations revealed the anticipated hydrophobic and steric interactions that control the binding modes for β-ionone and its orientation relative to the catalytically active Fe-oxo (Compound I) species.

MD simulations for R1A variant showed that β-ionone explores a major binding mode, in which only the pro-*R* C4−H is well oriented in a catalytically relevant NAC toward Compound I for HAT (Figures 5A and S9). This catalytically

https://doi.org/10.1021/acscatal.3c00702
ACS Catal. 2023, 13, 8963−8972

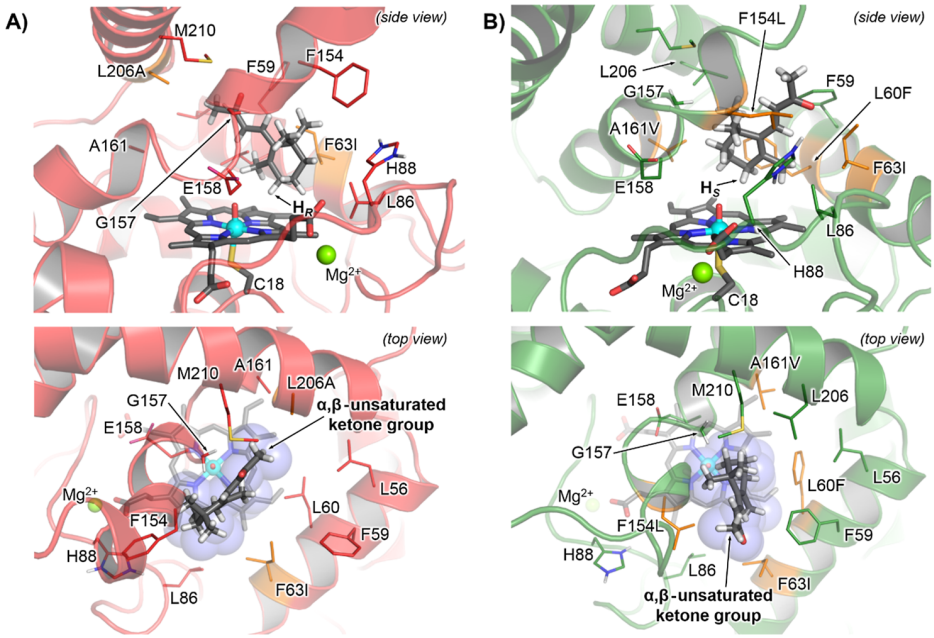


Figure 5. *β*-Ionone-binding modes characterized from MD simulations in (A) variant R1A (F63I/L206A, in red) and (B) variant R2B (L60F/F63I/F154L/A161V, in green). See Supporting Information for additional details.

relevant binding mode is stabilized due to the space generated by the initial F63I mutation, which is next to the region where C4 is oriented in the active site. The newly introduced L206A mutation reduces the steric hindrance, and it enables placing the *α,β*-unsaturated ketone group of the substrate and, thus, accommodating the *β*-ionone allylic chain in this region of the active site (Figure 5A). Within this favored binding mode of *β*-ionone, the pro-*R* C4−H bond is in a preorganized reactive conformation to promote a selective, pro-*R* HAT.

An opposite *β*-ionone positioning is observed in the R2B variant, where only the pro-*S* C4−H can explore reactive conformations relative to the Fe-oxo species (Figures 5B and S13). Unrestricted MD simulations revealed that when *β*-ionone is bound in R2B's active site, it preferentially explores a binding mode in which the *α,β*-unsaturated ketone group is oriented toward F154L position while keeping the six-membered ring in a similar position in the active site, as observed in R1A. This binding mode is favored by the A161V mutation that, together with the L206 residue, introduces bulkiness at this region of the active site. This sterically disfavors the positioning of the *α,β*-unsaturated ketone group in this area of the active site—contrary to R1A. The *α,β*-unsaturated ketone group is, thus, oriented toward the opposite side of the binding pocket where the F154L mutation is creating more empty space for accommodating this allylic substituent of *β*-ionone. The L60F mutation reduces the accessible space in that inner region of the active site, sterically disfavoring the positioning of the dimethyl and methyl *β*-ionone ring substituents there. All these factors synergistically favor the orientation of pro-*S* C4−H in a NAC toward the Fe-oxo species to selectively react via HAT.

These computational insights are in perfect agreement with the experimentally observed properties of the highly enantioselective variants, R1A and R2B.

CONCLUSIONS

The presented work revealed the first substrate-targeted engineering of an UPO toward enantioselective C−H hydroxylation, starting from a poor enantioselectivity. Based on *in silico* predictions of the positioning for the substrate in the active site, excellent activities as well as regio- and enantioselectivities could be achieved by introducing very few mutations. In only two rounds of mutagenesis and the screening of 1600 variants, it was possible to evolve the wild-type *Mth*UPO with poor e.r. of 76:24 (*R*/*S*) toward the two enantiodivergent variants *Mth*UPO F63I, L206A (e.r. 96.6:3.4, (*R*)-selective) and *Mth*UPO L60F, F63I, F154L, A161V (e.r. 0.3:99.7, (*S*)-selective). In addition to high enantioselectivity, both variants also showed excellent regioselectivities with more than 99.2% and the activity increased up to 17-fold achieving TONs of 41,000.

The active site of *Mth*UPO seems to be particularly well-suited for selectivity engineering approaches due to its high hydrophobicity. This leads to two advantages: (i) significant high affinity toward different aliphatic substrates and (ii) the possibility of using a substantially reduced amino acid alphabet, exclusively consisting of hydrophobic amino acids, for active site reshaping while keeping the hydrophobicity and lowering the screening effort dramatically.

Further engineering of regio- and enantioselective functionalization of terpenes, terpenoids, linear- or branched hydrocarbon chains could be performed, similar to the presented work with the model substrate *β*-ionone.

Our MISER screening method based on GC−MS analysis offers a broad substrate flexibility, only limited to volatile, thermal stable molecules. This is an advantage in contrast to systems in which only chromogenic substrates can be examined. Even with the MISER approach, the screening remains the time-limiting factor; hence, the generation of a small smart library is still of utmost importance for reducing the overall engineering time. This is where the advantages

https://doi.org/10.1021/acscatal.3c00702
ACS Catal. 2023, 13, 8963−8972

offered by *in silico* analysis will be even more relevant in the future.

The present work paves the way toward the rapid engineering of UPOs for enantioselective conversion of terpene derivatives and other substrates and, hence, solves one of the last remaining challenges in UPO research.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acscatal.3c00702.

Material and methods; computational methods and protocols; gene and protein sequence *Mth*UPO; energies, thermochemistry parameters and Cartesian coordinates of DFT optimized structures; oligonucleotides for cloning; initial screening data; saturation pairs for directed evolution campaign; TONs, regioselectivities and enantioselectivities for all variants; GC−MS parameters; chemical structures of all products; MD simulations and DFT calculations; screening data from directed evolution campaign; calibration curves; GC−MS chromatograms of non-chiral and chiral measurements; and NMR spectra (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Marc Garcia-Borràs − *Institut de Química Computacional i Catàlisi and Departament de Química, Universitat de Girona, 17003 Girona, Catalonia, Spain;* orcid.org/0000-0001-9458-1114; Email: marc.garcia@udg.edu

Martin J. Weissenborn − *Department of Chemistry, Martin Luther-University Halle-Wittenberg, 06120 Halle (Saale), Germany; Leibniz Institute of Plant Biochemistry, 06120 Halle (Saale), Germany;* orcid.org/0000-0002-1200-4485; Email: martin.weissenborn@chemie.uni-halle.de

Authors

Judith Münch − *Department of Chemistry, Martin Luther-University Halle-Wittenberg, 06120 Halle (Saale), Germany; Leibniz Institute of Plant Biochemistry, 06120 Halle (Saale), Germany*

Jordi Soler − *Institut de Química Computacional i Catàlisi and Departament de Química, Universitat de Girona, 17003 Girona, Catalonia, Spain*

Nicole Hünecke − *Department of Chemistry, Martin Luther-University Halle-Wittenberg, 06120 Halle (Saale), Germany*

Dominik Homann − *Department of Chemistry, Martin Luther-University Halle-Wittenberg, 06120 Halle (Saale), Germany; Leibniz Institute of Plant Biochemistry, 06120 Halle (Saale), Germany*

Complete contact information is available at: https://pubs.acs.org/10.1021/acscatal.3c00702

Author Contributions

#J.M. and J.S. contributed equally. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

J.M. thanks the Friedrich-Naumann-Stiftung für die Freiheit for a PhD scholarship. D.H. thanks the Friedrich-Ebert Stiftung for a PhD scholarship. M.J.W. and J.M. thank the Bundesministerium für Bildung und Forschung (Maßgeschneiderte Inhaltsstoffe 2, 031B0834A) and M.J.W., J.M., and D.H. thank the German Research Foundation (DFG, project ID 43649874, TP A05, RTG 2670) for generous funding. M.G.B. thanks the Spanish MICINN (Ministerio de Ciencia e Innovación) for PID2019-111300GA-I00 project and the Ramón y Cajal program via the RYC 2020-028628-I fellowship. J.S. thanks the Spanish MIU (Ministerio de Universidades) for a predoctoral FPU fellowship FPU18/02380. We would like to thank Dr. Franziska Seifert (Martin Luther University Halle-Wittenberg) for discussions and providing access to the DSF device for thermostability measurements.

■ ABBREVIATIONS

CAST, combinatorial active-site saturation test; DFT, density functional theory; GC−MS, gas chromatography−mass spectrometry; HAT, hydrogen-atom transfer; MD, Molecular Dynamics; MISER, multiple injection in a single chromatographical run for gas chromatography−mass spectrometry; NAC, near-attack conformation; T_M, transition temperature; TON, turnover number; TS, transition state; UPO, unspecific peroxygenase

■ REFERENCES

(1) (a) White, M. C.; Zhao, J. Aliphatic C−H oxidations for late-stage functionalization. *J. Am. Chem. Soc.* **2018**, *140*, 13988−14009. (b) Saint-Denis, T. G.; Zhu, R.-Y.; Chen, G.; Wu, Q.-F.; Yu, J.-Q. Enantioselective C (sp3)−H bond activation by chiral transition metal catalysts. *Science* **2018**, *359*, 6377. (c) Crabtree, R. H.; Lei, A. *Introduction: CH Activation*; ACS Publications, 2017; Vol. *117*, pp 8481−8482. (d) Chakrabarty, S.; Wang, Y.; Perkins, J. C.; Narayan, A. R. Scalable biocatalytic C−H oxyfunctionalization reactions. *Chem. Soc. Rev.* **2020**, *49*, 8137−8155.

(2) Chapman, J.; Ismail, A. E.; Dinu, C. Z. Industrial applications of enzymes: Recent advances, techniques, and outlooks. *Catalysts* **2018**, *8*, 238.

(3) Münch, J.; Püllmann, P.; Zhang, W.; Weissenborn, M. J. Enzymatic Hydroxylations of sp3-Carbons. *ACS Catal.* **2021**, *11*, 9168−9203.

(4) Urlacher, V. B.; Girhard, M. Cytochrome P450 monooxygenases in biotechnology and synthetic biology. *Trends Biotechnol.* **2019**, *37*, 882−897.

(5) (a) Beltrán-Nogal, A.; Sánchez-Moreno, I.; Méndez-Sánchez, D.; Gómez de Santos, P.; Hollmann, F.; Alcalde, M. Surfing the wave of oxyfunctionalization chemistry by engineering fungal unspecific peroxygenases. *Curr. Opin. Struct. Biol.* **2022**, *73*, 102342. (b) Wang, Y.; Lan, D.; Durrani, R.; Hollmann, F. Peroxygenases en route to becoming dream catalysts. What are the opportunities and challenges? *Curr. Opin. Chem. Biol.* **2017**, *37*, 1−9. (c) Molina-Espeja, P.; Santos, P. G. d.; Alcalde, M. Directed evolution of unspecific peroxygenase. In *Directed enzyme evolution: advances and applications*; Springer, 2017, pp 127−143.

(6) Hofrichter, M.; Kellner, H.; Pecyna, M. J.; Ullrich, R. Fungal unspecific peroxygenases: heme-thiolate proteins that combine peroxidase and cytochrome P450 properties. In *Monooxygenase, peroxidase and Peroxygenase properties and mechanisms of cytochrome P450*; Springer, 2015, pp 341−368.

(7) Molina-Espeja, P.; Garcia-Ruiz, E.; Gonzalez-Perez, D.; Ullrich, R.; Hofrichter, M.; Alcalde, M. Directed evolution of unspecific peroxygenase from Agrocybe aegerita. *Appl. Environ. Microbiol.* **2014**, *80*, 3496−3507.

(8) Püllmann, P.; Knorrscheidt, A.; Münch, J.; Palme, P. R.; Hoehenwarter, W.; Marillonnet, S.; Alcalde, M.; Westermann, B.; Weissenborn, M. J. A modular two yeast species secretion system for the production and preparative application of unspecific peroxygenases. *Commun. Biol.* **2021**, *4*, 562.
(9) Püllmann, P.; Weissenborn, M. J. Improving the heterologous production of fungal peroxygenases through an episomal Pichia pastoris promoter and signal peptide shuffling system. *ACS Synth. Biol.* **2021**, *10*, 1360−1372.
(10) Molina-Espeja, P.; Cañellas, M.; Plou, F. J.; Hofrichter, M.; Lucas, F.; Guallar, V.; Alcalde, M. Synthesis of 1-Naphthol by a Natural Peroxygenase Engineered by Directed Evolution. *ChemBioChem* **2016**, *17*, 341−349.
(11) Gomez de Santos, P.; Cañellas, M.; Tieves, F.; Younes, S. H.; Molina-Espeja, P.; Hofrichter, M.; Hollmann, F.; Guallar, V.; Alcalde, M. Selective synthesis of the human drug metabolite 5′-hydroxypropranolol by an evolved self-sufficient peroxygenase. *ACS Catal.* **2018**, *8*, 4789−4799.
(12) Martin-Diaz, J.; Paret, C.; García-Ruiz, E.; Molina-Espeja, P.; Alcalde, M. Shuffling the neutral drift of unspecific peroxygenase in Saccharomyces cerevisiae. *Appl. Environ. Microbiol.* **2018**, *84*, e00808−e00818.
(13) Gomez de Santos, P.; Hoang, M. D.; Kiebist, J.; Kellner, H.; Ullrich, R.; Scheibner, K.; Hofrichter, M.; Liers, C.; Alcalde, M. Functional expression of two unusual acidic peroxygenases from Candolleomyces aberdarensis in yeasts by adopting evolved secretion mutations. *Appl. Environ. Microbiol.* **2021**, *87*, No. e0087821.
(14) Knorrscheidt, A.; Soler, J.; Hünecke, N.; Püllmann, P.; Garcia-Borràs, M.; Weissenborn, M. J. Accessing Chemo-and Regioselective Benzylic and Aromatic Oxidations by Protein Engineering of an Unspecific Peroxygenase. *ACS Catal.* **2021**, *11*, 7327−7338.
(15) Carro, J.; González-Benjumea, A.; Fernández-Fueyo, E.; Aranda, C.; Guallar, V.; Gutiérrez, A.; Martínez, A. T. Modulating Fatty Acid Epoxidation vs Hydroxylation in a Fungal Peroxygenase. *ACS Catal.* **2019**, *9*, 6234−6242.
(16) Knorrscheidt, A.; Soler, J.; Hünecke, N.; Püllmann, P.; Garcia-Borràs, M.; Weissenborn, M. J. Simultaneous screening of multiple substrates with an unspecific peroxygenase enabled modified alkane and alkene oxyfunctionalisations. *Catal. Sci. Technol.* **2021**, *11*, 6058−6064.
(17) (a) González-Benjumea, A.; Linde, D.; Carro, J.; Ullrich, R.; Hofrichter, M.; Martínez, A. T.; Gutiérrez, A. Regioselective and stereoselective epoxidation of n-3 and n-6 fatty acids by fungal peroxygenases. *Antioxidants* **2021**, *10*, 1888. (b) Bassanini, I.; Ferrandi, E. E.; Vanoni, M.; Ottolina, G.; Riva, S.; Crotti, M.; Brenna, E.; Monti, D. Peroxygenase-Catalyzed Enantioselective Sulfoxidations. *Eur. J. Org. Chem.* **2017**, *2017*, 7186−7189. (c) Xu, X.; Paul, C.; Hollmann, F. Peroxygenase-catalysed enantioselective hydroxylation of ethylbenzene derivatives driven by formate oxidase. In *Expanding the Scope of H_2O_2-Driven Biocatalysis*; Doctoral Thesis; Delft University of Technology, 2022; p 99.
(18) Gomez de Santos, P.; Mateljak, I.; Hoang, M. D.; Fleishman, S. J.; Hollmann, F.; Alcalde, M. Repertoire of Computationally Designed Peroxygenases for Enantiodivergent C−H Oxyfunctionalization Reactions. *J. Am. Chem. Soc.* **2023**, *145*, 3443−3453.
(19) Christianson, D. W. Structural and chemical biology of terpenoid cyclases. *Chem. Rev.* **2017**, *117*, 11570−11648.
(20) Acevedo-Rocha, C. G.; Gamble, C. G.; Lonsdale, R.; Li, A.; Nett, N.; Hoebenreich, S.; Lingnau, J. B.; Wirtz, C.; Fares, C.; Hinrichs, H.; et al. P450-catalyzed regio-and diastereoselective steroid hydroxylation: efficient directed evolution enabled by mutability landscaping. *ACS Catal.* **2018**, *8*, 3395−3410.
(21) Le-Huu, P.; Heidt, T.; Claasen, B.; Laschat, S.; Urlacher, V. B. Chemo-regio-and stereoselective oxidation of the monocyclic diterpenoid β-cembrenediol by P450 BM3. *ACS Catal.* **2015**, *5*, 1772−1780.
(22) Beekwilder, J.; van Rossum, H. M.; Koopman, F.; Sonntag, F.; Buchhaupt, M.; Schrader, J.; Hall, R. D.; Bosch, D.; Pronk, J. T.; van Maris, A. J.; et al. Polycistronic expression of a β-carotene biosynthetic pathway in Saccharomyces cerevisiae coupled to β-ionone production. *J. Biotechnol.* **2014**, *192*, 383−392.
(23) Parker, G. L.; Smith, L. K.; Baxendale, I. R. Development of the industrial synthesis of vitamin A. *Tetrahedron* **2016**, *72*, 1645−1652.
(24) (a) Ansari, M.; Emami, S. β-Ionone and its analogs as promising anticancer agents. *Eur. J. Med. Chem.* **2016**, *123*, 141−154. (b) Aloum, L.; Alefishat, E.; Adem, A.; Petroianu, G. Ionone is more than a violet's fragrance: A review. *Molecules* **2020**, *25*, 5822.
(25) Luetragoon, T.; Pankla Sranujit, R.; Noysang, C.; Thongsri, Y.; Potup, P.; Suphrom, N.; Nuengchamnong, N.; Usuwanthim, K. Anti-Cancer effect of 3-Hydroxy-β-Ionone identified from moringa oleifera Lam. Leaf on human squamous cell carcinoma 15 cell line. *Molecules* **2020**, *25*, 3563.
(26) Wertz, P. W.; Kensler, T. W.; Mueller, G. C. Inhibition of phorbol ester action in lymphocytes by 5, 6-epoxy-β-ionone. *Biochem. Biophys. Res. Commun.* **1978**, *83*, 138−143.
(27) Zhou, J.; Geng, G.; Wu, J. H. Synthesis and in vitro characterization of ionone-based chalcones as novel antiandrogens effective against multiple clinically relevant androgen receptor mutants. *Invest. New Drugs* **2010**, *28*, 291−298.
(28) Babot, E. D.; Aranda, C.; del Río, J. C.; Ullrich, R.; Kiebist, J.; Scheibner, K.; Hofrichter, M.; Martínez, A. T.; Gutiérrez, A. Selective oxygenation of ionones and damascones by fungal peroxygenases. *J. Agric. Food Chem.* **2020**, *68*, 5375−5383.
(29) Urlacher, V. B.; Makhsumkhanov, A.; Schmid, R. D. Biotransformation of β-ionone by engineered cytochrome P450 BM-3. *Appl. Microbiol. Biotechnol.* **2006**, *70*, 53−59.
(30) Çelik, A.; Flitsch, S. L.; Turner, N. J. Efficient terpene hydroxylation catalysts based upon P450 enzymes derived from actinomycetes. *Org. Biomol. Chem.* **2005**, *3*, 2930−2934.
(31) Marumoto, S.; Shimizu, R.; Tanabe, G.; Okuno, Y.; Miyazawa, M. In Vitro Regio- and Stereoselective Oxidation of β-Ionone by Human Liver Microsomes. *Planta Med.* **2016**, *83*, 292−299.
(32) Kakeya, H.; Sugai, T.; Ohta, H. Biochemical preparation of optically active 4-hydroxy-β-ionone and its transformation to (S)-6-hydroxy-α-ionone. *Agric. Biol. Chem.* **1991**, *55*, 1873−1876.
(33) (a) Wijma, H. J.; Floor, R. J.; Bjelic, S.; Marrink, S. J.; Baker, D.; Janssen, D. B. Enantioselective enzymes by computational design and in silico screening. *Angew. Chem., Int. Ed.* **2015**, *54*, 3726−3730. (b) Cadet, F.; Fontaine, N.; Li, G.; Sanchis, J.; Ng Fuk Chong, M.; Pandjaitan, R.; Vetrivel, I.; Offmann, B.; Reetz, M. T. A machine learning approach for reliable prediction of amino acid interactions and its application in the directed evolution of enantioselective enzymes. *Sci. Rep.* **2018**, *8*, 16757−16815. (c) Li, G.; Dong, Y.; Reetz, M. T. Can machine learning revolutionize directed evolution of selective enzymes? *Adv. Synth. Catal.* **2019**, *361*, 2377−2386. (d) Rosales, A. R.; Wahlers, J.; Limé, E.; Meadows, R. E.; Leslie, K. W.; Savin, R.; Bell, F.; Hansen, E.; Helquist, P.; Munday, R. H.; et al. Rapid virtual screening of enantioselective catalysts using CatVS. *Nat. Catal.* **2018**, *2*, 41−45. (e) Romero-Rivera, A.; Garcia-Borràs, M.; Osuna, S. Computational tools for the evaluation of laboratory-engineered biocatalysts. *Chem. Commun.* **2017**, *53*, 284−297.
(34) Knorrscheidt, A.; Püllmann, P.; Schell, E.; Homann, D.; Freier, E.; Weissenborn, M. J. Identification of novel unspecific peroxygenase chimeras and unusual YfeX axial heme ligand by a versatile high-throughput GC-MS approach. *ChemCatChem* **2020**, *12*, 4788−4795.
(35) Reetz, M. T.; Wang, L. W.; Bocola, M. Directed evolution of enantioselective enzymes: iterative cycles of CASTing for probing protein-sequence space. *Angew. Chem., Int. Ed.* **2006**, *45*, 1236−1241.
(36) Reetz, M. T.; Bocola, M.; Carballeira, J. D.; Zha, D.; Vogel, A. Expanding the range of substrate acceptance of enzymes: combinatorial active-site saturation test. *Angew. Chem., Int. Ed.* **2005**, *44*, 4192−4196.
(37) (a) Acevedo-Rocha, C. G.; Hoebenreich, S.; Reetz, M. T. Iterative saturation mutagenesis: a powerful approach to engineer proteins by systematically simulating Darwinian evolution. In *Directed Evolution Library Creation*; Springer, 2014, pp 103−128. (b) Reetz, M. T.; Carballeira, J. D. Iterative saturation mutagenesis (ISM) for rapid

https://doi.org/10.1021/acscatal.3c00702
ACS Catal. 2023, 13, 8963−8972

directed evolution of functional enzymes. *Nat. Protoc.* **2007**, *2*, 891−903.
(38) Li, D.; Wu, Q.; Reetz, M. T. Focused rational iterative site-specific mutagenesis (FRISM). *Methods in Enzymology*; Elsevier, 2020; Vol. *643*, pp 225−242.
(39) (a) Acevedo-Rocha, C. G.; Li, A.; D'Amore, L.; Hoebenreich, S.; Sanchis, J.; Lubrano, P.; Ferla, M. P.; Garcia-Borràs, M.; Osuna, S.; Reetz, M. T. Pervasive cooperative mutational effects on multiple catalytic enzyme traits emerge via long-range conformational dynamics. *Nat. Commun.* **2021**, *12*, 1621. (b) Bocola, M.; Otte, N.; Jaeger, K. E.; Reetz, M. T.; Thiel, W. Learning from directed evolution: theoretical investigations into cooperative mutations in lipase enantioselectivity. *ChemBioChem* **2004**, *5*, 214−223. (c) Reetz, M. T.; Puls, M.; Carballeira, J. D.; Vogel, A.; Jaeger, K. E.; Eggert, T.; Thiel, W.; Bocola, M.; Otte, N. Learning from directed evolution: further lessons from theoretical investigations into cooperative mutations in lipase enantioselectivity. *ChemBioChem* **2007**, *8*, 106−112.
(40) Miton, C. M.; Tokuriki, N. How mutational epistasis impairs predictability in protein evolution and design. *Protein Sci.* **2016**, *25*, 1260−1272.
(41) Sun, Z.; Lonsdale, R.; Li, G.; Reetz, M. T. Comparing different strategies in directed evolution of enzyme stereoselectivity: single- versus double-code saturation mutagenesis. *ChemBioChem* **2016**, *17*, 1865−1872.
(42) Reetz, M. T.; Kahakeaw, D.; Lohmer, R. Addressing the numbers problem in directed evolution. *ChemBioChem* **2008**, *9*, 1797−1804.
(43) (a) van der Meer, J.-Y.; Poddar, H.; Baas, B.-J.; Miao, Y.; Rahimi, M.; Kunzendorf, A.; van Merkerk, R.; Tepper, P. G.; Geertsema, E. M.; Thunnissen, A.-M. W.; et al. Using mutability landscapes of a promiscuous tautomerase to guide the engineering of enantioselective Michaelases. *Nat. Commun.* **2016**, *7*, 10911−10916. (b) Zhang, D.; Chen, X.; Chi, J.; Feng, J.; Wu, Q.; Zhu, D. Semi−Rational engineering a carbonyl reductase for the enantioselective reduction of *β*-amino ketones. *ACS Catal.* **2015**, *5*, 2452−2457. (c) Chen, K.; Arnold, F. H. Engineering cytochrome P450s for enantioselective cyclopropenation of internal alkynes. *J. Am. Chem. Soc.* **2020**, *142*, 6891−6895.

https://doi.org/10.1021/acscatal.3c00702
ACS Catal. 2023, 13, 8963−8972

7 | Chapter V

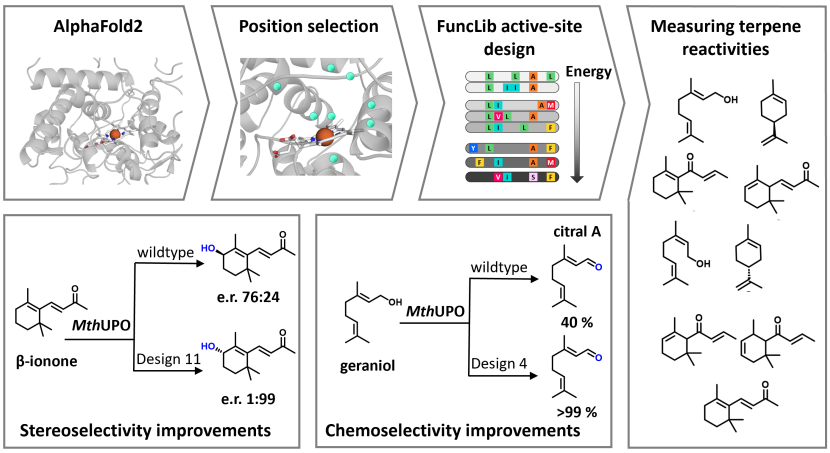
This chapter has been published as:

Computationally Designed Peroxygenases That Exhibit Diverse and Selective Terpene Oxyfunctionalization

by: **Judith Münch**, Jordi Soler, Ofir Gildor-Christal, Sarel J. Fleishman, Marc Garcia-Borràs, and Martin J. Weissenborn

in: *ACS Catalysis* **2025**, *15* (15), 12741–12755; doi: https://doi.org/10.1021/acscatal.5c02412

Supplementary information can be assessed free of charge at the article's landing page.



This chapter describes the application of the FuncLib algorithm to generate a small yet diverse library of 50 *Mth*UPO variants, designed from an AlphaFold2 structural model to remodel the active site for improved terpene oxyfunctionalization. All variants retained 100% functional expression, and each showed turnover on at least one substrate from the test panel. For terpene oxidations, the most active design for each substrate exhibited 2.2-fold to 7.1-fold higher activity compared to the wild type, with up to 26 designs surpassing the wild type by $\geq$ 2-fold for the top-performing substrate. Selectivity was also improved: regioselectivity for 3-hydroxy-β-damascone increased from 3% to 46%, chemoselectivity for geraniol oxidation to citral A rose from 40% to >99%, and citral B from 72% to 89%. Several designs enabled novel reactivities, including isopiperitenol formation from limonene and epoxide products from geraniol and nerol. In β-ionone hydroxylation, enantioselectivity was inverted from an (R):(S) ratio of 76:24 to 1:99. These gains often arose from combinations of mutations with strong epistatic effects, underscoring the power of multi-residue, structure-guided enzyme design.

ACS Catalysis

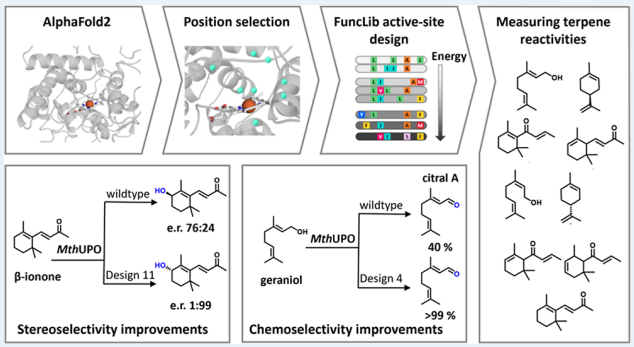
pubs.acs.org/acscatalysis
Research Article

Computationally Designed Peroxygenases That Exhibit Diverse and Selective Terpene Oxyfunctionalization

Judith Münch, Jordi Soler, Ofir Gildor-Cristal, Sarel J. Fleishman,* Marc Garcia-Borràs,* and Martin J. Weissenborn*

Cite This: *ACS Catal.* 2025, 15, 12741−12755

ABSTRACT: The selective oxyfunctionalization of terpenes remains a major challenge in chemical synthesis and is of significant industrial importance. This study presents a computational enzyme design approach based on an AlphaFold2 model of an unspecific peroxygenase (*Mth*UPO). Using the FuncLib algorithm, only 50 variants were required, and they exhibit remarkable advancements. All 50 designs retained 100% measurable activity across the tested substrate panel, with each design showing activity on at least one substrate. Among the terpene substrates, improvements in activity varied considerably: while some substrates had only a single design exhibiting a ≥2-fold increase in activity, the top-performing substrate had 26 such designs. The most active design per terpene substrate showed enhancements ranging from 2.2-fold to 7.1-fold relative to the wild type. In addition to increased activity, many designs also demonstrated useful and dramatic shifts in regio-, chemo-, and stereoselectivity. Regioselectivity for the energetically less favored 3-hydroxy-β-damascone increased from 3 to 46%. Particularly striking is the dramatic improvement in chemoselectivity for the oxidation of geraniol and nerol to citral A (>99%) and citral B (89%), respectively. While wild-type *Mth*UPO exhibited only a moderate selectivity of 40% for citral A and 72% for citral B, our computationally designed variants displayed significantly enhanced product preference and up to a 4.5-fold increase in activity. Additionally, further products not found with the wild-type enzyme, such as isopiperitenol from limonene and epoxides from geraniol and nerol, were synthesized. For the hydroxylation of β-ionone, the enantioselectivity was inverted to a ratio of 1:99 from (*R*)- to (*S*)-4-hydroxy-β-ionone. FuncLib-enabled active-site remodeling allowed us to generate a small yet highly diverse enzyme panel that significantly outperformed the wild type across multiple synthetic challenges. The best-performing variants, such as design 4 and design 11 (both 4 mutations), exhibit improvements that result from epistatic effects. MD simulations demonstrated that these mutations collectively reshape the active site, allowing for regio- and chemoselectivities that are difficult to achieve by single-point mutations. Herein, we demonstrate the potential of in silico-guided approaches to rapidly develop highly selective biocatalysts for synthetic applications.



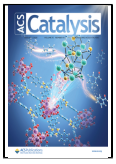
KEYWORDS: *computational chemistry, enzymes, FuncLib, oxyfunctionalization, Saccharomyces cerevisiae, terpenes, unspecific peroxygenase, yeast*

INTRODUCTION

Selective oxyfunctionalization ranks among the most challenging and desirable reactions in synthetic chemistry. These reactions are often critical for selective C−H activation and alkene epoxidation of complex organic molecules.[1−4] Since their discovery in 2004, fungal unspecific peroxygenases (UPOs) have attracted great interest for their ability to perform versatile oxyfunctionalization reactions on a broad scope of substrates.[5−7] These enzymes have several advantages relative to other versatile oxidases, such as P450s:[8,9] they are typically stable,[10] use prereduced hydrogen peroxide as a cosubstrate instead of molecular oxygen and expensive reductants such as NAD(P)H, and exhibit a broad substrate scope[11] and high turnover numbers (TON) up to 900,000.[12]

Despite their remarkable oxidative capabilities, unspecific peroxygenases (UPOs) also exhibit certain limitations. Many UPOs display only moderate regio- or enantioselectivity, which can restrict their application in fine chemical synthesis.[13,14] Additionally, their broad substrate scope—while advantageous in some contexts—can lead to undesired side reactions such as overoxidation or low product specificity.[6,15] These challenges

Received: May 8, 2025
Revised: June 20, 2025
Accepted: June 23, 2025
Published: July 14, 2025



https://doi.org/10.1021/acscatal.5c02412
ACS Catal. 2025, 15, 12741−12755

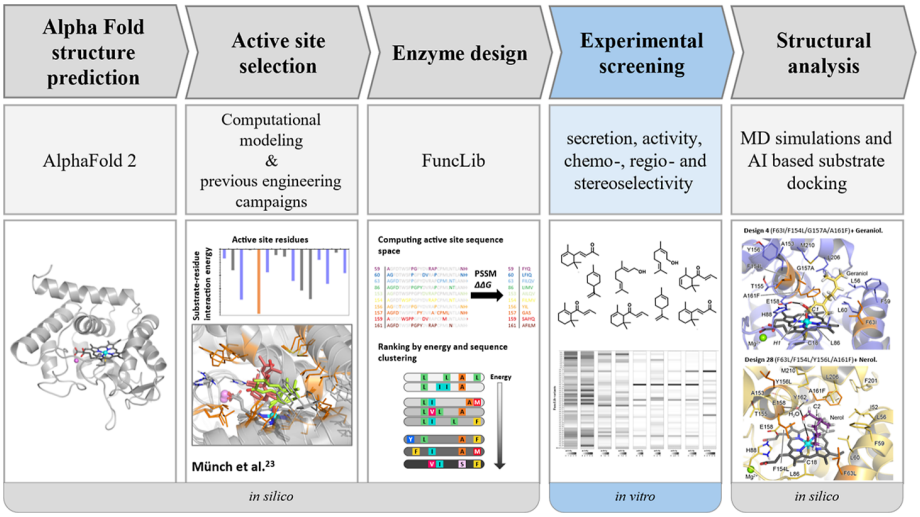


Figure 1. Overview of the different steps of the work protocol.

highlight the importance of engineering strategies to fine-tune UPO selectivity and expand its practical utility in biocatalysis.

UPOs can be engineered for overexpression in rapidly proliferating host organisms, such as yeast, through the implementation of protein[16] and signal-peptide[17] engineering and promoter-shuffling techniques.[18] Recent studies yielded enhanced variants characterized by increased activity,[19] augmented thermo-,[20] pH-,[21] and solvent-stabilities,[20] and large shifts in chemo-, regio-, and stereoselectivities.[22−27] These studies frequently employed directed evolution,[16,23,26,28] a prominent protein-engineering approach that emulates the engineering prowess of natural evolution through iterative rounds of random or semirational mutation and selection of variants that exhibit desirable properties.[29] But despite yielding exceptional results, directed evolution is labor-intensive and time-consuming. It is especially impractical when seeking enzymes that exhibit improvements across multiple substrates, as mutations that are favorable for one substrate rarely benefit others.[30] Furthermore, measuring oxyfunctionalization products is not amenable to chromogenic or fluorogenic measurement, limiting options for medium to high throughput screening, which is often essential to successful in vitro evolution campaigns. These limitations are currently addressed through small, "smart" libraries, using insights from molecular structures and mechanistic information, reducing the screening effort from thousands of variants to several hundred per round of evolution.[31−33] As a pertinent example, we recently performed an engineering campaign starting from the UPO from *Myceliophthora thermophila* (*Mth*UPO) resulting in two UPOs that selectively produced (*S*)- and (*R*)-4-hydroxy-*β*-ionone, with enantiomeric ratios of 97:3 (*R*) and 0.3:99.7 (*S*), respectively.[26] This study establishes a benchmark for rational engineering of a challenging UPO: we screened in total 2,500 variants in three rounds of evolution, 75% of which displayed lower activity than the parental variant.

Here, we ask whether recent developments in computational protein design and engineering[34] can address the limitations of rational engineering of enzyme active sites through a limited experimental effort.[35] Such protein-design methods can be applied to crystallographic structures and can be generalized to AI-based model structures.[26,36,37] For example, FuncLib is an automated method for designing diverse combinations of multipoint mutations within the active site of an enzyme.[37] This method uses phylogenetic analysis and Rosetta atomistic design calculations to generate variants that exhibit large diversity in active-site geometry and electrostatics without impairing the stability, foldability, and primary activity of the enzyme. The resulting designs may exhibit strong epistatic dependencies among the mutations[37−39] that are rarely observed in natural and lab-evolved variants; yet these designs exhibit diverse activities suggesting that FuncLib may uncover sequences and activities that are difficult for evolutionary processes to reach.

We focus on oxyfunctionalization of small terpenes which constitute the largest class of secondary plant metabolites.[40] Terpenes and their oxyfunctionalized derivatives, terpenoids, frequently showcase pharmacological activity[41] and find applications in the flavor and fragrance industries.[42,43] Chemo-, regio-, and stereoselective oxygenation presents a challenge for chemical catalysis, given that many terpenes feature multiple sites where C−H hydroxylation and C═C epoxidation are chemically and energetically almost indistinguishable. Enzymes are naturally stereoselective and can position substrates to promote reactions at desired positions that are not kinetically favored or chemically activated; thus, accurate control of the UPO active-site pocket may enable oxyfunctionalization reactions that overcome the dictates of intrinsic chemical reactivity of the substrate.

Starting from an AlphaFold2 model of *Mth*UPO, we generated 50 *Mth*UPO FuncLib designs. All proved to be functionally secreted from yeast, allowing us to compare their influence on various substrates and to find large activity improvements and substantial shifts in chemo-, regio-, and stereoselectivity through a limited experimental screening effort (Figure 1). AI-based modeling and molecular dynamics simulations of substrate−enzyme pairs provided insights into some of the dramatic specificity changes that may extend to other UPOs.

MATERIALS AND METHODS

Chemicals. See the Supporting Information.

Bacterial and Yeast Strains. For all cloning purposes and plasmid propagation, *E. coli* DH10B cells (ThermoFisher Scientific, Waltham, US) were utilized. All work regarding *S.*

https://doi.org/10.1021/acscatal.5c02412
ACS Catal. 2025, 15, 12741−12755

cerevisiae was performed utilizing the INVSc1 strain (ThermoFisher Scientific, Waltham, MA, US).

Oligonucleotides and Gene Parts. All oligonucleotides were purchased in the lowest purification grade "desalted" and minimal quantity at Eurofins Genomics (Ebersberg, DE). The genes of the *Mth*UPO FuncLib library were purchased as gene parts from Twist Bioscience (San Francisco, US).

AlphaFold2 Model. The AlphaFold2 model of *Mth*UPO was extracted from the AlphaFold Protein Structure Database (UniProt entry G2QID2). Source: https://alphafold.ebi.ac.uk/entry/G2QID2 (16.02.2023).

FuncLib Design. We used a C-terminally truncated (−18 residues) AlphaFold2 model of *Mth*UPO with its natural signal peptide automatically refined by the FuncLib algorithm (using Rosetta) as the protein structure. The loop at the original C-terminus entered the substrate channel and active side during the simulations, impeding residue variation wherefore we used the truncated model for further calculations. We subsequently defined positions C18 (axial ligand) as well as E158 and H88 (catalytic cascade) as essential amino acids which must not be subject to any variation. The number of mutations per design was set to two to four with at least two variations between each design. We did not apply an algorithm like PROSS[44] for introducing stabilizing mutations into the enzyme before running the FuncLib campaign, even though introducing several mutations in the active side may compromise enzyme stability. In our previous work with *Mth*UPO, we did not encounter stability issues after introducing mutations. Further prior work on *Aae*UPO variant PADA-I showed that an aggressive FuncLib campaign on unspecific peroxygenase can be tolerated without introducing further stabilizing mutations beforehand.[38] For diversification, positions were selected that enhanced activity, chemo-, regio-, and stereoselectivity of *Mth*UPO for different substrates in prior directed evolution campaigns.[22,25,26] Altering positions F59, L60, F63, A153, F154, Y156, G157, S159, A161, L206, and M210 showed substantial influence in previous studies, and we additionally selected position L86 due to its proximity to the active site. In total, we selected positions F59, L60, F63, L86, A153, F154, Y156, G157, S159, A161, L206, and M210 to be diversified. Charged amino acids (Arg, His, Lys, Glu, and Asp) were excluded from the allowed sequence space to keep the hydrophobic character of the active site. Thr was also excluded as including all mutations to the initial sequence space file, that were found in previous projects (e.g., F59Q), exceeded the maximum permitted number of variants. The highly polar amino acid Thr was not considered to be of most interest for the project compared to other more hydrophobic amino acids. Initial FuncLib runs showed steric overlap of mutations containing aromatic identities at position L86, mutations larger than Ala and Ser at position G157 might obstruct access to the heme, and the mutation A161F appeared to clash with the presumed heme site; therefore, such mutations were prohibited. In the future, the usage of AlphaFill or AlphaFold3 to include heme into the active site could prevent the introduction of mutations such as the aforementioned. Finally, the following sequence space was explored: Table S1.

Expression Plasmids. Expression plasmids for *Mth*UPO expression in *S. cerevisiae* were constructed as previously described—all parts are available at AddGene.[17] The gene parts of the FuncLib library were first cloned as individual level 0 standard modules into the universal level 0 acceptor plasmid (pAGM9121) and afterward released upon BsaI restriction digest. Golden Gate reactions were performed to combine the *Mth*UPO FuncLib genes with the *Sce*-α.galactosidase signal peptide and a TwinStrep-GFP11-Tag for purification and secretion detection in a level 1 expression plasmid (pAGT572).

Microtiter plate cultivation of *S. cerevisiae*, shake flask cultivation of *S. cerevisiae*, supernatant ultrafiltration and protein purification, and heme−CO complex measurements were performed as described previously.[17]

Colorimetric Screening Assays. Four colorimetric assays were performed: ABTS, DMP, NBD, and splitGFP assay. All assays were performed as described before.[45] For NBD, ABTS, and DMP, the absorbance difference between five minutes ($t1$) and 0 minutes ($t0$) was evaluated. The DMP assay was performed with enzyme secreted in a medium without additional heme to avoid a background reaction.

Bioconversion in a Microtiter Plate. All reactions with noncolorimetric substrates were initially performed in a microtiter plate. 100 μL of supernatant derived from enzyme expression and secretion in *S. cerevisiae* in a microtiter plate was transferred to a 96-deep-well plate (CR1496, EnzyScreen, Heemstede, NL). 400 μL of a reaction Mastermix was added to achieve final concentrations of 1 mM substrate, 1 mM H_2O_2, 5% acetone, and 100 mM Kpi (pH 7.0). Reactions were performed for 1 h at 30 °C under continuous shaking at 300 rpm. The extraction was accomplished through the addition of 500 μL of EtOAc (GC Ultra grade) containing 0.25 mM of an internal standard (Table S3) and further shaking for 30 min at 25 °C and 300 rpm. Microtiter plates were centrifuged to separate the phases (3000 rpm, 10 min) and 300 μL of the organic phase was transferred to a glass-coated microtiter plate utilizing the Platemaster (Gilson, Middelton, US) for subsequent GC−MS analysis.

Bioconversion in Single Vials. All bioconversions for regioselectivity determination were performed in triplicates in single vials with direct addition of H_2O_2. 250 nM enzyme supernatant, derived from enzyme expression and secretion in *S. cerevisiae* in a shake flask after ultrafiltration and concentration determination was transferred to a glass vial. A reaction Mastermix was added to achieve final concentrations of 1 mM substrate, 5% acetone, 1 mM H_2O_2, and 100 mM Kpi (pH 7.0) to give a total volume of 500 μL. Reactions were performed for 1 h at 30 °C under continuous shaking. The extraction was accomplished through the addition of 500 μL of EtOAc (GC Ultra grade) containing 0.25 mM of an internal standard (Table S3), and the organic phase was transferred to a new glass vial for subsequent GC−MS analysis.

Bioconversion in Single Vials with a Syringe Pump. All bioconversions for TON determination were performed in triplicates in single vials utilizing a syringe pump system. 250 nM enzyme supernatant, derived from enzyme expression and secretion in *S. cerevisiae* in a shake flask after ultrafiltration and concentration determination, was transferred to a glass vial. A reaction Mastermix was added to achieve final concentrations of 1 mM substrate, 5% acetone, and 100 mM Kpi (pH 7.0) to give a total volume of 400 μL. 100 μL of H_2O_2 (stock solution 5 mM, final concentration 1 mM) was added over the period of the reaction via a syringe pump. Reactions were performed for 1 h at 30 °C under continuous shaking. The extraction was accomplished through the addition of 500 μL of EtOAc (GC Ultra grade) containing 0.25 mM of an internal standard (Table S3), and the organic phase was transferred to a new glass vial for subsequent GC−MS analysis.

https://doi.org/10.1021/acscatal.5c02412
ACS Catal. 2025, 15, 12741−12755

GC−MS Analysis. All GC−MS measurements were performed on a Shimadzu GCMS-QP2010 Ultra (Shimadzu, Kyoto, JP) with helium as carrier gas. The detector voltage of the secondary electron multiplier was adjusted in relation to the tuning results with perfluorotributylamine. The GC−MS parameters were controlled with GCMS Real Time Analysis, and for data evaluation, GCMS Postrun Analysis (GCMSsolution Version 4.45, Shimadzu, Kyoto, JP) was used. The temperature programs utilized are listed in Table S3. Ionization was obtained by electron impact with a voltage of 70 V. Calibration and quantification were implemented in scan mode.

Nonchiral Gas Chromatography−Mass Spectrometry (GC−MS). Measurements were performed on an SH-Rxi-5Sil MS column (30 m × 0.25 mm, 0.25 μm film, Shimadzu, Kyoto, JP). One microliter of each sample was injected with a split ratio of 1:20 (inlet temperature 200 °C). The temperature of the ion source was 280 °C. All initial screening measurements were performed in single replicates.

Chiral Gas Chromatography−Mass Spectrometry (GC−MS). Measurements were performed on a Lipodex E column (25 × 0.25 mm, Macherey-Nagel, Düren, DE). One microliter of each sample was injected with a split ratio of 1:10 (inlet temperature 200 °C). The temperature of the ion source was 200 °C.

Quantum Mechanics (QM) Calculations. A truncated computational model was used to model the C−H activation and epoxidation of geraniol (1) and nerol (2) substrates. The truncated model [Fe═O(Por)(SCH_3)(substrate)] includes the active Fe-oxo species (Fe═O), the porphyrin pyrrole core (Por), a methyl thiolate group (-SCH_3) to mimic cysteine axial ligand, and 1 or 2 as substrate. Density Functional Theory (DFT) calculations were carried out using the Gaussian16 software package.[46] Geometry optimizations and frequency calculations were performed using the unrestricted hybrid (U)B3LYP[47−49] functional with an ultrafine integration grid[50] and including the CPCM polarizable conductor model (dichloromethane, ε = 8.9)[52,53] to have an estimation of the dielectric permittivity in the enzyme active site.[53] The 6-31G(d) basis set was used for all atoms but Fe, where the SDD basis set and related SDD pseudopotential were employed. The optimized geometries were verified as minima by vibrational frequency analysis, and transition state geometries have a single imaginary frequency consistent with the reaction coordinate. Enthalpies and entropies were calculated for 1 atm and 298.15 K. A correction to the harmonic oscillator approximation, as discussed by Truhlar and co-workers,[54,55] was also applied to the enthalpy calculations by raising all frequencies below 100 cm^{-1} to 100 cm^{-1} using the Goodvibes v.1.0.1 Python script.[56] Single-point energy calculations were performed using the functional (U)B3LYP with the Def2TZVP basis set on all atoms and within the CPCM polarizable conductor model (dichloromethane, ε = 8.9)[51,52] and an ultrafine integration grid.[50] Empirical Grimme D3 dispersion corrections with Becke−Johnson (GD3BJ) damping are also included in single-point calculations.[57] All structures have a total neutral charge and calculations were performed with doublet (d) or quartet (q) multiplicities consistent with the expected electronic states of the Fe. Figures of DFT structures were rendered using CYLview[58] and MolUP VMD extension[59] was used for output visualization.

Homology Model and Molecular Dynamics (MD) Simulations. The homology model for the *Mth*UPO structure (283 residues) obtained from our previous work has been used as the starting point.[22] Mutations were introduced using the Mutagenesis tool in PyMOL.[60] Molecular Dynamics (MD) simulations in explicit water were performed using the AMBER18 package.[61,62] Parameters for the geraniol (**1**) and nerol (**2**) substrates were generated within the antechamber[63] module in the AMBER18 package using the general AMBER force field (gaff2),[64] with partial charges set to fit the electrostatic potential generated at the B3LYP/6-31G(d) level by the RESP model.[65] The charges were calculated according to the Merz−Singh−Kollman scheme[66,67] using the Gaussian16 package. Parameters for the heme compound I (Cpd I) and the axial Cys were taken from ref 68. The protein was solvated in a pre-equilibrated cubic box with a 12 Å buffer of TIP3P[69] water molecules using the AMBER18 leap module, resulting in the addition of ∼12,500 solvent molecules. The systems were neutralized by the addition of explicit counterions (Na^+ and Cl^-). All subsequent calculations were done using the AMBER force field 14 Stony Brook (ff14SB).[70] A two-stage geometry optimization approach was performed. The first stage minimizes the positions of solvent molecules and ions imposing positional restraints on the solute by a harmonic potential with a force constant of 500 kcal·mol^{-1} $Å^{-2}$, and the second stage is an unrestrained minimization of all of the atoms in the simulation cell. The systems were gently heated using six 50 ps steps, incrementing the temperature by 50 K for each step (0−300 K) under constant-volume and periodic-boundary conditions. Water molecules were treated with the SHAKE algorithm such that the angle between the hydrogen atoms was kept fixed. Long-range electrostatic effects were modeled using the particle-mesh-Ewald method.[71] An 8 Å cutoff was applied to Lennard-Jones and electrostatic interactions. Harmonic restraints of 30 kcal·mol^{-1} were applied to the solute, and the Langevin equilibration scheme was used to control and equalize the temperature. The time step was kept at 1 fs during the heating stages, allowing potential inhomogeneities to self-adjust. Each system was then equilibrated for 2 ns with a 2 fs time step at a constant pressure of 1 atm and a temperature of 300 K without restraints. Once the systems were equilibrated in the *NPT* ensemble, production trajectories were then run under the *NVT* ensemble and periodic-boundary conditions. In particular, a total of 1500 ns from 3 independent replicas (500 ns each) were accumulated for each of the following systems: design 2 (F63L/A153I/F154I/G157A), design 4 (F63I/F154L/G157A/A161F), design 26 (F63I/F154I/G157A/A161L), and design 28 (F63L/F154L/Y156L/A161F). Trajectories were processed and analyzed using the CPPtraj[72] module from AmberTools utilities. VMD visualization software was used to visualize MD simulations.[73] Protein structures were rendered using PyMOL.[60]

Docking and Protocol Used for Substrate-Bound MD Simulations. Docking calculations were performed using AutoDock Vina.[74] The most populated clusters (based on backbone clustering analysis) obtained from MD simulations carried out in the absence of a substrate were used, and docking predictions were then utilized as starting points for substrate-bound MD simulations. The following systems were prepared: design 4 + geraniol, design 2 + geraniol, design 26 + nerol, and design 28 + nerol. The same protocol for MD simulations described above has been employed. Three replicas of 500 ns were carried out on each system without any external restraints on the substrate, thus accumulating a total of 1500 ns

https://doi.org/10.1021/acscatal.5c02412
ACS Catal. 2025, 15, 12741−12755

for each system. Trajectories were processed and analyzed using the CPPtraj28 module from the AmberTools utilities. VMD visualization software was used to visualize MD simulations.[73] Protein structures were rendered using PyMOL.[60]

AI-Based Docking for Design/Substrate Modeling. Models for each design/substrate pair were generated using Chai-1 (downloaded from https://github.com/chaidiscovery/chai-lab), an AlphaFold3-based and open-access predictor of biomolecular interactions. Each docking trajectory produces five docked models, and we launched several runs with different seeds to generate up to 50 models per design pair (Table S11).

We used the Simplified Molecular Input Line Entry System (SMILES) notation for heme-oxo species (compound I) for modeling:

C=CC1=C(C)C2=CC(C(C)=C3CCC([O-])=O)=
[N]4C3=CC5=[N]6C(C(C)=C5CCC([O-])=O)=
CC7=C(C=C)C(C)=C(C=C1[N-]28)[N-]7[Fe]
468=O

In addition, we used SMILES notations extracted from PubChem for each of the terpenes (Table S11).

Models that exhibited ≤4.5 Å between the oxidized carbon and compound I oxygen were defined as forming a near-attack conformation (NAC). Models that exhibited a ≤3.4 Å distance between any atoms on the ligand and the enzyme were considered as forming steric overlaps.

Preparative Work. *2,3-Epoxy Nerol.* Nerol (590 mg, 3.8 mmol) was dissolved in acetone (7.5 mL) and poured into a solution of potassium phosphate buffer (100 mM, 111 mL, pH 7.0) and 250 nM *Mth*UPO Var 28 (stock solution of 20.5 μM, 1.8 mL). Thirty millimolar hydrogen peroxide (stock solution 150 mM, 30 mL) was added via a syringe pump over the period of the reaction time. The solution (total: 150 mL) was stirred at 30 °C, overnight. Afterward, the mixture was extracted three times with ethyl acetate. The organic phase was washed with brine, dried with sodium sulfate, filtered, and concentrated under reduced pressure. The crude product was purified by column chromatography on silica gel using *n*-hexane/ethyl acetate (9/1 → 2/1) obtaining 156 mg (70%) of 2,3-epoxy nerol as a pale-yellow oil. GC−MS analysis showed impurities of silicone grease leading to a total purity of 85% which corresponds to a yield of 132 mg (55%) of 2,3-epoxy nerol.

^{1}H NMR (400 MHz, $CDCl_3$) δ: 5.07 (t, *J* 8 Hz, 1H), 3.37 (m, 1H), 3.65 (m, 1H), δ: 2.94 (q, *J* 4 Hz, 1H), 2.17−2.01 (m, 2H), 1.68−1.63 (m, 4H), 1.60 (s, 3H), 1.33 (s, 3H).

Isopiperitenol. (*S*)-(−)-Limonene (360 mg, 2.4 mmol) was dissolved in acetone (7.5 mL) and poured into a solution of potassium phosphate buffer (100 mM, 93 mL, pH 7.0) and 200 nM *Mth*UPO Var 28 (stock solution of 1.5 μM, 20 mL). Twenty millimolar hydrogen peroxide (stock solution 100 mM, 30 mL) was added via a syringe pump over the period of the reaction time. The solution (total: 150 mL) was stirred at 30 °C, overnight. Afterward, the mixture was extracted three times with ethyl acetate. The organic phase was washed with brine, dried with sodium sulfate, filtered, and concentrated under reduced pressure. The crude product was purified by column chromatography on silica gel using *n*-hexane/ethyl acetate (19/1 → 5/1) obtaining 3 mg (1%) of isopiperitenol as a pale-yellow oil. GC−MS analysis showed impurities of silicone grease leading to an even lower yield of isolated product.

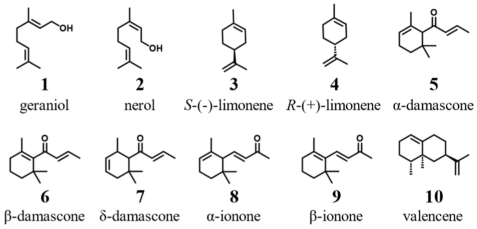
^{1}H NMR (400 MHz, $CDCl_3$) δ: 5.42 (s, 1H), 4.86 (d, *J* 20 Hz, 2H), 4.15−4.05 (m, 2H), 2.11 (s, 1H), 2.05 (m, 1H), 1.71 (s, 3H), 1.68 (s, 3H), 1.41 (d, *J* 8 Hz, 2H).

RESULTS

Modeling and Design Calculations. Because the molecular structure of *Mth*UPO has not been determined experimentally, we started from an AlphaFold2 model (UniProt entry: G2QID2) of the core enzymatic domain (amino acid positions 1−227).[75] Because AlphaFold2 does not model cofactors, we compared the model to the structure of an artificial UPO (artUPO, PDB entry: 7ZNV), a homologue from *Marasmius rotula* (*Mro*UPO).[76] Visual inspection verified that the side chain conformations of amino acids that are in direct contact with heme in other UPOs were aligned with those observed in experimentally determined UPOs.[26,76] The proximal axial ligand Cys18 and the catalytic dyad Glu158/His88 were kept fixed in their modeled conformations in all design calculations to maintain the core catalytic activity. Active-site positions were chosen for design to alter the active-site cavity based on our previous work and modeling (Table S1).[26] The designs were selected to contain 2−4 mutations each and to differ by at least two mutations from each other. 50 lowest-energy FuncLib designs (Table S2) were selected for experimental screening (see section *FuncLib design* in the SI for further details). Among the 50 designs, some positions exhibited no mutation (Leu206 and Met210) or only low sequence diversity, reflecting the high sequence conservation and energetic sensitivity of active-site positions, whereas others showed radical mutations with position Gly157 mutated in 88% of the designs.

Designs Generate Chemically Challenging Oxyfunctionalized Terpenoids. We selected ten small cyclic and noncyclic terpenes and terpenoids (Scheme 1) to assess the catalytic capabilities of the designed *Mth*UPO variants. Additionally, we chose three standard colorimetric substrates that serve as indicators for peroxygenase or peroxidase activity (ABTS (2,2′-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid), DMP (2,6-dimethoxyphenol), and NBD (5-nitro-1,3-benzodioxole) and performed a splitGFP assay to determine the protein secretion levels independently of activity.[45]

Scheme 1. Terpenoid Substrates Investigated in This Study



Among the substrates, the oxyfunctionalized products of geraniol and nerol are of outstanding interest, as partial oxidation of their terminal alcohol to aldehyde produces citral A (geranial, **33**) and citral B (neral, **31**), which are commonly used aroma compounds in perfumery due to their strong lemon (citrus) scents.[77] Moreover, citral is a key building block for the synthesis of vitamin A[78] and shows anti-inflammatory,

https://doi.org/10.1021/acscatal.5c02412
ACS Catal. 2025, 15, 12741−12755

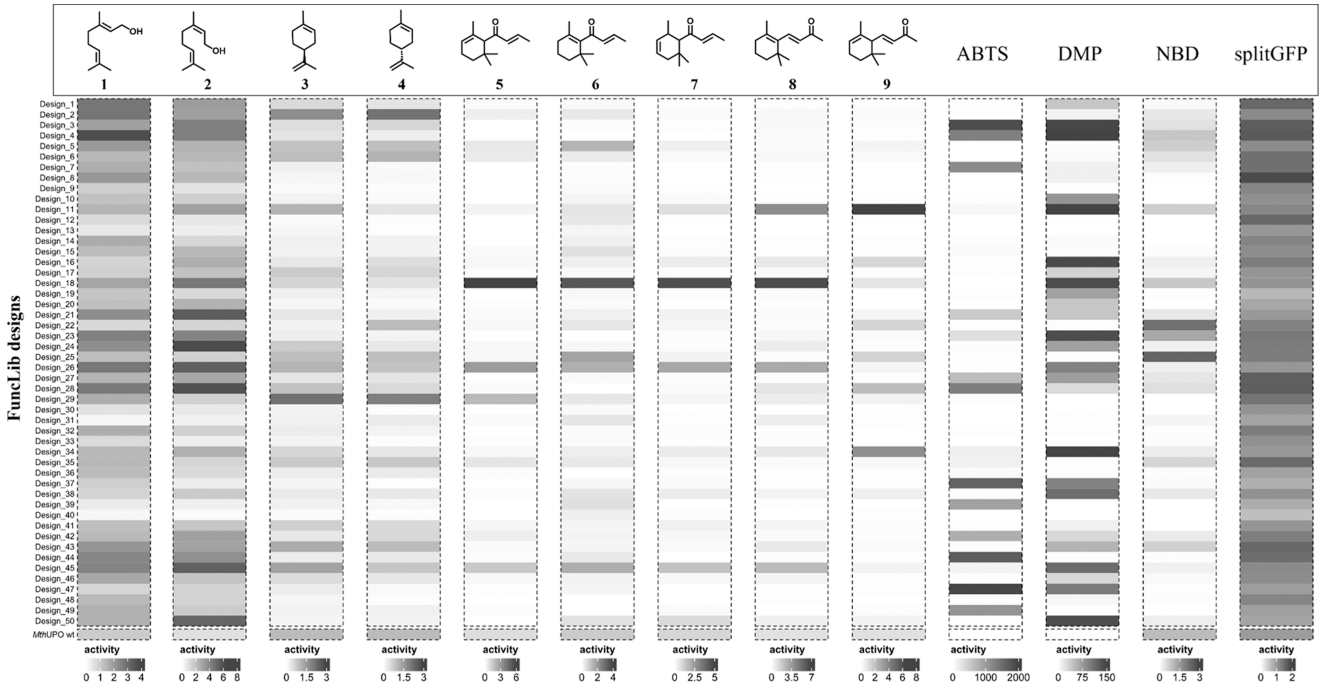


Figure 2. Total activity (sum of all products when multiple products were formed) relative to the activity of wild-type *Mth*UPO. The activity of wild-type *Mth*UPO is set to 1 separately for each substrate. All measurements were performed in a single replicate in microtiter plates. ABTS, DMP, and NBD were analyzed photometrically, while product formation for all other substrates was determined by GC−MS.

antitumor, and antibacterial activity.[79−81] Limonene and its oxyfunctionalized products are widely used in the fragrance[82,83] and polymer[84,85] industry and have important pharmacological effects.[86−89] Wild-type *Mth*UPO, however, yields a diverse array of products from limonene (Figure S3), including multiple oxyfunctionalizations, similar to the results observed with PaDa-I and artUPO.[76] Additionally, the volatility of limonene renders upscaling reactions for product isolation and identification challenging.

Wild-type *Mth*UPO is readily secreted in functional form from yeast, and remarkably, all 50 designs were successfully expressed and secreted. Secretion levels were between 0.6-fold and 1.9-fold of wild-type levels (Figure 2), indicating that active-site design using FuncLib preserves the expressibility of the parental enzyme. Furthermore, among the designed enzymes, we measured improvements in activity relative to the wild type for nine of the ten terpenes, with only valencene (**10**), the only substrate comprising two conjugated rings in our set, not showing detectable activity in the wild type or any of the designs. The greatest improvement was seen in designs 3 and 47 relative to ABTS (1,880-fold and 1,950-fold increase in absorbance related to product formation, respectively). Design 34 demonstrated a 200-fold activity improvement over DMP (Figure 2). Both design 3 and 47 performed excellently as well for DMP, but underperformed with all other substrates, indicating that those designs improved the peroxidase rather than peroxygenase activity. By contrast, designs 22 and 25 showed 2-fold improvement on the peroxygenase model substrate NBD.

All FuncLib designs showed at least baseline activity toward geraniol (**1**) and nerol (**2**), confirming that all designs were functionally secreted. The largest improvement toward geraniol was seen with design 4 exhibiting a 3.8-fold increase compared to the wild type. The highest activity for nerol (**2**) is seen by design 24 with 6.2-fold improvement. The highest overall activity toward all damascones and α-ionone (**8**) is seen in design 18 with 3.1- to 6.7-fold increases. The highest activity on β-ionone (**9**) was revealed by design 11 with a 7-fold increase relative to the wild type. For limonene, the activity gain reached 1.7- and 2.2-fold with designs 2 and 29, respectively, for (*R*)-(+)-limonene (**4**) and (*S*)-(−)-limonene (**3**). Thus, a variety of different designs exhibited excellent levels of oxyfunctionalization of different substrates.

Novel Products and Outstanding Improvements in Regioselectivity. Aside from the increase in activity, we also observed dramatic shifts in regioselectivity among the designs (Figure 3). Design 45 displays a notable shift in the formation of 3-hydroxy-β-damascone (**15**), which increased from 3% for wild-type *Mth*UPO to 46%. This regioselectivity shift is remarkable as the aliphatic C3-position is substantially less activated compared to the allylic C4-position of the main product 4-hydroxy-β-damascone (**14**). Thus, the designs can substantially change the profile of the resulting products toward ones that are less favored for chemical reactivity.

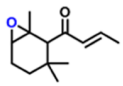
We also saw shifts in chemo- and regioselectivity for limonene. For instance, wild-type *Mth*UPO mainly generates the epoxide limonene oxide (**21**), whereas design 45 mainly produces (+)-carveol (**22**) (Figure 3B) from (*R*)-(+)-limonene, alongside a 5.6-fold increase in turnover number (TON). This change in chemoselectivity is noteworthy, as hydroxylation reactions are energetically less favored than epoxidation reactions. Moreover, this reaction provides access to a valuable fragrance and pharmacologically active compound.[86,87] The main product of (*S*)-(−)-limonene with design 45 was identified as isopiperitenol (**19**), which could only be found in trace amounts after conversion by wild-type *Mth*UPO. This activity is boosted 25-fold, reaching 310 TONs.

α- and β-Ionones are mainly converted by *Mth*UPO to 3-hydroxy-α-ionone (**26**) and 4-hydroxy-β-ionone (**28**), respectively.[26] With the FuncLib designs, the main products

https://doi.org/10.1021/acscatal.5c02412
ACS Catal. 2025, 15, 12741−12755

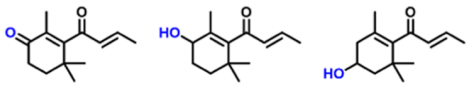
A

α-damascone



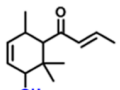
11 12

β-damascone



13 14 15

δ-damascone



16

Enzyme	4-epoxy-α-D 11	n.i.[a]	3-OH-α-D 12
Design 18	1 %	22 %	77 %
Design 26	1 %	7 %	91 %
Design 29	2 %	3 %	96 %
wt	7 %	5 %	88 %

[a] not identified, could be 2-hydroxy-α-damascone

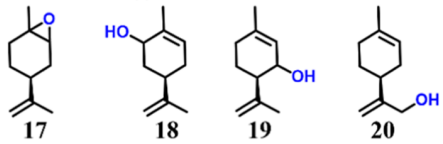
Enzyme	4-CO-β-D 13	4-OH-β-D 14	3-OH-β-D 15
Design 18	0 %	91 %	9 %
Design 25	28 %	72 %	0 %
Design 45	1 %	53 %	46 %
wt	12 %	85 %	3 %

Enzyme	n.i.[a]	2-OH-δ-D 16
Design 18	2 %	98 %
Design 26	18 %	82 %
Design 45	55 %	45 %
wt	6 %	94 %

[a] not identified

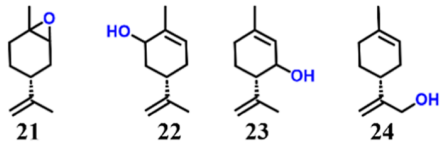
B

S-(-)-limonene

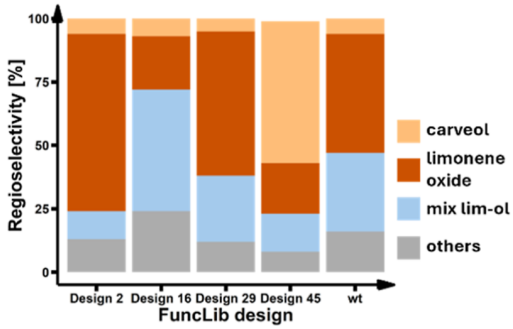


17 18 19 20

R-(+)-limonene



21 22 23 24

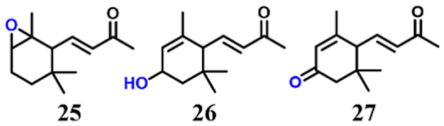


Enzyme	Substrate	-fold wt activity[a]	Regioselectivity: Limonene oxide	Carveol	Mix Lim-ol[b]	Others
Design 2	*R*-(+)-limonene	2.2	70 %	6 %	11 %	13 %
Design 16		0.5	21 %	7 %	48 %	24 %
Design 29		2.0	57 %	5 %	26 %	12 %
Design 45		0.8	20 %	56 %	15 %	8 %
wt		1.0	47 %	6 %	31 %	16 %
Design 2	*S*-(-)-limonene	1.7	71 %	7 %	9 %	12 %
Design 24		0.8	6 %	13 %	77 %	3 %
Design 29		2.2	52 %	19 %	6 %	23 %
Design 45		1.4	28 %	5 %	54 %	13 %
wt		1.0	49 %	7 %	7 %	37 %

[a] total activity compared to wt *Mth*UPO [b] mixture of isopiperitenol and limonene-10-ol, eluting both at the same time

C

α-ionone

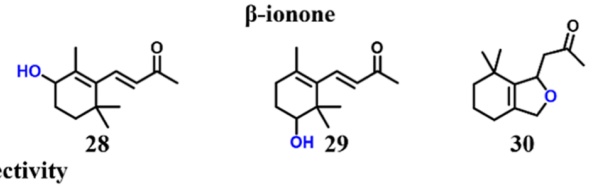


25 26 27

Enzyme	Regioselectivity 26
Design 18	98 %
Design 45	99 %
wt	91 %

D

β-ionone



28 29 30

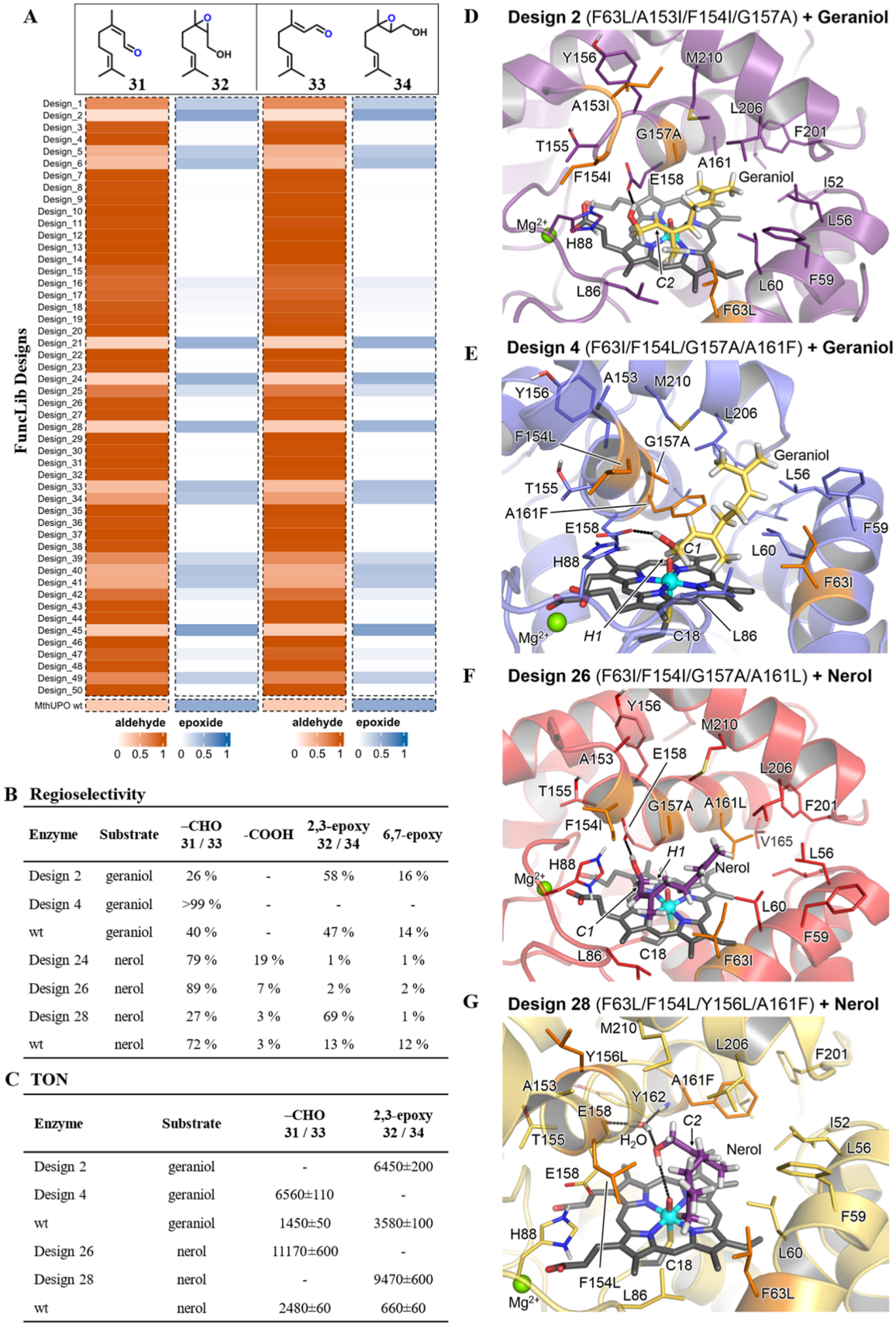
Regioselectivity

Enzyme	EME 30	4-OH-β-I 28	n.i.	2-OH-β-I 29
Design 11	0 %	87 %	12 %	0 %
Design 34	1 %	98 %	2 %	0 %
wt	4 %	92 %	1 %	3 %

Enzyme	-fold wt activity [a] 28
Design 11	7.1
Design 34	3.7
wt	1.0

Figure 3. Relative abundance of selected products after conversion with different FuncLib designs. All data are analyzed by GC−MS. (A) Conversion of *α*-, *β*-, and *δ*-damascone. (B) Conversion of (*S*)-limonene and (*R*)-limonene stacked bar chart displays the relative abundance of products after conversion of (*R*)-limonene leading to carveol (**22**) (orange), limonene oxide (**21**) (red), mix lim-ol (**23**/**24**) (blue), and others (gray). Isopiperitenol elutes on the GC−MS simultaneously with a second limonene alcohol (presumably limonene-10-ol), elution peaks are not separable, but mass spectra indicate the presence of two different moieties, the sum of both products is referred to as *mix lim-ol*. (C) Conversion of *α*-ionone. (D) Conversion of *β*-ionone reveals differing relative product abundance and increased enzyme activity compared to wild-type *Mth*UPO.

https://doi.org/10.1021/acscatal.5c02412
ACS Catal. 2025, 15, 12741−12755



B Regioselectivity

Enzyme	Substrate	–CHO 31 / 33	-COOH	2,3-epoxy 32 / 34	6,7-epoxy
Design 2	geraniol	26 %	-	58 %	16 %
Design 4	geraniol	>99 %	-	-	-
wt	geraniol	40 %	-	47 %	14 %
Design 24	nerol	79 %	19 %	1 %	1 %
Design 26	nerol	89 %	7 %	2 %	2 %
Design 28	nerol	27 %	3 %	69 %	1 %
wt	nerol	72 %	3 %	13 %	12 %

C TON

Enzyme	Substrate	–CHO 31 / 33	2,3-epoxy 32 / 34
Design 2	geraniol	-	6450±200
Design 4	geraniol	6560±110	-
wt	geraniol	1450±50	3580±100
Design 26	nerol	11170±600	-
Design 28	nerol	-	9470±600
wt	nerol	2480±60	660±60

Figure 4. Oxyfunctionalization of geraniol (**1**) and nerol (**2**) with FuncLib designs of *Mth*UPO. (A) Proportion of 2,3-epoxide (**32**/**34**) (blue) and aldehyde (**31**/**33**) (orange) formation in the overall reaction. Measurements were performed in single replicates with enzyme from the microtiter plate supernatant. (B) Regioselectivity of selected enzyme variants toward citral A (**33**)/citral B (**31**), neric acid and 2,3-epoxy (**32**/**34**), and 6,7-

https://doi.org/10.1021/acscatal.5c02412
ACS Catal. 2025, 15, 12741−12755

Figure 4. continued

epoxy formation measurements was performed in triplicates. (C) Turnover number of selected products; data are mean ± s.d. of measurements in triplicates. (D−G) Most representative structure of the binding mode of geraniol/nerol in different FuncLib designs as determined by clustering analysis.

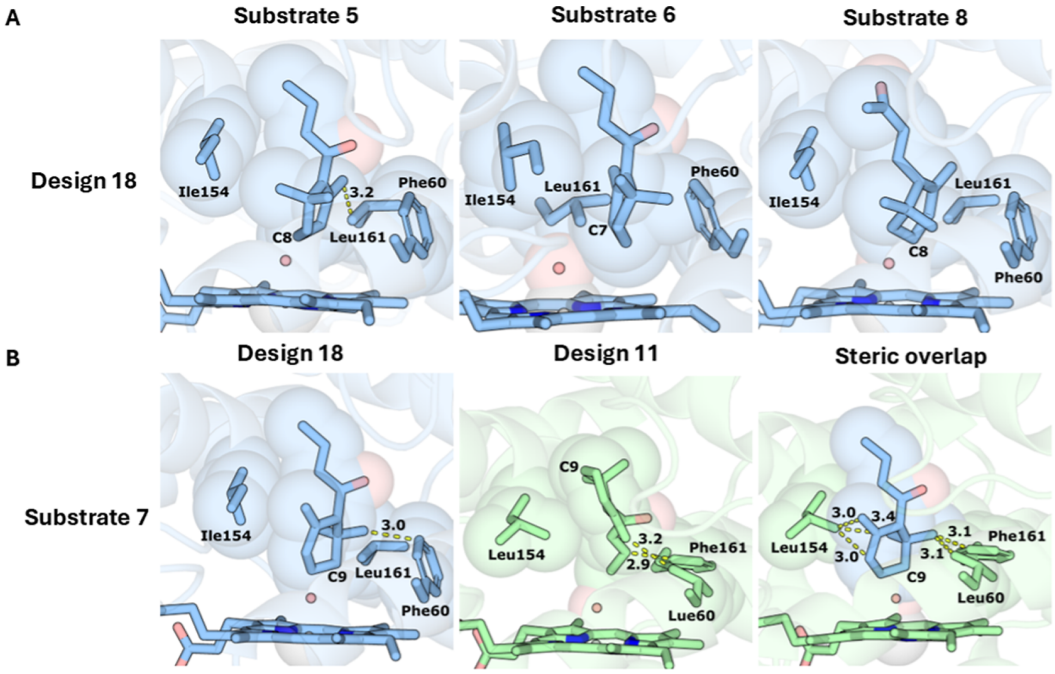


Figure 5. AI-based docking provides a molecular basis for the observed reactivities in design/substrate pairs. (A) Design 18, substrates 5, 6, and 8, and compound I were modeled using Chai-1, generating up to 50 models for each. For each pair, we selected a model that formed a near-attack conformation (NAC) defined as exhibiting distance <4.5 Å between the oxidized carbon and the compound I oxygen. (B) Substrate 7 exhibits much higher activity with design 18 than with design 11. In both cases, 10 models were generated with Chai-1. (left) The modeled interaction places the oxidized C9 atom of substrate 7 in a NAC. (center) Design 11 does not accommodate the NAC with this substrate, and (right) superimposing the substrate conformation modeled in design 18 in the active-site pocket of design 11 reveals significant steric overlap with designed amino acid residues that are close to the heme cofactor. Carbons in blue or green; compound I oxygen in red sphere; heme in sticks; steric overlap and interatomic distances (Å) indicated with dashed yellow lines.

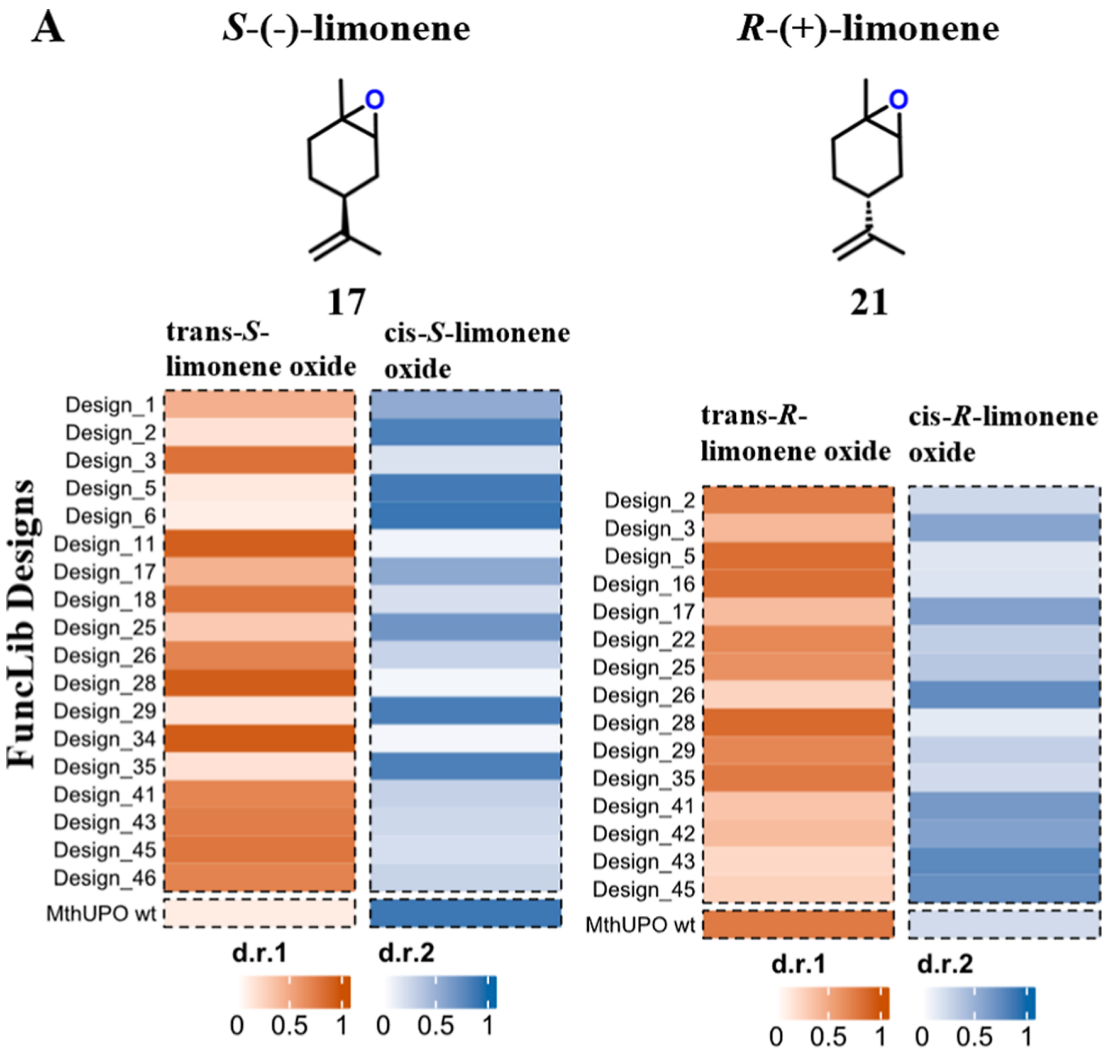
remained **26** and **28**, respectively, but the regioselectivity increased strongly. For the formation of **26**, it increased from 91% (wild-type *Mth*UPO) to 98% (design 18) and 99% (design 45), respectively. For **28**, the improvement was from 92% (wild-type *Mth*UPO) to 98% (design 34).

Overall, our experiments revealed large shifts in the main substrates, the formation of novel products, and an increase from satisfactory to outstanding regioselectivities, thus underscoring the efficacy of FuncLib in generating useful functional diversity. In addition, some FuncLib designs significantly enhanced reactivities and produced regioisomers that are energetically disfavored, demonstrating how changes to the active-site geometry and electrostatics can dramatically impact the outcome of the reactions (see below).

Large Chemoselectivity Shifts among FuncLib Designs. The FuncLib-enabled mutations lead to large shifts in chemoselectivity, as demonstrated by geraniol (**1**) and nerol (**2**). Design 4 displays excellent selectivity (>99%) for citral A (geranial, **33**, Figure 4) and a 4.5-fold improvement in TON compared to wild-type *Mth*UPO, which only showed 40% chemoselectivity for converting the terminal alcohol group. The most pronounced chemoselectivity shift toward citral B (neral, **31**) was achieved with design 26 with 89% relative abundance reaching 11,170 TONs, which corresponds to a 4.5-fold activity gain, but only a minor chemoselectivity shift compared to the 72% of wild-type *Mth*UPO. Further products are 2,3-epoxides, with a relative abundance of up to 58% for 2,3-epoxy geraniol (**34**) with design 2 and 69% for 2,3-epoxy nerol (**32**) with design 24. Some designs further exhibited overoxidation from nerol to neric acid, a reaction that was not observed with geraniol (Figure 4B). The selectively oxidized aldehydes citral A and citral B are of great importance for industrial applications, which renders design 4 with its outstanding chemoselectivity an exciting biocatalyst. Upscaling reactions with design 28 were successfully performed and yielded more than 150 mg of 2,3-epoxy nerol, which was subsequently used to determine TONs of both nerol and geraniol 2,3 oxides. Design 2 displayed a 1.8-fold increase in TON for **34** when using geraniol. Design 28 displayed a 14.3-fold boost for **32** compared to the wild-type *Mth*UPO enzyme, resulting in 9,470 TONs for product **32** (Figure 4C).

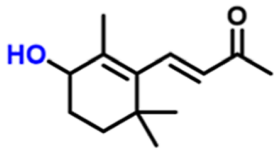
Molecular Basis of Chemoselectivity Changes. To shed light on the molecular basis for the chemoselectivity changes observed on the noncyclic terpene substrates, we conducted computational modeling based on Density Functional Theory (DFT) calculations using truncated models and molecular dynamics (MD) simulations. DFT calculations are used to elucidate intrinsic reactivity trends and provide

https://doi.org/10.1021/acscatal.5c02412
ACS Catal. 2025, 15, 12741−12755



B

β-ionone



28

4-OH-β-I

Enzyme	e.r. (*R*:*S*)
Design 11	1.1:98.9
Design 34	1.1:98.9
wt	76:24

Figure 6. Influence of FuncLib designs on stereoselectivity. (A) Shifts in the diastereomeric ratio of *trans*-limonene oxide (orange) and *cis*-limonene oxide (blue). Measurements were performed as single replicates with enzyme from the supernatant in the microtiter plate scale. (B) Shifts in the enantiomeric ratio of 4-hydroxy-*β*-ionone (**28**) giving access to *S*-hydroxy-*β*-ionone with design 11 and design 34. Measurements are performed as single replicates.

descriptions of optimal geometries for the competing C−H activation and epoxidation transition states, while MD simulations offer structural and molecular insights into key substrate−enzyme interactions that ultimately govern the selectivities observed across selected variants. Additionally, the AI-based structure predictor Chai-1 was used to explore catalytically relevant substrate binding modes and to assess its predictive power.

For both geraniol and nerol, DFT calculations in the absence of enzyme show that allylic hydroxylation on the terminal carbon atom (C1) and epoxidation at the 2,3 double bond are the energetically most favored oxyfunctionalizations, with the terminal allylic C−H activation (via Hydrogen Atom Transfer, HAT) slightly preferred (Figures S6 and S7 in the Supporting Information). This is consistent with the major products detected experimentally for these substrates with all variants.

https://doi.org/10.1021/acscatal.5c02412
ACS Catal. 2025, 15, 12741−12755

We chose to conduct Molecular Dynamics (MD) simulations (see Supporting Information for details) to analyze the large selectivity shifts seen in design 4 vs design 2 for geraniol oxidation and design 28 vs 26 for nerol. According to these simulations, geraniol establishes a persistent polar interaction with the Glu158 side chain of design 4, thus positioning both C1−H and C2−C3 bonds near the catalytically active oxo-ferryl cation radical complex species (compound I, see Figures 4E and S8 in the Supporting Information). The presence of the bulky Phe161 (mutated from wild-type Ala) side chain prevents the substrate from approaching the catalytic species in alternative near-attack conformations. Considering that terminal allylic hydroxylation is energetically slightly preferred, the citral A product is preferentially formed from this binding mode. On the other hand, design 2 led mainly to 2,3-epoxy geraniol with high activity and moderate chemoselectivity (1.8-fold increase and 58% selectivity). The models show that design 2 has a less dense active site, and the C2−C3 double bond of geraniol may approach the catalytic iron-oxo group. At the same time, C1−H bonds are not geometrically well-oriented for effective HAT (Figures 4D and S9 in the Supporting Information), leading to effective epoxidation by subtle substrate repositioning in the active site as compared to design 4. Therefore, MD simulations describe that active-site reshaping due to mutations repositions the substrate in the active-site cavity, which is finally responsible to switch the selectivity observed for design 4 vs design 2.

A similar scenario is found for nerol and designs 26 and 28. In design 26, nerol is preferentially bound by placing both the allylic terminal position and the C2−C3 double bond geometrically preorganized for oxyfunctionalization (Figures 4F and S10 in the Supporting Information). Nevertheless, mutations included in design 28, particularly Tyr156Leu, induce structural changes in the active-site cavity that favor the positioning of a water molecule that persistently interacts with the Glu158 and Tyr162 backbones. These structural changes, stemming from cooperation among several mutations, reposition nerol in the active site placing the C2−C3 double bond in a near-attack conformation for epoxidation by compound I while geometrically disfavoring terminal C−H activation (Figures 4G and S11 in the Supporting Information).

Recently developed AI-based structure predictors, such as AlphaFold3,[90] can accurately predict biomolecular interactions between proteins, cofactors, and ligands. We asked whether the open-access predictor Chai-1[91] could be used to shed light on the promiscuity and specificity of some of the designed enzymes modeling selected design/substrate pairs (Figure 5). We focused on design 18, which exhibits high activity toward substrates 5−8 (Figure 2), and generated ab initio models for the design, substrates, heme, and compound I oxygen. The models show that the substrates may all orient similarly, with the compound I oxygen atom placed close to the specific carbon atom that undergoes oxidation (Figures 5A and S14). In addition, we asked why design 18 oxidized substrate 7 much more efficiently than design 11. Modeling suggests that three mutated positions close to the compound I oxygen determine this specificity profile: in design 18, Phe60, Ile154, and Leu161 provide sufficient room for the ring to come close to the reactive oxygen, whereas Leu60, Leu154, and Phe161 in design 11 prevent this close approach (Figure 5B). Therefore, the Chai-1 structure predictive tool is able to capture and describe steric modifications in the active-site cavity due to mutations that impact substrate catalytically relevant binding poses.

FuncLib Designs Impact Stereoselectivity of Limonene Oxide and 4-Hydroxy-β-ionone. The major influence of the FuncLib library on stereoselectivity is shown by shifts in the diastereomeric ratio of 1,2-epoxy limonene (Figure 6A) ranging from 86:14 (design 28) for *trans*-(*R*)-limonene oxide to 23:77 (design 43) for *cis*-(*R*)-limonene oxide and 94:6 (design 28 and 34) for *trans*-(*S*)-limonene oxide to 10:90 (design 6) for *cis*-(*S*)-limonene oxide. We further see outstanding changes in the enantioselectivity of 4-hydroxy-β-ionone (**28**) from an enantiomeric ratio of 76:24 for wild-type *Mth*UPO to 1:99 for both design 11 and 34 producing (*S*)-4-hydroxy-β-ionone (Figure 6B). A lack of enantioselectivity of the FuncLib designs toward the (*R*)-4-hydroxyl-β-ionone product is not surprising as the key position for controlling *R*-selectivity in β-ionone hydroxylation (Leu206)[26] was not diversified in the FuncLib library (Table S2). These shifts confirm that FuncLib not only influences regio- and chemoselectivity but can further exert major influences on stereoselectivity in accordance with previous work on small aromatic substrates.[38]

DISCUSSION

In both industrial and academic contexts, the ability to produce a small set of highly active enzymes with diverse selectivities is extremely valuable to determining the feasibility of desired reactions. Previously, such diversity was generated, including in UPOs, using natural or engineered enzymes from a variety of sources. We demonstrated here that substantial diversity can also be provided by the FuncLib algorithm starting from a single enzyme that exhibits favorable stability and activity profiles.

All FuncLib designs we tested experimentally were functional and exhibited large and potentially useful changes in activity and selectivity profiles. This is remarkable, especially given that the designs were based on an AlphaFold model and that the UPO family is challenging for protein engineering due to difficulties in heterologous expression, a limited number of solved crystal structures, and its high glycosylation. For example, in a previous engineering effort in our lab, we screened 2,500 *Mth*UPO active-site variants, of which 75% displayed lower activity than the parental variant. By contrast, screening 50 designed variants in the current study led to substantial improvements in activity, chemo-, regio-, and stereoselectivity such as >99% chemo- and regioselectivity for the fragrant and pharmacological active citral A (starting from 40% for wild-type *Mth*UPO), more than 1,900-fold activity increase for ABTS, variants with at least doubled activity for all substrates, 98% regioselectivity and an inversion of enantioselectivity to an e.r. of 1:99 (starting from 76:24) for (*S*)-4-hydroxy-β-ionone, and increased regioselectivity for the formation of the energetically less favored 3-hydroxy-β-damascone (from 3% to 46%) as well as novel products and significant regioselectivity shifts with limonene. While achieving excellent selectivity may require further improvements in design methodology, FuncLib can identify substantially enhanced starting sequences for subsequent engineering campaigns, including using traditional engineering approaches, presenting a fascinating opportunity to accelerate enzyme engineering efforts toward new oxyfunctionalization selectivities.

https://doi.org/10.1021/acscatal.5c02412
ACS Catal. 2025, 15, 12741−12755

In addition to high predictive value, the computational chemistry modeling retrospectively suggested mechanistic explanations and molecular bases for the high reactivities observed in several design variants. These calculations are highly sensitive to structural details and crucial for describing critical active-site remodeling and substrate positioning, although they require a significant computational effort. We are encouraged that the high reliability of AlphaFold and FuncLib calculations enables such sophisticated simulations, yielding valuable insights into the structural rearrangements that enable efficient catalysis. Such insights may be useful in the design of efficient and selective enzymes in other UPOs targeting different substrates. The high level of conservation of UPO structures around the heme binding site suggests that mutations in these positions may impact substrate specificity in other members of this family. We also found that the recent AI-based biomolecular modeling tools[90] produce model structures that may help rationalize the observed specificity profiles. Our results underscore the complementary nature of MD simulations and AI-based structure prediction tools such as Chai-1. While MD provides fine-grained mechanistic insight by capturing the conformational flexibility and dynamic substrate−enzyme interactions that govern catalytic selectivity, Chai-1 offers a fast and accessible approach for identifying potential near-attack conformations. This makes it a valuable initial screening tool for exploring catalytically relevant binding modes. However, its limited resolution in accounting for subtle dynamic effects, such as those influencing enantioselectivity, highlights the need to integrate both approaches for a more comprehensive understanding of enzyme function. Importantly, our benchmarking of Chai-1 is intended as a preliminary test of its applicability in this context. A more systematic evaluation of its predictive power across broader data sets and different enzymatic systems will be a valuable direction for future work. Regardless, this is an encouraging sign that in the near future, we may be able to focus design calculations on specific desired substrates using a combination of atomistic design and AI-based substrate docking calculations and MD simulations.

Notwithstanding the high success rate of the FuncLib design library, for some substrates, we found a limited improvement in catalytic efficiency. We note that FuncLib usually designs thousands of active-site variants, but cost and time considerations preclude screening such large sets. We recently demonstrated an economical approach to design large combinatorial libraries for screening (htFuncLib).[92] We envision that such an approach, in combination with mechanistic insights from computational modeling, may enable finding many more high-efficiency and specificity designs, including against substrates for which we did not find high-efficiency designs.

Recent progress in AI-based modeling and evolution-guided atomistic design provides unanticipated opportunities to address protein-engineering challenges at the forefront of biocatalysis that have frustrated conventional in vitro evolution approaches. The new approaches could be further refined by considering additional mechanistic insights from atomistic models obtained from computational chemistry calculations. In this study, we demonstrated that this combination effectively samples the functional space of an enzyme active site toward diverse regio- and enantioselectivity outcomes while accounting for high activities. Using this approach, in principle, any natural enzyme can be engineered quickly and effectively to tune its activity profile for basic or applied needs. These findings underscore the potential of this strategy to contribute meaningfully to future developments in green and medicinal chemistry.

■ ASSOCIATED CONTENT

Data Availability Statement

Raw data for all activities found in the screening process and the sequences of the 50 FuncLib designs can be found on the following dataverse server: Muench, Judith, 2024, "UPO FuncLib Paper", https://doi.org/10.7910/DVN/ZPKEAI, Harvard Dataverse, DRAFT VERSION

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acscatal.5c02412.

Material and methods; gene and protein sequence; table containing mutations for all FuncLib designs; GC−MS temperature programs; utilized linker and tags; and structural characterization data for selected substrates and variants (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Sarel J. Fleishman − *Department of Biomolecular Sciences, Weizmann Institute of Science, Rehovot 7600001, Israel;* orcid.org/0000-0003-3177-7560; Email: sarel@weizmann.ac.il

Marc Garcia-Borràs − *Institut de Qui'mica Computacional i Catàlisi and Departament de Qui'mica, Universitat de Girona, Girona 17003 Catalonia, Spain;* orcid.org/0000-0001-9458-1114; Email: marc.garcia@udg.edu

Martin J. Weissenborn − *Institute of Chemistry, Martin Luther-University Halle-Wittenberg, Halle (Saale) 06120, Germany;* orcid.org/0000-0002-1200-4485; Email: martin.weissenborn@chemie.uni-halle.de

Authors

Judith Münch − *Institute of Chemistry, Martin Luther-University Halle-Wittenberg, Halle (Saale) 06120, Germany*

Jordi Soler − *Institut de Qui'mica Computacional i Catàlisi and Departament de Qui'mica, Universitat de Girona, Girona 17003 Catalonia, Spain*

Ofir Gildor-Cristal − *Department of Biomolecular Sciences, Weizmann Institute of Science, Rehovot 7600001, Israel*

Complete contact information is available at: https://pubs.acs.org/10.1021/acscatal.5c02412

Author Contributions

Conceptualization: J.M., M.J.W, S.J.F., and M.G.B.; methodology: J.M, S.J.F., and J.S.; validation: J.M., J.S., S.J.F., and M.G.B.; formal analysis: J.M., J.S., and OGC; investigation: J.M., OGC, and J.S.; resources: M.J.W, M.G.B., and S.J.F; data curation: J.M. and J.S.; writing—original draft: J.M and J.S.; writing—review and editing: J.M., J.S., S.J.F., M.G.B., and M.J.W; visualization: J.M., OGC, and J.S.; supervision: S.J.F., M.G.B., and M.J.W.; project administration: J.M.

Notes

During the preparation of this work, the authors used ChatGPT in order to improve readability and language of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

The authors declare the following competing financial interest(s): JM, SJF, and MJW are named inventors on a patent filing on designs reported here. SJF is a named inventor on patents related to the design methods used in this manuscript and advises companies on protein design.

■ ACKNOWLEDGMENTS

J.M. thanks the Friedrich-Naumann-Stiftung für die Freiheit for a PhD scholarship. M.J.W. and J.M. thank the Bundesministerium für Bildung und Forschung (Maβgeschneiderte Inhaltsstoffe 2, 031B0834A) and the German Research Foundation (DFG, project ID 43649874, TP A05, RTG 2670) for generous funding. M.G.-B. thanks the Spanish MICINN (Ministerio de Ciencia e Innovacio'n) for projects PID2022-141676NB-I00 and TED2021-130173B−C42 and the Ramo'n y Cajal RYC 2020-028628-I fellowship. Part of the computational resources used were funded by FEDER (Fondo Europeo de Desarrollo Regional) and Spanish MINECO (Ministerio de Economi'a, Comercio y Empresa) through projects CTQ2014-54306-P, CTQ2015-69363-P, and MPCUdG2016/096, as well as by the European Research Council (ERC) under the European Union's ERC-StG-2015 (Grant agreement No. 679001). J.S. thanks the Ramo'n Areces Foundation for a postdoctoral fellowship.

■ ABBREVIATIONS

HAT. hydrogen atom transfer; Mix-10-ol. mixture of limonene-10-ol and isopiperitenol; TON. turnover number; UPO. unspecific peroxygenase

■ REFERENCES

(1) White, M. C.; Zhao, J. Aliphatic C−H oxidations for late-stage functionalization. *J. Am. Chem. Soc.* **2018**, *140* (43), 13988−14009.
(2) Saint-Denis, T. G.; Zhu, R.-Y.; Chen, G.; Wu, Q.-F.; Yu, J.-Q. Enantioselective C (sp3)−H bond activation by chiral transition metal catalysts. *Science* **2018**, *359* (6377), No. eaao4798.
(3) Crabtree, R. H.; Lei, A. Introduction: CH activation. *Chem. Rev.* **2017**, *117* (13), 8481−8482.
(4) Chakrabarty, S.; Wang, Y.; Perkins, J. C.; Narayan, A. R. Scalable biocatalytic C−H oxyfunctionalization reactions. *Chem. Soc. Rev.* **2020**, *49* (22), 8137−8155.
(5) Beltra'n-Nogal, A.; Sa'nchez-Moreno, I.; Me'ndez-Sa'nchez, D.; Go'mez de Santos, P.; Hollmann, F.; Alcalde, M. Surfing the wave of oxyfunctionalization chemistry by engineering fungal unspecific peroxygenases. *Curr. Opin. Struct. Biol.* **2022**, *73*, 102342.
(6) Monterrey, D. T.; Me'nes-Rubio, A.; Keser, M.; Gonzalez-Perez, D.; Alcalde, M. Unspecific peroxygenases: The pot of gold at the end of the oxyfunctionalization rainbow? *Curr. Opin. Green Sustainable Chem.* **2023**, *41*, 100786.
(7) Kinner, A.; Rosenthal, K.; Lütz, S. Identification and expression of new unspecific peroxygenases−Recent advances, challenges and opportunities. *Front. Bioeng. Biotechnol.* **2021**, *9*, 705630.
(8) Guengerich, F. P.; Waterman, M. R.; Egli, M. Recent structural insights into cytochrome P450 function. *Trends Pharmacol. Sci.* **2016**, *37* (8), 625−640.
(9) Manikandan, P.; Nagini, S. Cytochrome P450 structure, function and clinical significance: a review. *Curr. Drug Targets* **2018**, *19* (1), 38−54.
(10) Hofrichter, M.; Kellner, H.; Pecyna, M. J.; Ullrich, R. Fungal unspecific peroxygenases: heme-thiolate proteins that combine peroxidase and cytochrome P450 properties. In *Monooxygenase, Peroxidase and Peroxygenase Properties and Mechanisms of Cytochrome P450*; Springer, 2015; pp 341−368.
(11) Hobisch, M.; Holtmann, D.; de Santos, P. G.; Alcalde, M.; Hollmann, F.; Kara, S. Recent developments in the use of peroxygenases−Exploring their high potential in selective oxyfunctionalisations. *Biotechnol. Adv.* **2021**, *51*, 107615.
(12) Hobisch, M.; De Santis, P.; Serban, S.; Basso, A.; Byström, E.; Kara, S. Peroxygenase-Driven Ethylbenzene Hydroxylation in a Rotating Bed Reactor. *Org. Process Res. Dev.* **2022**, *26* (9), 2761−2765.
(13) Kiebist, J.; Schmidtke, K. U.; Zimmermann, J.; Kellner, H.; Jehmlich, N.; Ullrich, R.; Zänder, D.; Hofrichter, M.; Scheibner, K. A peroxygenase from Chaetomium globosum catalyzes the selective oxygenation of testosterone. *ChemBioChem* **2017**, *18* (6), 563−569.
(14) Wang, Y.; Lan, D.; Durrani, R.; Hollmann, F. Peroxygenases en route to becoming dream catalysts. What are the opportunities and challenges? *Curr. Opin. Chem. Biol.* **2017**, *37*, 1−9.
(15) Hofrichter, M.; Ullrich, R. Heme-thiolate haloperoxidases: versatile biocatalysts with biotechnological and environmental significance. *Appl. Microbiol. Biotechnol.* **2006**, *71* (3), 276−288.
(16) Molina-Espeja, P.; Garcia-Ruiz, E.; Gonzalez-Perez, D.; Ullrich, R.; Hofrichter, M.; Alcalde, M. Directed evolution of unspecific peroxygenase from Agrocybe aegerita. *Appl. Environ. Microbiol.* **2014**, *80* (11), 3496−3507.
(17) Püllmann, P.; Knorrscheidt, A.; Münch, J.; Palme, P. R.; Hoehenwarter, W.; Marillonnet, S.; Alcalde, M.; Westermann, B.; Weissenborn, M. J. A modular two yeast species secretion system for the production and preparative application of unspecific peroxygenases. *Commun. Biol.* **2021**, *4* (1), 562.
(18) Püllmann, P.; Weissenborn, M. J. Improving the heterologous production of fungal peroxygenases through an episomal Pichia pastoris promoter and signal peptide shuffling system. *ACS Synth. Biol.* **2021**, *10*, 1360.
(19) Gomez de Santos, P.; Cañellas, M.; Tieves, F.; Younes, S. H.; Molina-Espeja, P.; Hofrichter, M.; Hollmann, F.; Guallar, V.; Alcalde, M. Selective synthesis of the human drug metabolite 5′-hydroxypropranolol by an evolved self-sufficient peroxygenase. *ACS Catal.* **2018**, *8* (6), 4789−4799.
(20) Martin-Diaz, J.; Paret, C.; Garci'a-Ruiz, E.; Molina-Espeja, P.; Alcalde, M. Shuffling the neutral drift of unspecific peroxygenase in Saccharomyces cerevisiae. *Appl. Environ. Microbiol.* **2018**, *84* (15), e00808−e00818.
(21) Gomez de Santos, P.; Hoang, M. D.; Kiebist, J.; Kellner, H.; Ullrich, R.; Scheibner, K.; Hofrichter, M.; Liers, C.; Alcalde, M. Functional expression of two unusual acidic peroxygenases from Candolleomyces aberdarensis in yeasts by adopting evolved secretion mutations. *Appl. Environ. Microbiol.* **2021**, *87* (19), No. e0087821.
(22) Knorrscheidt, A.; Soler, J.; Hünecke, N.; Püllmann, P.; Garcia-Borràs, M.; Weissenborn, M. J. Accessing chemo-and regioselective benzylic and aromatic oxidations by protein engineering of an unspecific peroxygenase. *ACS Catal.* **2021**, *11* (12), 7327−7338.
(23) Molina-Espeja, P.; Cañellas, M.; Plou, F. J.; Hofrichter, M.; Lucas, F.; Guallar, V.; Alcalde, M. Synthesis of 1-naphthol by a natural peroxygenase engineered by directed evolution. *ChemBioChem* **2016**, *17* (4), 341−349.
(24) Carro, J.; Gonza'lez-Benjumea, A.; Ferna'ndez-Fueyo, E.; Aranda, C.; Guallar, V.; Gutie'rrez, A.; Marti'nez, A. T. Modulating fatty acid epoxidation vs hydroxylation in a fungal peroxygenase. *ACS Catal.* **2019**, *9* (7), 6234−6242.
(25) Knorrscheidt, A.; Soler, J.; Hünecke, N.; Püllmann, P.; Garcia-Borràs, M.; Weissenborn, M. J. Simultaneous screening of multiple substrates with an unspecific peroxygenase enabled modified alkane and alkene oxyfunctionalisations. *Catal. Sci. Technol.* **2021**, *11* (18), 6058−6064.
(26) Münch, J.; Soler, J.; Hünecke, N.; Homann, D.; Garcia-Borràs, M.; Weissenborn, M. J. Computational-Aided Engineering of a Selective Unspecific Peroxygenase toward Enantiodivergent β-Ionone Hydroxylation. *ACS Catal.* **2023**, *13*, 8963−8972.
(27) Swoboda, A.; Pfeifenberger, L. J.; Duhovic', Z.; Bürgler, M.; Oroz-Guinea, I.; Bangert, K.; Weißensteiner, F.; Parigger, L.; Ebner, K.; Glieder, A.; et al. Enantioselective High-Throughput Assay Showcased for the Identification of (R)-as well as (S)-Selective

https://doi.org/10.1021/acscatal.5c02412
ACS Catal. 2025, 15, 12741−12755

Unspecific Peroxygenases for C− H Oxidation. *Angew. Chem., Int. Ed. Engl.* **2023**, *62* (46), No. e202312721.
(28) Martin-Diaz, J.; Molina-Espeja, P.; Hofrichter, M.; Hollmann, F.; Alcalde, M. Directed evolution of unspecific peroxygenase in organic solvents. *Biotechnol. Bioeng.* **2021**, *118* (8), 3002−3014.
(29) Arnold, F. H.; Volkov, A. A. Directed evolution of biocatalysts. *Curr. Opin. Chem. Biol.* **1999**, *3* (1), 54−59.
(30) Hammer, S. C.; Knight, A. M.; Arnold, F. H. Design and evolution of enzymes for non-natural chemistry. *Curr. Opin. Green Sustainable Chem.* **2017**, *7*, 23−30.
(31) Sun, Z.; Wikmark, Y.; Bäckvall, J. E.; Reetz, M. T. New concepts for increasing the efficiency in directed evolution of stereoselective enzymes. *Chem.—Eur. J.* **2016**, *22* (15), 5046−5054.
(32) Chica, R. A.; Doucet, N.; Pelletier, J. N. Semi-rational approaches to engineering enzyme activity: combining the benefits of directed evolution and rational design. *Curr. Opin. Biotechnol.* **2005**, *16* (4), 378−384.
(33) Wang, Y.; Xue, P.; Cao, M.; Yu, T.; Lane, S. T.; Zhao, H. Directed evolution: methodologies and applications. *Chem. Rev.* **2021**, *121* (20), 12384−12444.
(34) Listov, D.; Goverde, C. A.; Correia, B. E.; Fleishman, S. J. Opportunities and challenges in design and optimization of protein function. *Nat. Rev. Mol. Cell Biol.* **2024**, *25*, 639−653.
(35) Lipsh-Sokolik, R.; Fleishman, S. J. Addressing epistasis in the design of protein function. *Proc. Natl. Acad. Sci. U.S.A.* **2024**, *121* (34), No. e2314999121.
(36) Goldenzweig, A.; Goldsmith, M.; Hill, S. E.; Gertman, O.; Laurino, P.; Ashani, Y.; Dym, O.; Unger, T.; Albeck, S.; Prilusky, J.; et al. Automated structure-and sequence-based design of proteins for high bacterial expression and stability. *Mol. Cell* **2016**, *63* (2), 337−346.
(37) Khersonsky, O.; Lipsh, R.; Avizemer, Z.; Ashani, Y.; Goldsmith, M.; Leader, H.; Dym, O.; Rogotner, S.; Trudeau, D. L.; Prilusky, J.; et al. Automated design of efficient and functionally diverse enzyme repertoires. *Mol. Cell* **2018**, *72* (1), 178−186.e5.
(38) Gomez de Santos, P.; Mateljak, I.; Hoang, M. D.; Fleishman, S. J.; Hollmann, F.; Alcalde, M. Repertoire of computationally designed peroxygenases for enantiodivergent C−H oxyfunctionalization reactions. *J. Am. Chem. Soc.* **2023**, *145* (6), 3443−3453.
(39) Bengel, L. L.; Aberle, B.; Egler-Kemmerer, A. N.; Kienzle, S.; Hauer, B.; Hammer, S. C. Engineered Enzymes Enable Selective N-Alkylation of Pyrazoles With Simple Haloalkanes. *Angew. Chem., Int. Ed.* **2021**, *60* (10), 5554−5560.
(40) Christianson, D. W. Structural and chemical biology of terpenoid cyclases. *Chem. Rev.* **2017**, *117* (17), 11570−11648.
(41) Cox-Georgian, D.; Ramadoss, N.; Dona, C.; Basu, C. Therapeutic and medicinal uses of terpenes. *Med. Plants* **2019**, 333−359.
(42) Tetali, S. D. Terpenes and isoprenoids: a wealth of compounds for global use. *Planta* **2019**, *249*, 1−8.
(43) Breitmaier, E. *Terpenes: Flavors, Fragrances, Pharmaca, Pheromones*; John Wiley & Sons, 2006.
(44) Weinstein, J. J.; Goldenzweig, A.; Hoch, S.; Fleishman, S. J. PROSS 2: a new server for the design of stable and highly expressed protein variants. *Bioinformatics* **2021**, *37* (1), 123−125.
(45) Dietz, N.; Wan, L.; Münch, J.; Weissenborn, M. J. Secretion and directed evolution of unspecific peroxygenases in S. cerevisiae. *Methods Enzymol.* **2023**, *693*, 267−306.
(46) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; et al. *Gaussian 16*, Revision C.01; Gaussian, Inc.: Wallingford, CT, 2016.
(47) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37* (2), 785.
(48) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A:At., Mol., Opt. Phys.* **1988**, *38* (6), 3098.
(49) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98* (7), 5648−5652.
(50) Bootsma, A. N.; Wheeler, S. Popular integration grids can result in large errors in DFT-computed free energies. *ChemRxiv* **2019**.
(51) Barone, V.; Cossi, M. Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model. *J. Phys. Chem. A* **1998**, *102* (11), 1995−2001.
(52) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. *J. Comput. Chem.* **2003**, *24* (6), 669−681.
(53) Schutz, C. N.; Warshel, A. What are the dielectric "constants" of proteins and how to validate electrostatic models? *Proteins* **2001**, *44* (4), 400−417.
(54) Ribeiro, R. F.; Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Use of solution-phase vibrational frequencies in continuum models for the free energy of solvation. *J. Phys. Chem. B* **2011**, *115* (49), 14556−14562.
(55) Zhao, Y.; Truhlar, D. G. Computational characterization and modeling of buckyball tweezers: density functional study of concave−convex $\pi\cdots\pi$ interactions. *Phys. Chem. Chem. Phys.* **2008**, *10* (19), 2813−2818.
(56) Funes-Ardoiz, I.; Paton, R. *GoodVibes: GoodVibes v1.0.2*; Zenodo, 2016.
(57) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32* (7), 1456−1465.
(58) Legault, C. *CYLview, 1.0 B*; Universite' de Sherbrooke Canada, 2009.
(59) Fernandes, S. H.; Ramos, M. J.; Cerqueira, M. F. S. A. *molUP: A VMD Plugin to Handle QM and ONIOM Calculations Using the Gaussian Software*; Wiley Online Library, 2018.
(60) Likova, E.; Petkov, P.; Ilieva, N.; Litov, L. *The PyMOL Molecular Graphics System*, version 2.0; Schrödinger, LLC: New York, 2015.
(61) *AMBER 2018*; University of California: San Fransisco, 2018.
(62) Cruzeiro, V. W. D.; Amaral, M. S.; Roitberg, A. E. Redox potential replica exchange molecular dynamics at constant pH in AMBER: Implementation and validation. *J. Chem. Phys.* **2018**, *149* (7), 072338.
(63) Li, P.; Merz, K. M., Jr. MCPB.py: a python based metal center parameter builder. *J. Chem. Inf. Model.* **2016**, *56*, 599−604.
(64) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25* (9), 1157−1174.
(65) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **1993**, *97* (40), 10269−10280.
(66) Besler, B. H.; Merz, K. M.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Comput. Chem.* **1990**, *11* (4), 413−439.
(67) Singh, U.; Kollman, P. An approach to computing electro static charges for molecules. *Comput. Chem.* **1984**, *5*, 129−145.
(68) Shahrokh, K.; Orendt, A.; Yost, G. S.; Cheatham III, T. E. Quantum mechanically derived AMBER-compatible heme parameters for various states of the cytochrome P450 catalytic cycle. *J. Comput. Chem.* **2012**, *33* (2), 119−133.
(69) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79* (2), 926−935.
(70) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11* (8), 3696−3713.
(71) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N·log (N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98* (12), 10089−10092.
(72) Roe, D. R.; Cheatham, T. E., III PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* **2013**, *9* (7), 3084−3095.

https://doi.org/10.1021/acscatal.5c02412
ACS Catal. 2025, 15, 12741−12755

(73) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **1996**, *14* (1), 33−38.
(74) Trott, O.; Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31* (2), 455−461.
(75) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583−589.
(76) Melling, B.; Mielke, T.; Whitwood, A. C.; O'Riordan, T. J. C.; Mulholland, N.; Cartwright, J.; Unsworth, W. P.; Grogan, G. Complementary specificity of unspecific peroxygenases enables access to diverse products from terpene oxygenation. *Chem Catal.* **2024**, *4*, 100889.
(77) Southwell, I. Backhousia citriodora F. Muell.(Lemon Myrtle), an unrivalled source of citral. *Foods* **2021**, *10* (7), 1596.
(78) Parker, G. L.; Smith, L. K.; Baxendale, I. R. Development of the industrial synthesis of vitamin A. *Tetrahedron* **2016**, *72* (13), 1645−1652.
(79) Sharma, S.; Habib, S.; Sahu, D.; Gupta, J. Chemical properties and therapeutic potential of citral, a monoterpene isolated from lemongrass. *Med. Chem.* **2020**, *17* (1), 2−12.
(80) Bailly, C. Targets and pathways involved in the antitumor activity of citral and its stereo-isomers. *Eur. J. Pharmacol.* **2020**, *871*, 172945.
(81) Zielin'ska, A.; Martins-Gomes, C.; Ferreira, N. R.; Silva, A. M.; Nowak, I.; Souto, E. B. Anti-inflammatory and anti-cancer activity of citral: Optimization of citral-loaded solid lipid nanoparticles (SLN) using experimental factorial design and LUMiSizer®. *Int. J. Pharm.* **2018**, *553* (1−2), 428−440.
(82) Santos, R. A. M. d.; Souza, F. d. O.; Pilau, E. J.; Porto, C.; Gonçalves, J. E.; Oliveira, A. J. B. d.; Gonçalves, R. A. C. Biotransformation of (+)-carvone and (−)-carvone using human skin fungi: A green method of obtaining fragrances and flavours. *Biocatal. Biotransform.* **2018**, *36* (5), 396−400.
(83) Kashid, S.; Joshi, K.; More, S.; Shinde, A.; Nene, S. Enhanced Productivity of Fragrance Compounds: Biotransformation of d-limonene Using Whole Cell Immobilization of Pseudomonas putida and Rhodococcus erythropolis. *J. Inst. Eng. (India): E* **2023**, *104* (1), 83−93.
(84) Parrino, F.; Fidalgo, A.; Palmisano, L.; Ilharco, L. M.; Pagliaro, M.; Ciriminna, R. Polymers of limonene oxide and carbon dioxide: polycarbonates of the solar economy. *ACS Omega* **2018**, *3* (5), 4884−4890.
(85) Carrodeguas, L. P.; Chen, T. T.; Gregory, G. L.; Sulley, G. S.; Williams, C. K. High elasticity, chemically recyclable, thermoplastics from bio-based monomers: carbon dioxide, limonene oxide and ε-decalactone. *Green Chem.* **2020**, *22* (23), 8298−8307.
(86) Ahmed, M. S.; Khan, A.-u.; Kury, L. T. A.; Shah, F. A. Computational and pharmacological evaluation of carveol for antidiabetic potential. *Front. Pharmacol.* **2020**, *11*, 919.
(87) Alvi, A. M.; Al Kury, L. T.; Alattar, A.; Ullah, I.; Muhammad, A. J.; Alshaman, R.; Shah, F. A.; Khan, A. U.; Feng, J.; Li, S. Carveol attenuates seizure severity and neuroinflammation in pentylenetetrazole-kindled epileptic rats by regulating the Nrf2 signaling pathway. *Oxid. Med. Cell. Longev.* **2021**, *2021*, 9966663.
(88) Chen, T. C.; Da Fonseca, C. O.; Schönthal, A. H. Preclinical development and clinical use of perillyl alcohol for chemoprevention and cancer therapy. *Am. J. Cancer Res.* **2015**, *5* (5), 1580.
(89) Chen, T. C.; Da Fonseca, C. O.; Schönthal, A. H. Perillyl alcohol and its drug-conjugated derivatives as potential novel methods of treating brain metastases. *Int. J. Mol. Sci.* **2016**, *17* (9), 1463.
(90) Abramson, J.; Adler, J.; Dunger, J.; Evans, R.; Green, T.; Pritzel, A.; Ronneberger, O.; Willmore, L.; Ballard, A. J.; Bambrick, J.; et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **2024**, *630*, 493−500.
(91) Discovery, C.; Boitreaud, J.; Dent, J.; McPartlon, M.; Meier, J.; Reis, V.; Rogozhnikov, A.; Wu, K. Chai-1: Decoding the molecular interactions of life. *bioRxiv* **2024**, 10.10.615955.
(92) Weinstein, J. Y.; Marti'-Go'mez, C.; Lipsh-Sokolik, R.; Hoch, S. Y.; Liebermann, D.; Nevo, R.; Weissman, H.; Petrovich-Kopitman, E.; Margulies, D.; Ivankov, D.; et al. Designed active-site library reveals thousands of functional GFP variants. *Nat. Commun.* **2023**, *14* (1), 2890.

https://doi.org/10.1021/acscatal.5c02412
ACS Catal. 2025, 15, 12741−12755

8 General Discussion and Perspectives

This section briefly reviews the main findings of this dissertation. While the individual chapters already include detailed discussions and conclusions related to each study, the focus here is to integrate the results, highlight common themes, and place them in the context of current knowledge on the engineering of unspecific peroxygenases. Furthermore, this section outlines perspectives on future research and anticipated developments in the field.

8.1 Motivation for Engineering UPOs

The body of work presented in this dissertation underscores the remarkable biocatalytic potential of unspecific peroxygenases (UPOs). These fungal enzymes operate using only hydrogen peroxide as a co-substrate, eliminating the need for costly cofactors or redox partners, and display extraordinary versatility—catalyzing C(sp^3)-, C(sp^2)-, and aromatic hydroxylations, epoxidations, halogenations, and heteroatom oxidations. Their sequence diversity, coupled with robust heterologous expression in yeast (*P. pastoris*, *S. cerevisiae*), *E. coli* and industrial mold strains, along with several successful engineering campaigns, highlights their adaptability for tailored applications. Some UPO variants have already reached impressive turnover numbers—up to several hundred thousands—demonstrating high catalytic throughput and broad substrate tolerance.[230,231] Moreover, UPOs often display fine chemo-, regio-, and enantioselectivity, making them particularly attractive for selective synthesis and late-stage functionalization in pharmaceutical and chemical contexts.[19,72,73]

More broadly, enzymes have become an increasingly important component of industrial biotechnology. Compared with traditional chemical catalysts, they can in many cases offer advantages such as higher reaction rates, specificity, product yields, or product quality, while also helping to reduce energy consumption, hazardous reagents, and toxic by-products. Their selectivity can minimize downstream separation steps, which is particularly beneficial in processes where isomeric purity is required.[232,233] Reflecting this relevance, the global enzyme market was valued at USD10.9-13.97billion in 2023 and is projected to reach USD21-24billion by the early 2030s.[234–236] Demand arises from diverse sectors, including food, agriculture, biofuels, pharmaceuticals, textiles, leather, cosmetics, and waste management.[237]

Industrial enzyme production relies predominantly on microorganisms: around 60% of enzymes originate from fungi, 24% from bacteria, 4% from yeast, and 10% from plants or animals.[238] Microbial expression systems are attractive because they enable high yields, reproducibility, and cost-effective cultivation, while also being amenable to optimization and strain engineering. Yeast systems such as *P. pastoris* and *S. cerevisiae* are of particular interest since they allow extracellular secretion of target proteins. Although secretion is not advantageous in every case, it can considerably simplify purification and process development. For UPOs—fungal enzymes that can be

successfully expressed and secreted in yeast—this property represents a significant step toward scalability.[239]

However, establishing a viable industrial biocatalytic process requires more than enzyme discovery or engineering alone. It is a multidimensional endeavor that encompasses efficient downstream processing and immobilization strategies, strain improvement, media optimization, bioreactor design, and the decision between whole-cell or extracellular production systems.[163] Improvements in enzyme activity and production titers are therefore just as critical as tailoring catalytic properties, and both must be aligned with scalable process design.

At the same time, wild type enzymes rarely fulfill the stringent demands of industry. Robust biocatalysis requires not only high activity and selectivity, but also long half-life, tolerance to non-natural substrates, resistance to extremes of pH, temperature, and solvent conditions, as well as minimized substrate and product inhibition.[240] Achieving these parameters typically necessitates targeted protein engineering. Indeed, nearly 90% of all commercial enzymes are engineered to meet such performance demands,[233] underscoring the central role of engineering in translating promising enzymes such as UPOs into practical industrial biocatalysts.

8.2 Expression Systems and Upscaling Strategies

Before addressing specific engineering strategies, it is important to consider factors that, while not directly related to protein engineering, play a crucial role in the overall success of an enzyme development campaign. In this section, we focus on two such aspects: the choice and optimization of heterologous expression systems used for producing unspecific peroxygenases, and the strategies employed for upscaling and performing enzymatic conversions at preparative scale. These topics frame the practical and biotechnological context in which the engineering efforts took place and are essential for ensuring that optimized enzymes can be produced efficiently and applied effectively.

8.2.1 Yeast Expression Systems: *P. pastoris* and *S. cerevisiae*

One of the major challenges in working with UPOs has been achieving efficient enzyme production, including correct protein folding, heme incorporation, post-translational modifications, and suitable glycosylation.[241] These processes are strongly influenced by the choice of expression host, and various heterologous systems have been explored to overcome these limitations. Heterologous expression of UPOs has been studied in several hosts, including *Aspergillus niger*,[242] *Aspergillus oryzae*,[35] *P. pastoris*,[37–39,243] *S. cerevisiae*,[38,62] and *E. coli*,[43] with highly variable expression efficiencies. The work in this dissertations utilizes a previously established heterologous expression systems for the secretion of UPOs in yeast hosts. Modular expression platforms had been developed for both *P. pastoris* and *S. cerevisiae*, incorporating promoter and signal-peptide shuf-

fling libraries that enabled robust initial secretion of diverse UPOs.[38,39] Both shuffling methods proofed largely important for either initial secretion of some UPOs or improvment of the enzyme secretion levels. Secretion levels of up to 22.4 mg/L could be reached for *Mth*UPO with the P_{HpFMD} promotor and the *Gga*-lysozym signal-peptide in *P. pastoris* and 5 mg/L for *Mth*UPO with the *Sce*-α-galactosidase signal-peptide in *S. cerevisiae*.[38,39] Also in most other cases the secretion levels were higher in *P. pastoris* compared to *S. cerevisiae*. Both *P. pastoris* and *S. cerevisiae* were employed in the studies of this dissertation, integrated with additional strategies—such as structure-based mutagenesis—to address secretion challenges and formed the starting point for the enzyme engineering campaigns.

P. pastoris is particularly attractive for industrial-scale protein production due to its high secretion capacity, availability of strong methanol-inducible promoters, and its Crabtree-negative metabolism, which minimizes ethanol accumulation and enhances respiratory growth.[244] The *P. pastoris* expression system utilized in this work included an episomal plasmid for high-throughput mutagenesis and screening, and an integrative plasmid designed for stable, high-yield expression suitable for batch fermentation. This dual system was employed to improve secretion and stability of novel UPOs, which initially failed to express or secrete adequately using signal-peptide shuffling alone (see Chapter III). To overcome these limitations, we combined the signal-peptid shuffling approaches with structure-guided stabilizing mutations predicted by the PROSS algorithm. As a result, nine out of ten engineered variants showed functional expression in *P. pastoris*, compared to only one of the corresponding wild type enzymes. Notably, this included three UPOs from oomycetes—representing the first experimentally characterized non-fungal UPOs—demonstrating the robustness of the approach even for phylogenetically distant enzymes.

The same constructs that yielded optimal expression in *P. pastoris* were subsequently expressed in *S. cerevisiae* to assess the influence of host-dependent factors on enzyme stability. Apparent transition temperatures (T_m) were determined and revealed substantial variability between hosts, with differences exceeding 15 °C in some cases. These differences likely stem from variations in glycosylation patterns, but may also be influenced by the use of different C-terminal tags in the constructs. In several instances, the same enzyme exhibited a higher T_m in one host over the other, without a consistent trend favoring either organism. Thus, while no generalizable conclusion can be drawn regarding host-dependent stability, the magnitude of the observed differences underscores the importance of carefully selecting the expression host based on the intended application.

Despite the potential of the *P. pastoris* system, several challenges were encountered. Notably, enzyme secretion levels in microtiter plate screenings using episomal plasmids were highly variable, complicating the identification of improved variants. This variability was partially mitigated by supplementing the culture medium with 0.5% (w/v) casamino acids, which stabilized secretion levels. However, the requirement to subsequently transfer selected variants into integrative plasmids, followed by linearization and transformation into production strains, added significant

experimental overhead. These factors ultimately prompted a shift toward the use of *S. cerevisiae* for other directed evolution campaigns.

S. cerevisiae is widely adopted in protein engineering of eukaryotic enzymes[62,176,245–248] due to its ability to perform post-translational modifications, its high transformation efficiency and the large availability of genome editing tools due to its well-annotated genome.[244] In this work, it enabled consistent UPO secretion and was used for engineering activity and selectivity toward various terpene substrates, as described in Chapters IV and V. Its stable secretion facilitated more reliable screening and characterization workflows compared to *P. pastoris*.

8.2.2 Upscaling and Biocatalytic Conversions at Preparative Scale

The outstanding oxyfunctionalization capabilities of unspecific peroxygenases (UPOs) underscore their potential as versatile biocatalysts. However, considering the significant effort required for their discovery, heterologous expression, and engineering, it is crucial to demonstrate their applicability beyond analytical-scale assays. Preparative-scale biotransformations serve as an essential benchmark to assess catalytic performance under conditions more reflective of synthetic and industrial environments. They enable the use of substrates with commercial or pharmacological relevance, offering insights into process scalability and economic feasibility. Moreover, such reactions are often indispensable for isolating sufficient product quantities for structural elucidation (e.g., NMR spectroscopy or optical rotation) or for preparing authentic standards needed for turnover number (TON) determination. Thus, transitioning from microscale assays to preparative-scale applications is pivotal in evaluating the practical utility and robustness of UPOs.

The first preparative-scale biotransformation in this work was carried out using *Mth*UPO with *N*-protected phenethylamine, yielding the benzylic hydroxylation product *S*-(+)-2-*N*-phthaloyl-1-phenylethanol with 98.6% ee and 57% isolated yield (9.70 mg). This transformation provides access to pharmacologically important scaffolds such as β-blockers and sympathomimetics.[56] Concentrated *P. pastoris* supernatant containing the secreted enzyme was used directly without enzyme purification (see Chapter II).

In the subsequent study, a directed evolution campaign improved the enantioselectivity of *Mth*UPO variants toward 4-hydroxy-β-ionone. A preparative-scale reaction with one evolved variant yielded 34 mg (55%) of *S*-4-hydroxy-β-ionone, but the optical rotation alone was insufficient to unambiguously assign the absolute configuration due to the minimal difference between the *S*- and *R*-forms. Therefore, a second preparative synthesis using the enantiocomplementary variant was performed, followed by derivatization, affording 51.1 mg (37%) of *R*-4-acetoxy-β-ionone. The absolute configurations was conclusively determined. Both reactions used non-purified *S. cerevisiae* supernatant (see Chapter IV).

Further preparative-scale experiments were conducted with UPO variants designed via FuncLib. Reactions with nerol and *S*-(-)-limonene yielded 132 mg (55%) of 2,3-epoxy nerol and

3 mg (1%) of isopiperitenol, respectively. The former was synthesized as it is not commercially available and was required to determine the TONs of various UPO designs. The latter was prepared due to limitations in GC-MS identification, necessitating NMR-based structural analysis. In both cases, concentrated *S. cerevisiae* supernatant was used directly (see Chapter V).

The use of non-purified supernatants streamlines the workflow, enabling rapid transition from enzyme production to catalysis without labor-intensive purification steps. Control experiments with host strains carrying empty vectors confirmed the absence of background activity for the tested substrates, affirming that all observed product formation originated from the UPOs. Despite high UPO expression levels in both *S. cerevisiae* and *P. pastoris*, significant quantities of other proteins remain in the supernatant, even after medium exchange via ultrafiltration. This complexity poses challenges for downstream processing: phase separation during product extraction with organic solvents was often time-consuming or incomplete. In such cases, water was removed using sodium sulfate, though this sometimes reduced yields. Consequently, several reactions had to be repeated with increased starting material to ensure sufficient product quantities for analysis. Nonetheless, these preparative-scale reactions convincingly demonstrate that UPOs are not limited to microliter-scale experiments but are amenable to larger-scale synthetic applications.

Recent advances in UPO production further support their preparative and industrial relevance. For example, fed-batch fermentation has achieved titers of 217 mg/L for PaDa-I expressed in *P. pastoris*,[37] while submerged fermentation of *Aae*UPO in *A. aegerita* yielded 265 U/L using vinasse, a by-product of bioethanol production, as a nutrient source.[249] Preparative applications include the synthesis of 1.4 g (ω-1)-hydroxytetradecanoic acid,[250] and the production of 520 g cyclohexanol/cyclohexanone using recombinant *Aae*UPO in *P. pastoris*.[251] Complementary developments such as total turnover numbers exceeding 400,000 for the conversion of ethylbenzene to *R*-1-phenylethanol,[29,231] and continuous-flow UPO catalysis with a space-time yield of 0.97 g L^{-1} h^{-1}[252] further strengthen the case for UPOs as emerging catalysts of industrial relevance.

8.3 Engineering Strategies and Objectives

While suitable expression systems and preparative-scale reactions are essential for demonstrating the feasibility of UPOs in synthetic applications, their catalytic performance and substrate versatility often require targeted engineering. The following section presents the engineering strategies employed to enhance UPO stability, enantioselectivity, and substrate scope, highlighting how smart library design and computational tools were integrated to address these objectives within the context of this work.

8.3.1 Ab Initio Structure Prediction Tools for Stability Engineering

The challenging heterologous expression and high glycosylation of fungal UPOs are key reasons why, to date, only four crystal structures have been reported. These limited structural resources

were initially used to create homology models to guide rational enzyme engineering, as reflected in earlier work in this thesis (Chapter IV). However, recent advances in ab initio structure prediction such as AlphaFold2[191] have dramatically expanded structural accessibility for many enzymes and across the diverse UPO sequence space. With thousands of putative UPOs available in genomic databases, high-confidence structural models can now be generated without relying on experimental crystallography.

In this thesis, ab initio predicted structures served as a foundation for several enzyme engineering efforts. In two studies, AlphaFold2 models enabled structure-guided mutational design using the tools PROSS[226] and FuncLib,[228] which incorporate both sequence conservation and structural context to suggest stabilizing or functionally relevant mutations (Chapters III and V). Notably, both campaigns achieved their respective goals—enabeling functional enzyme expression and modulating activity and selectivity for several terpene susbtrates—despite relying entirely on predicted rather than experimental structures. These results confirm the practical utility of ab initio models in diverse engineering objectives for UPOs.

AlphaFold2 does not include cofactors such as heme in its predictions. Accordingly, in the discussed studies, heme groups were manually positioned by structural alignment with known UPO crystal structures. Recent tools such as AlphaFold3[253] and AlphaFill[254] now address this limitation by enabling integrated modeling of protein-cofactor complexes, thereby further streamlining computational enzyme design workflows—especially for cofactor-dependent enzymes like UPOs.

In Chapter III, AlphaFold2-predicted structures were used in a multi-step strategy combining signal-peptide shuffling with PROSS-based mutagenesis. This approach aimed to enhance secretion and functional expression in yeast hosts and was successful for 9 out of 10 UPOs tested, whereas only one wild type sequence showed detectable secretion and activity. The study illustrates how structure prediction can facilitate early-stage enzyme development by enabling access to candidates previously inaccessible due to poor expression. In addition, this strategy offers a valuable route for stabilizing enzymes with promising catalytic profiles, serving as a foundational step prior to further functional optimization.

Improvements in activity or selectivity through protein engineering often come at the cost of reduced stability.[255–257] Enhancing enzyme stability can therefore serve as a foundation for further discovery: mutations that improve activity or selectivity may otherwise remain undetectable if overall enzyme activity is lost due to instability. Beyond this enabling role, stability is also a key requirement for industrial applications, where enzymes must withstand elevated temperatures and prolonged process times. Increased stability not only extends enzyme lifetime but also supports faster reactions and higher space-time yields. Thus, stabilizing an enzyme is frequently a critical early step—both to unlock new mutational opportunities in the laboratory and to pave the way for industrial implementation.

For identifying a suitable starting point toward a desired reaction, commercially available UPO panels provide a convenient entry into application. However, to systematically explore the vast di-

versity of uncharacterized UPOs, a structure-driven approach combining accurate ab initio models with stability-focused design can be employed. While the former route enables rapid initial screening, the latter has the potential to uncover enzymes with superior properties and to provide robust foundations for subsequent functional optimization, enhancing the UPO toolbox.

8.3.2 Improving Enantioselectivity of Oxyfunctionalizing Enzymes

In addition to activity and stability, the selectivity of enzymes represents a major advantage in biocatalysis. Amino acid residues in the active site can position substrates for chemo-, regio-, and stereoselective transformations. This inherent capability enables biocatalysts to access synthetic routes to chiral products, which are often otherwise obtained using costly chiral chemical catalysts. However, engineering the enantioselectivity of enzymes remains a significant challenge. In contrast to activity engineering, which frequently relies on rapid colorimetric assays or short GC and HPLC runs, the similar physical and chemical properties of enantiomers complicate their separation and quantification. This often necessitates longer retention times in chromatographic methods and renders colorimetric differentiation infeasible, making large-scale variant screening time-intensive.

In this work, for engineering the enantioselectivity of the aliphatic hydroxylation of β-ionone (see Chapter IV), an initial activity-focused screening was performed using MISER-GC-MS (see Chapter I). Enantioselectivity was subsequently assessed for the most active variants from each round using chiral GC-MS. The underlying hypothesis was that increased activity may reflect improved substrate recognition and binding, which could concurrently affect enantioselectivity. This strategy proved successful, yielding enantiodivergent variants with improved activity and enantiomeric ratios (e.r.) of 96.6:3.4 (*R*) and 0.3:99.7 (*S*), respectively. This campaign represents the first successful attempt to improve targeted the enantioselectivity of an oxyfuncitionalzation reaction with an unspecific peroxygenase displaying initially low selectivity.

Many reactions catalyzed by cytochrome P450s (P450s) and UPOs exhibit inherently high enantioselectivity. For instance, a P411 heme protein catalyzing carbene transfer to internal alkynes achieved >99.9% enantiomeric excess in cyclopropene synthesis without prior engineering. Directed evolution was subsequently employed to enhance activity while preserving this high enantioselectivity.[258] However, enantioselectivity engineering is often necessary, particularly for promiscuous reactions.

In another study, both activity and enantioselectivity of the α-hydroxylation of β-ketoesters catalyzed by a UPO from *Aspergillus niger* were improved using two rounds of site-directed saturation mutagenesis, screening 5100 variants in total. The resulting variant achieved 99:1 e.r., 4140 total turnover number (TTN), and 97% yield. Each reaction was analyzed via reverse-phase HPLC using a chiral column. While this approach is advantageous in capturing both highly active and highly selective variants, the analytical demand is considerable. With an average HPLC runtime of 15 minutes per sample, the complete screen required over 50 days of instrument time.[259]

An alternative strategy involving initial colorimetric activity screening followed by selective enantioselectivity analysis was applied in the engineering of 4-oxolocrotonate tautomerase for the epoxidation of α,β-unsaturated aldehydes such as citral. This artificial cofactor-independent peroxygenase utilizes hydroperoxides (e.g., *t*-BuOOH or H_2O_2), and substrate consumption was monitored via spectrophotometric kinetic assays. Enantioselectivity was determined only for the best-performing variants. Following two rounds of single-site saturation mutagenesis, variants with up to 98% ee, high conversion (up to 98%), and product yields of 50–80% were identified. Notably, the choice of hydroperoxide had a pronounced impact on the stereochemical outcome, in some cases even reversing enantioselectivity.[260]

In a further study, a spectrophotometric method enabling the simultaneous detection of activity and enantioselectivity was established for UPO-catalyzed hydroxylations yielding secondary alcohols. The system employed an enzyme cascade where each enantiomer of the product alcohol was selectively oxidized by a specific alcohol dehydrogenase (ADH), generating either NADH or NADPH. Fluorescence readout of NAD(P)H formation provided a proxy for the enantiomeric excess. This setup enabled the screening of 44 UPOs for enantioselectivity in the hydroxylation of 1-phenylethanol and *n*-octane. Although efficient, the method requires the availability of highly selective reporter enzymes for each enantiomer, which poses a limitation when targeting novel or non-commercially available products.[261]

A further strategy to reduce screening effort is the design of small, smart libraries. One study applied this approach to engineer a non-natural P450 for *R*-selective styrene epoxidation. A dual-functional small molecule (DFSM) was introduced to assist H_2O_2 activation and influenced substrate orientation, improving ee from 7% to 84%. Over three rounds of rational mutagenesis, introducing single and double mutations and combining beneficial mutations, a final variant with four mutations was obtained after screening around 60 variants, exhibiting >4000 turnover number and 98% ee. Analytical evaluation relied on single chiral GC and HPLC measurements.[262]

A similar strategy was applied in the present study (Chapter V) for terpene functionalization using a computationally designed library of 50 *Mth*UPO variants. Among the targeted substrates was again β-ionone, previously examined in Chapter IV. Two variants from this library displayed e.r. values of 1.1:98.9 in favor of *S*-4-hydroxy-β-ionone. Design 11 showed 16,000 TON with 87% regioselectivity, while Design 34 reached 8,500 TON with 98% regioselectivity. These performances were slightly lower than the best variant from the directed evolution campaign in Chapter IV, which exhibited an e.r. of 0.3:99.7 and 41,000 TON for the same product. Additionally, the directed evolution campaign yielded an enantiodivergent variant with an e.r. of 96.6:3.4 for the *R*-enantiomer and 28,000 TON. Notably, the positions selected for diversification in the design of the library in Chapter V included all positions that were mutated in the top-performing variants from Chapter IV, indicating that, in principle, the same high-performing designs could have been sampled within the computationally designed library. While the designed library achieved comparable selectivity with significantly fewer variants for one enantiomere, it did not produce any variants

with high activity and selectivity for the *R*-selective product. Small libraries, whether rationally or computationally designed, can provide an efficient shortcut to identifying variants with improved enantioselectivity; however, in many cases, these variants may serve better as starting points for subsequent directed evolution rather than yielding optimal solutions in a single iteration.

Overall, these studies illustrate that the engineering of enantioselective oxyfunctionalization remains a complex task with no universally applicable strategy. Spectrophotometric assays can significantly reduce screening time but are generally limited to substrates with colorimetric or fluorescent properties or rely on elaborate enzyme cascades. The MISER-GC-MS method enables rapid activity screening across diverse substrates, but provides no direct information on enantioselectivity. Nevertheless, an approach that prioritizes activity screening followed by enantioselectivity analysis, as applied in this work and others, can yield highly enantioselective and active variants.

Looking forward, machine learning (ML) tools may offer new opportunities in the engineering of enantioselective enzymes. Particularly in campaigns involving multiple rounds of evolution or combination of beneficial mutations, ML-guided design could streamline variant selection and enhance hit rates. The potential of ML in this context will be discussed in more detail in 8.4.2.

8.3.3 Multisubstrate Engineering Using Smart Libraries

Regardless of the chosen approach, engineering an enzyme for high activity or selectivity toward a single substrate is typically a laborious and time-intensive process. Extending this process to multiple substrates multiplies the experimental workload. If biocatalysts for several substrates are required, it is advantageous to incorporate multiple substrates directly into the screening strategy—either in an iterative manner or in direct competition. This enables the assessment of enzyme promiscuity, identification of mutational hot-spots, and the simultaneous discovery of variants with improved properties toward different substrates. Such approaches are variously termed fingerprinting, substrate cocktail screening, or multisubstrate/multiplexed assays.[263] When coupled with appropriate analytical methods, these strategies also reveal shifts in chemo-, regio-, and enantioselectivity, and can uncover entirely new products.

In a study using substrate-multiplexed screening (SUMS), an L-tryptophan decarboxylase (*Rgn*TDC) and a previously engineered β-subunit of tryptophan synthase from the thermophilic archaeon *Pyrococcus furiosus* (*Pf*TrpB2B9, or 2B9) were diversified by site-saturation mutagenesis, and used to convert different tryptophane analogs forming tryptamines. The product mixtures were analyzed simultaneously via UPLC-MS with single-ion monitoring of distinct products. The multiplex approach revealed pronounced shifts in substrate promiscuity and activity. Comparison with single-substrate runs confirmed the same general trends, validating multiplex screening as an efficient strategy for guiding protein engineering toward improved activity across multiple substrates, mapping altered substrate scope, and pinpointing mutational hot-spots in a single, highly customizable experiment.[263]

A similar concept was applied to engineer *Mth*UPO, using simultaneous reactions with three substrates—octane, cyclohexane, and cyclohexene—and analyzing six products from a single reaction mixture. Site-saturation mutagenesis at nine positions was followed by one-pot reactions, with product identification by GC-MS (electron-impact ionization, EI) in MISER mode, enabling the rapid screening of up to 100 reactions within 2 h. To resolve octanol regioisomers, *N*-trimethylsilyl imidazole derivatization was used, enabling regiospecific fragmentation patterns. Significant shifts in product distribution were observed; notably, variant A161L produced 1-octanol, a terminal hydroxylation product rarely formed by unspecific peroxygenases. This product was detected only upon rescreening individual variants with single substrates, illustrating a limitation of the single-ion detection mode in multiplex GC-MS, where unexpected products may be missed. Additional drawbacks include the need to find optimal ion traces for each product and the extra derivatization step. Nevertheless, extending MISER-GC-MS to multisubstrate formats and detecting product distribution shifts highlights its value for multisubstrate enzyme engineering.[71] In general, scan-mode or high-resolution MS analysis is advisable in exploratory campaigns to avoid overlooking unanticipated reaction products.

The increasing availability of computational tools has opened new opportunities for rationally designing enzymes with activity toward multiple substrates from the outset. Structure prediction methods such as AlphaFold2 now provide high-confidence three-dimensional models even in the absence of experimental structures,[191] offering a crucial starting point for subsequent design and analysis. These structural models can be directly used with dedicated approaches such as multichemical state analysis (MCSA),[264] which explicitly models broad-specificity catalysts, or with tools such as FuncLib to generate libraries of diverse but stable enzyme variants, whose structural and chemical diversity increases the likelihood of displaying broad or altered activity profiles.[228] Such methods are particularly powerful when combined with knowledge of active-site architecture, dynamic bottlenecks, and substrate-access tunnel topology, the latter of which can be identified using tools such as `CAVER`,[218,219] or with stability-enhancing algorithms such as PROSS.[225,226]

FuncLib was previously used to generate a library of the engineered *Aae*UPO variant PaDa-I, producing 30 designs each containing 4–5 mutations. Of these, 24 were functionally expressed, and several exhibited pronounced changes in enantiodivergence and regioselectivity across a set of aromatic substrates, in some cases shifting oxyfunctionalization from alkyl to aromatic hydroxylation. These shifts were often the result of positive epistatic effects, underlining the importance of exploring combinatorial mutations rather than relying solely on single-residue substitutions.[265]

In the present work (Chapter V), FuncLib was applied to engineer *Mth*UPO simultaneously for different terpene substrates. Unlike in the PaDa-I study, all 50 designs (100% expression success) were active with at least one tested substrate. Notably, some designs exhibited over 1,880-fold improved activity toward the peroxidase substrate 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS). The same designs also performed exceptionally with 2,6-dimethoxyphenol (DMP), another peroxidase substrate, but were poor performers with peroxygenase substrates.

Conversely, another design displayed a 3.1- to 6.7-fold improvement across four structurally related rose-ketone terpenes, illustrating that generalist catalysts for structurally or mechanistically related substrates can be obtained via computationally guided, one-shot library design. The choice between evolving generalists versus highly specialized catalysts should be made deliberately, as broad scope often comes at the cost of peak performance on individual targets.[266,267]

Several designs also exhibited pronounced shifts in chemo- and regioselectivity, accompanied by the formation of additional novel products. For example, wild type *Mth*UPO predominantly epoxidizes limonene to limonene oxide, with only minor hydroxylation to carveol (<10%). Several designs shifted this distribution substantially, yielding up to 56% carveol in some cases, and as much as 77% of the hydroxylated product isopiperitenol in others—a compound detected only in trace amounts in the wild type reaction. For geraniol, the wild type enzyme primarily oxyfunctionalized the terminal alcohol group to citral A (40%), while also oxidizing each of the two double bonds (47% and 14%, respectively). In contrast, one FuncLib design exhibited >99% conversion to citral A. Strong effects on enantioselectivity were likewise observed (see 8.3.2). Collectively, these alterations in activity, chemo-, regio-, and enantioselectivity underscore the substantial functional diversity attainable through this design strategy, particularly when evaluated across a chemically diverse substrate panel.

The accumulated data provide a valuable basis for elucidating residue–substrate interactions within the active site and for dissecting epistatic effects arising from combinations of mutations. These insights can be amplified through molecular dynamics simulations of selected variants (as applied to certain designs in this study) or by integrating activity and selectivity data into machine-learning frameworks. In the multisubstrate engineering context, such approaches are particularly powerful, as they enable simultaneous modelling of variant performance across diverse substrates, revealing structure-function patterns and quantifying trade-offs between generalist and specialist profiles.[268–270] Computational integration can thus predict mutation effects across related substrates, prioritize promising variants for experimental validation, and support rational, data-driven navigation of complex multisubstrate fitness landscapes.[221,271]

In (Chapter V), *Mth*UPO was not optimized for stability prior to initiating the engineering campaign, for example by employing the PROSS algorithm. This decision was based on prior studies showing that *Mth*UPO retains sufficient stability even when incorporating activity- and selectivity-enhancing mutations.[52,71,272] Nevertheless, it should be noted that gains in activity or selectivity can sometimes come at the expense of stability.[273] The positions selected for mutagenesis in this work had previously been shown to positively influence activity or selectivity; such prior knowledge is not always available for other enzyme systems.

8.4 Integrative and Emerging Approaches in Enzyme Engineering

Enzyme engineering sits at a crossroads where traditional biochemical methods intersect with cutting-edge computational technologies. The challenge of tailoring enzymes for specific industrial, environmental, or medical applications calls for strategies that simultaneously optimize multiple properties such as stability, selectivity, and substrate scope. While much of this work draws on insights from unspecific peroxygenases, the principles and approaches discussed extend well beyond this enzyme family. Recent advances in data-driven approaches, including machine learning and high-throughput screening, are transforming how these complex objectives are addressed, enabling exploration of vast sequence spaces that were once inaccessible. Yet, despite the power of in silico design and predictive algorithms, the tangible realities of enzyme function in a laboratory setting remain indispensable. This interplay between computation and experiment is redefining what is possible in enzyme engineering, offering new avenues to develop catalysts that are not only highly efficient and selective but also robust and adaptable to diverse real-world demands.

8.4.1 From Established Methods to Integrated Workflows

For enzymes with limited prior characterisation, a logical engineering strategy—once functional expression has been achieved—often begins with improving stability using tools such as PROSS.[226] Subsequent diversification can be introduced via algorithms such as FuncLib,[228] targeting active-site residues or positions lining the substrate-access tunnel. The resulting variant library can then be screened against either a defined set of substrates or a chemically diverse panel to identify both generalist and specialist candidates. Mapping screening outcomes onto an initial fitness landscape can help reveal promising regions of sequence space, thereby guiding subsequent iterative optimisation and enabling multi-substrate engineering across different enzyme classes.

Central to the success of any engineering effort is the clear definition of campaign objectives. Determining whether the goal is to obtain a generalist or specialist biocatalyst, to enhance stability, activity, selectivity, or a combination thereof, will dictate the choice and order of methods. This process also requires an assessment of available structural and functional information: whether high-resolution structures are accessible or require prediction (e.g., AlphaFold),[191] whether the active site and substrate tunnel are mapped, and whether expression levels or secretion efficiency are limiting factors. In some cases, targeted approaches such as signal-peptide shuffling can improve secretion, while thermostability enhancement may be a prerequisite in others. Likewise, the nature of the desired screen—high-throughput versus focused, chiral resolution versus activity-based—should be determined early, as it will influence both library design and analytical requirements.

This decision-making process can be envisioned as a decision tree, where each answer—

defining objectives, evaluating available data, selecting screening methods—helps narrow down the optimal combination of computational and experimental tools. By systematically addressing key questions such as: "Is structural data available or is modeling required?", "Is secretion a bottleneck?", "Is activity or selectivity the priority?", and "What throughput can the screening support?", an integrated and tailored workflow can be constructed that balances feasibility with engineering goals.

Engineering campaigns are most effective when iterative cycles of computational design and experimental validation are employed, allowing continuous refinement of hypotheses and models based on accumulating data. This cyclical approach helps to efficiently navigate the vast sequence space while avoiding overreliance on any single method. The choice of tools and strategies is often shaped by practical constraints such as available time, budget, and laboratory infrastructure, which in turn influence the scale and throughput of screening efforts. Furthermore, enzyme class-specific considerations—such as whether the enzyme is secreted or intracellular, or its catalytic mechanism—can impact the selection of engineering targets and experimental workflows, underscoring the need for adaptable, context-sensitive approaches.

Modern engineering campaigns increasingly benefit from integrating high-content datasets. When screening large variant libraries, the use of next-generation sequencing (NGS) allows the full mapping of sequence-function relationships, producing datasets that can subsequently be mined for fitness landscapes or serve as training material for machine learning models.[274] For broader exploration of sequence space, methods such as EVcouplings[275] enable the introduction of coordinated mutations derived from evolutionary covariation analyses, thus extending beyond the active-site-focused diversity of approaches like FuncLib.

A broad spectrum of computational tools is now accessible through user-friendly webservers, many of which require minimal prior training in computational chemistry or structural biology. Likewise, molecular biology resources such as ready-to-use cloning kits (e.g., signal-peptide shuffling kits from repositories such as Addgene) facilitate experimental implementation. However, the effective use of such tools requires an awareness of their assumptions, limitations, and the contexts in which they deliver reliable predictions. Without this, computational outputs risk being misinterpreted or misapplied. Consequently, the integration of computational chemistry and bioinformatics training into the core curriculum for chemistry and biochemistry is essential for equipping future researchers to navigate and exploit these resources effectively.

8.4.2 Data-Driven Horizons: Machine Learning and Beyond

Although many sophisticated enzyme engineering methods can improve desired properties, they typically explore only a limited fraction of the enzyme's fitness landscape. Directed evolution, for example, often risks becoming trapped in local optima, potentially overlooking variants with superior or desired traits that exist elsewhere in sequence space (**Fig. 2**).

Machine learning (ML) models address this challenge by constructing predictive mappings be-

tween protein sequences and functional outputs, effectively approximating the fitness landscape based on limited experimental data. These models can then predict the properties of unseen variants, enabling a more efficient search for improved enzymes. Bayesian approaches, such as Gaussian process models, are particularly valuable because they not only provide predictions but also quantify the uncertainty associated with those predictions. This uncertainty quantification allows for prioritization of variants that are both promising and underexplored, accelerating iterative engineering cycles.[274]

A key example demonstrated the use of Gaussian process ML to optimize thermostability in chimeric cytochrome P450 enzymes. Starting from a design space restricted to single- and double-crossover chimeras (10^{10} potential sequences) and a modest training set of 261 variants, the approach rapidly identified thermostable variants outperforming all previously known constructs. After seven iterative rounds of prediction, experimental validation, and retraining, the top variant exhibited a 5.3 °C improvement in melting temperature compared to the best predecessor, differing by 23 mutations.[274]

Despite its promise, the accessibility of ML approaches to enzyme engineers without specialized computational expertise remains variable. Unlike many established bioinformatics tools and webservers, ML workflows often require familiarity with programming languages, data preprocessing, model training, and hyperparameter optimization. While emerging platforms and user-friendly interfaces (e.g., EvoML, DeepEnzyme)[276] are beginning to lower these barriers, the integration of ML into routine enzyme engineering pipelines is still limited by the need for adequate training data, computational resources, and domain knowledge.[277,278]

Current limitations of ML in enzyme engineering include data scarcity and bias—small or narrowly distributed datasets can cause models to overfit or fail to generalize to distant sequence space. Additionally, the interpretability of ML models can be challenging, making it difficult to rationalize the mechanistic basis of predicted improvements. Furthermore, ML models typically require continuous retraining as new data become available, underscoring the importance of well-designed experimental campaigns and iterative workflows.[279,280]

Looking forward, advances in transfer learning, unsupervised learning on large protein sequence databases, and integration of structural and dynamic information promise to enhance model accuracy and generalizability. Increasing availability of large, standardized, and high-quality datasets—coupled with improvements in computational hardware and algorithmic efficiency—will further democratize ML tools. These developments are expected to transform enzyme engineering into a more predictive and data-driven discipline, reducing the reliance on laborious trial-and-error experimentation.[281,282]

Beyond ML, integration with molecular dynamics, quantum mechanics/molecular mechanics (QM/MM) simulations, and systems biology approaches will likely create powerful hybrid platforms that combine mechanistic understanding with data-driven predictions, enabling rational design of enzymes with complex, multi-parameter optimization objectives.

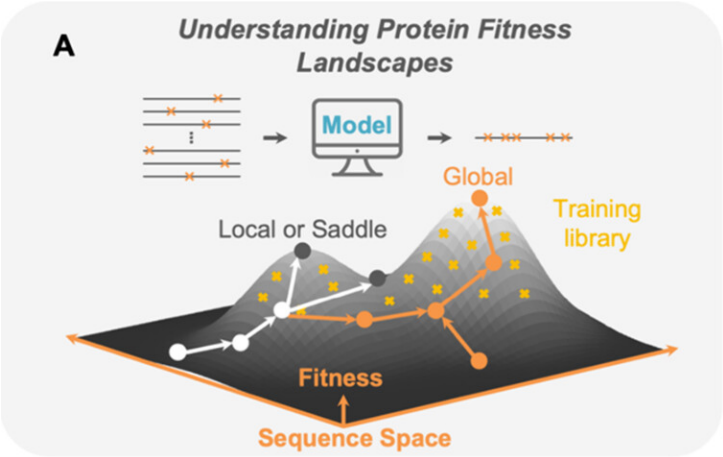


Figure 2 ML models can propose combinations of mutations that might not be found with traditional directed evolution approaches. This allows larger jumps in sequence space and enables finding variants with desired properties.[†]

8.4.3 The Relevance of the Wet Lab in a Computational Era

With the rapid advancement and increasing accessibility of computational tools, it is natural to question the evolving role of wet-lab researchers in enzyme engineering. Could the entire process eventually be performed in silico, or will hands-on experimental validation and optimization remain essential? While computational methods have transformed many aspects of enzyme design—from stability prediction to substrate specificity modeling—real-world experiments still serve as the definitive test of enzyme function.

Currently, computational approaches excel at hypothesis generation, narrowing down vast sequence spaces to manageable sets of promising candidates. Techniques such as molecular dynamics (MD) simulations can provide atomistic insights into enzyme-substrate interactions and stability; however, these simulations are computationally intensive. For example, running 800 MD trajectories to evaluate variant dynamics may require weeks or months of GPU-accelerated computing time, depending on simulation length and system complexity.[283,284] By contrast, modern high-throughput wet-lab screening platforms—such as droplet-based microfluidics and automated robotic workflows—can assay millions of enzyme variants per day at relatively low cost and with high precision.[285,286]

This throughput disparity underscores the complementary nature of computational and experimental approaches: computational pre-screening focuses efforts and reduces the number of variants requiring costly physical testing, while wet-lab assays provide critical validation of computational predictions. The experimental data generated also fuel iterative refinement of machine learning models and other in silico tools, closing the design-build-test-learn cycle.[274,278]

[†]*Reproduced* from **Yang, J.; Li, F.-Z.; Arnold, F. H.**, Opportunities and challenges for machine learning-assisted enzyme engineering, *ACS Central Science*, **2024**, *10* (2), 226–241.[277]

Despite increasing automation in the wet lab—through robotics, liquid handling systems, and integrated screening platforms—certain tasks still demand the experience and judgment of skilled experimentalists. Designing meaningful screening assays, interpreting ambiguous or unexpected results, troubleshooting protocols, and handling complex reaction conditions often require human expertise. Additionally, some enzyme properties such as long-term stability, conformational heterogeneity, and rare side reactions remain difficult to capture fully through computational means and require empirical investigation.[287–289]

Beyond these practical considerations, several broader aspects influence the interplay between computational and experimental enzyme engineering:

- Experimental variability and noise: Computational models typically assume idealized conditions, while wet-lab experiments reveal the biological and chemical complexities—including batch-to-batch variation and assay limitations—that affect enzyme performance and reproducibility.

- Costs and resource allocation: Balancing investment in compute time, reagents, and personnel is critical. While computational resources can be expensive and time-consuming, wet-lab consumables and labor costs can also be significant, influencing strategy decisions.

- Data sharing and reproducibility: Standardized data formats, open repositories, and transparent reporting practices are vital to enable effective feedback loops between computational predictions and experimental validations, fostering community progress.[290]

- Emerging technologies: Advances in AI-guided robotic platforms and ultra-high throughput screening techniques, such as cell-free expression systems and microfluidics, are rapidly blurring the boundaries between computational and experimental workflows, promising more integrated and efficient enzyme engineering pipelines.

Looking forward, fully in silico enzyme design and screening for complex traits remains challenging. Accurate simulations often require prohibitive computational resources, and predictive models must be continually validated and refined against experimental data. Standardized benchmarking of in silico predictions against wet-lab results is an important ongoing effort.[277] However, as automation and machine learning improve, the gap between computational prediction and experimental validation is expected to narrow.

In conclusion, the future of enzyme engineering will likely feature increasingly integrated workflows that leverage the strengths of both computational and experimental methods. Rather than one replacing the other, these approaches will continue to complement each other, with computational tools guiding design and experimentation providing the definitive proof, enabling iterative cycles of improvement.

9 | Authorship declaration

Introduction:

Enzymatic Hydroxylations of sp^3-Carbons

by: **Judith Münch**, Pascal Püllmann, Wuyuan Zhang, Martin J. Weissenborn

in: *ACS Catalysis* **2021**, *11* (15), 9168–9203; doi: https://doi.org/10.1021/acscatal.1c00759

Estimated overall contribution: 80%
Specific tasks: Conception of the review article (with MJW); research of the content; approx. 95% of the writing of the review article (further contributions from MJW, PP and WZ).

Chapter I:

Secretion and directed evolution of unspecific peroxygenases in *S. cerevisiae*

by: Niklas Dietz, Li Wan, **Judith Münch**, Martin J. Weissenborn

in: *Methods in Enzymology* **2023**, *693*, Elsevier; doi: https://doi.org/10.1016/bs.mie.2023.09.013

Estimated overall contribution: 15%
Specific tasks: Planning and execution of experimental work described in Chapter 7 and Chapter 8; design and writing of Chapter 8: MISER-GC-MS (content and figures), supportive part in writing of the manuscript (mainly ND, LW and MJW).

Chapter II:

A modular two yeast species secretion system for the production and preparative application of unspecific peroxygenases

by: Pascal Püllmann, Anja Knorrscheidt, **Judith Münch**, Paul R. Palme, Wolfgang Hoehenwarter, Sylvestre Marillonnet, Miguel Alcalde, Bernhard Westermann, Martin J. Weissenborn

in: *Communications Biology* **2021**, *4*, 562; doi: https://doi.org/10.1038/s42003-021-02076-3

Estimated overall contribution: 10%
Specific tasks: Creation of Fig. 1C, 2A, 3, 4B, 4D, 4E, 5B; design, execution and evaluation of all biocatalytical conversions with *N*-phthaloyl-phenethylamine, including product isolation of (*S*)-(+)-2-*N*-phthaloyl-1-phenylethanol and determination of the absolute configuration; design, execution and evaluation of the haem carbon monoxide assay to determine the enzyme concentration within the supernanat.

Chapter III:

Functionally Diverse Peroxygenases by AlphaFold2, Design, and Signal Peptide Shuffling

by: **Judith Münch***, Niklas Dietz*, Shiran Barber-Zucker, Franziska Seifert, Susanne Matschi, Pascal Püllmann, Sarel J. Fleishman, and Martin J. Weissenborn

in: *ACS Catalysis* **2024**, *14* (7), 4738–4738; doi: https://doi.org/10.1021/acscatal.4c00883

Estimated overall contribution: 55%
Specific tasks: Conception, design and experimental planning of the research project (with MJW), gathering and analysis of approx. 70% of laboratory experimental data (further contributions under my supervision by ND), writing of the manuscript (with MJW, SJF and ND).

Chapter IV:

Computational-Aided Engineering of a Selective Unspecific Peroxygenase toward Enantiodivergent β-Ionone Hydroxylation

by: **Judith Münch***, Jordi Soler*, Nicole Hünecke, Dominik Homann, Marc Garcia-Borràs, and Martin J. Weissenborn

in: *ACS Catalysis* **2023**, *13* (13), 8963–8972; doi: https://doi.org/10.1021/acscatal.3c00702

Estimated overall contribution: 60%
Specific tasks: Conception, design and experimental planning of the research project (with MJW), gathering and analysis of approx. 95% of laboratory experimental data (further contributions by NH and DH), writing of the manuscript (with MJW and MGB).

Chapter V:

Computationally Designed Peroxygenases That Exhibit Diverse and Selective Terpene Oxyfunctionalization

by: **Judith Münch**, Jordi Soler, Ofir Gildor-Christal, Sarel J. Fleishman, Marc Garcia-Borràs, and Martin J. Weissenborn

in: *ACS Catalysis* **2025**, *15* (15), 12741–12755; doi: https://doi.org/10.1021/acscatal.5c02412

*shared first authorship

Estimated overall contribution: 70%
Specific tasks: Conception, design and experimental planning of the research project (with MJW and SF), creation of the FuncLib library (supported by SF); gathering and analysis of all laboratory experimental data, writing of the manuscript (with MJW, SF and MGB).

10 Curriculum vitae

Judith Münch

Education

Since 09/2019	PhD in Chemistry, Martin-Luther University Halle-Wittenberg and Leibniz Institute for Plant Biochemistry, Halle
04/2017 - 07/2019	MSc. in biochemistry Martin Luther University Halle Wittenberg, Halle (09.07.2019)
10/2014 - 03/2017	BSc. in chemistry University of Leipzig, Faculty of Chemistry and Mineralogy (02.03.2017)
09/2009 - 07/2013	Abitur at Albert-Schweitzer-Gymnasium, Erfurt – school specialized in STEM (28.06.2013)
09/2005 - 08/2009	Secondary School Gymnasium Ernestinum, Gotha
09/2001 - 08/2005	Elementary School Sonneborn, Sonneborn

Skills

- Managing science-policy and research projects with national and international partners (e.g., Weizmann Institute, Israel; University of Girona, Spain, Volkswagen Foundation)
- Networking and advocacy with academic and political stakeholders at the national level
- Scientific and science-policy communication (peer-reviewed publications, policy statements, blogs, presentations, event moderation)
- Engineering enzymes for improved stability, activity, and selectivity
- Molecular cloning, heterologous expression, and purification of proteins in *E. coli*, *S. cerevisiae*, and *P. pastoris* (secreted and non-secreted)
- Applying GC-MS (chiral and non-chiral; medium- and high-throughput)
- Developing and implementing assays for protein expression and enzyme activity control

Professional experience

Since 06/2024	Coordinator of the German Postdoc Network, located at Helmholtz-Munich
09/2022-02/2024	PhD student at GRK 2670 Beyond Amphiphilicity (BEAM), RTG 2670 Self-Organization of Soft Matter via Multiple Noncovalent Interactions
05/2022	Development of enzyme cascades with UPOs, CARs, KREDs and IREDs Research stay at the Manchester Institute for Biotechnology, group of Prof. Sabine Flitsch and Prof. Nicolas Turner
06/2018 - 08/2018	Isolation and characterization of aldoxime dehydratases Research stay at the University of Toyama, Japan, group of Prof. Asano

Publication list

(1) Münch, J.; Soler, J.; Gildor-Christal, O.; Fleishman, S. J.; Garcia-Borràs, M.; Weissenborn, M. J.; Computationally Designed Peroxygenases That Exhibit Diverse and Selective Terpene Oxyfunctionalization. *ACS Catal.* **2025**, *15*, 12741-12755

(2) Münch, J.; Dietz, N.; Barber-Zucker, S.; Seifer, F.; Matschi, S.; Püllmann, P.; Fleishman, S.; Weissenborn, M. J.; Functionally diverse peroxygenases by AlphaFold2, design and signal peptide shuffling. *ACS Catal.* **2024**, *14*, 4738-4748
(3) Dietz, N.; Wan, L.; Münch, J.; Weissenborn, M. J. Secretion and directed evolution of unspecific peroxygenases in S. cerevisiae. *Methods Enzymol.* **2023**, *693*, 267-306.
(4) Münch, J.; Soler, J.; Hünecke, N.; Homann, D.; Garcia-Borràs, M.; Weissenborn, M. J. Computational-Aided Engineering of a Selective Unspecific Peroxygenase toward Enantiodivergent β-Ionone Hydroxylation. *ACS Catal.* **2023**, *13*, 8963-8972.
(5) Münch, J.; Püllmann, P.; Zhang, W.; Weissenborn, M. J. Enzymatic hydroxylations of sp3-carbons. *ACS Catal.* **2021**, *11* (15), 9168-9203.
(6) Püllmann, P.; Knorrscheidt, A.; Münch, J.; Palme, P. R.; Hoehenwarter, W.; Marillonnet, S.; Alcalde, M.; Westermann, B.; Weissenborn, M. J. A modular two yeast species secretion system for the production and preparative application of unspecific peroxygenases. *Commun. Biol.* **2021**, *4* (1), 562.
(7) Gáliková, M.; Klepsatel, P.; Münch, J.; Kühnlein, R. P. Spastic paraplegia-linked phospholipase PAPLA1 is necessary for development, reproduction, and energy metabolism in Drosophila. *Sci. Rep.* **2017**, *7* (1), 46516.

Miscellaneous

IT and computer skills	Good knowledge of Word, PowerPoint, Excel and R Basic knowledge of Python, Java, WordPress
Further training / courses	Project Management Training (12 hours), 2025 Good Manufacturing Practice (GMP) Training (12 hours), 2023
Languages	German: Native speaker English: Business fluent French: Basic communication skills
Scholarships	09/2019 - 08/2022 Doctoral scholarship from Friedrich Naumann Foundation for Freedom 06/2018 - 08/2018 JSPS Summer Program Research Fellowship in Japan 10/2013 - 03/2019 Scholarship holder of the Friedrich Naumann Foundation for Freedom
Conferences	Vitae Conference Birmingham 2024 (participant) Biotrans La Rochelle 2023 (poster) NextGenBiocat Graz 2023 (contributed talk) BEAM Symposium 2023 (contributed talk) Biotrans Graz 2021 (digital, poster)

11 List of publications

Peer-reviewed publications

1. **Computationally Designed Peroxygenases That Exhibit Diverse and Selective Terpene Oxyfunctionalization**
   by: **Judith Münch**, Jordi Soler, Ofir Gildor-Christal, Sarel J. Fleishman, Marc Garcia-Borràs, and Martin J. Weissenborn
   in: *ACS Catalysis* **2025**, *15* (15), 12741–12755; doi: https://doi.org/10.1021/acscatal.5c02412

2. **Functionally Diverse Peroxygenases by AlphaFold2, Design, and Signal Peptide Shuffling**
   by: **Judith Münch***, Niklas Dietz*, Shiran Barber-Zucker, Franziska Seifert, Susanne Matschi, Pascal Püllmann, Sarel J. Fleishman, and Martin J. Weissenborn
   in: *ACS Catalysis* **2024**, *14* (7), 4738–4738; doi: https://doi.org/10.1021/acscatal.4c00883

3. **Computational-Aided Engineering of a Selective Unspecific Peroxygenase toward Enantiodivergent β-Ionone Hydroxylation**
   by: **Judith Münch***, Jordi Soler*, Nicole Hünecke, Dominik Homann, Marc Garcia-Borràs, and Martin J. Weissenborn
   in: *ACS Catalysis* **2023**, *13* (13), 8963–8972; doi: https://doi.org/10.1021/acscatal.3c00702

4. **Secretion and directed evolution of unspecific peroxygenases in *S. cerevisiae***
   by: Niklas Dietz, Li Wan, **Judith Münch**, Martin J. Weissenborn
   in: *Methods in Enzymology* **2023**, *693*, Elsevier; doi: https://doi.org/10.1016/bs.mie.2023.09.013

5. **Enzymatic Hydroxylations of sp^3-Carbons**
   by: **Judith Münch**, Pascal Püllmann, Wuyuan Zhang, Martin J. Weissenborn
   in: *ACS Catalysis* **2021**, *11* (15), 9168–9203; doi: https://doi.org/10.1021/acscatal.1c00759

6. **A modular two yeast species secretion system for the production and preparative application of unspecific peroxygenases**
   by: Pascal Püllmann, Anja Knorrscheidt, **Judith Münch**, Paul R. Palme, Wolfgang

*shared first authorship

Hoehenwarter, Sylvestre Marillonnet, Miguel Alcalde, Bernhard Westermann, Martin J. Weissenborn
in: *Communications Biology* **2021**, *4,* 562; doi: https://doi.org/10.1038/s42003-021-02076-3

7. **Spastic paraplegia paraplegia-linked phospholipase PAPLA 1 is necessary for development, reproduction, and energy metabolism in Drosophila**
by: Martina Gáliková, Peter Klepsatel, **Judith Münch** and Ronald P. Kühnlein
in: *Sci. Rep.* **2017**, 7, 46516; doi: https://doi.org/10.1038/srep46516

12 Scholarship information

13 Affidavit (Eidesstattliche Versicherung)

Erklärung

Hiermit erkläre ich an Eid statt, dass ich mich mit der vorliegenden wissenschaftlichen Arbeit erstmals um die Erlangung des Doktorgrades bewerbe, die Arbeit selbstständig und ohne fremde Hilfe verfasst, nur die angegebenen Quellen und Hilfsmittel genutzt und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stelle als solche kenntlich gemacht habe.

Halle (Saale), den 16. September 2025

Judith Doan-Münch

References

(1) Münch, J.; Pullmann, P.; Zhang, W.; Weissenborn, M. J., Enzymatic hydroxylations of sp3-carbons, *ACS Catal.*, **2021**, *11* (15), 9168–9203.

(2) You, S.-L.; Xia, J.-B., Palladium-Catalyzed Aryl–Aryl Bond Formation Through Double C–H Activation, *CH Activation*, **2010**, 165–194.

(3) Chen, M. S.; White, M. C., A predictably selective aliphatic C–H oxidation reaction for complex molecule synthesis, *Science*, **2007**, *318* (5851), 783–787.

(4) White, M. C., Adding aliphatic C–H bond oxidations to synthesis, *Science*, **2012**, *335* (6070), 807–809.

(5) White, M. C.; Zhao, J., Aliphatic C–H oxidations for late-stage functionalization, *J. Am. Chem. Soc.*, **2018**, *140* (43), 13988–14009.

(6) Jacoby, C.; Ferlaino, S.; Bezold, D.; Jessen, H.; Müller, M.; Boll, M., ATP-dependent hydroxylation of an unactivated primary carbon with water, *Nat. Commun.*, **2020**, *11* (1), 3906.

(7) Hrycay, E. G.; Bandiera, S. M., Monooxygenase, peroxidase and peroxygenase properties and reaction mechanisms of cytochrome P450 enzymes, In *Monooxygenase, Peroxidase and Peroxygenase Properties and Mechanisms of Cytochrome P450*, **2015**, pp 1–61.

(8) Hausinger, R. P., Fe (II)/α-ketoglutarate-dependent hydroxylases and related enzymes, *Crit. Rev. Biochem. Mol. Biol.*, **2004**, *39* (1), 21–68.

(9) Barry, S. M.; Challis, G. L., Mechanism and catalytic diversity of Rieske non-heme iron-dependent oxygenases, *ACS Catal.*, **2013**, *3* (10), 2362–2370.

(10) Schwarz, G.; Mendel, R. R.; Ribbe, M. W., Molybdenum cofactors, enzymes and pathways, *Nature*, **2009**, *460* (7257), 839–847.

(11) Hille, R.; Hall, J.; Basu, P., The mononuclear molybdenum enzymes, *Chem. Rev.*, **2014**, *114* (7), 3963–4038.

(12) Luo, Y.-R. *Comprehensive handbook of chemical bond energies*; CRC press: 2007.

(13) Ross, M. O.; Rosenzweig, A. C., A tale of two methane monooxygenases, *J. Biol. Inorg. Chem.*, **2017**, *22*, 307–319.

(14) Banerjee, R.; Jones, J. C.; Lipscomb, J. D., Soluble methane monooxygenase, *Annu. Rev. Biochem.*, **2019**, *88*, 409–431.

(15) Ullrich, R.; Hofrichter, M., Enzymatic hydroxylation of aromatic compounds, *Cell. Mol. Life Sci.*, **2007**, *64* (3), 271–293.

(16) Munro, A. W.; McLean, K. J.; Grant, J. L.; Makris, T. M., Structure and function of the cytochrome P450 peroxygenase enzymes, *Bioch. Soc. Trans.*, **2018**, *46* (1), 183–196.

(17) Tang, M.-C.; Fu, C.-Y.; Tang, G.-L., Characterization of SfmD as a Heme peroxidase that catalyzes the regioselective hydroxylation of 3-methyltyrosine to 3-hydroxy-5-methyltyrosine in saframycin A biosynthesis, *J. Biol. Chem.*, **2012**, *287* (7), 5112–5121.

(18) Neusser, D.; Schmidt, H.; Spizèk, J.; Novotnà, J.; Peschke, U.; Kaschabeck, S.; Tichy, P.; Piepersberg, W., The genes lmbB1 and lmbB2 of Streptomyces lincolnensis encode enzymes involved in the conversion of L-tyrosine to propylproline during the biosynthesis of the antibiotic lincomycin A, *Arch. Microbiol.*, **1998**, *169*, 322–332.

(19) Wang, Y.; Lan, D.; Durrani, R.; Hollmann, F., Peroxygenases en route to becoming dream catalysts. What are the opportunities and challenges?, *Curr. Opin. Chem. Biol.*, **2017**, *37*, 1–9.

(20) Ullrich, R.; Nüske, J.; Scheibner, K.; Spantzel, J.; Hofrichter, M., Novel haloperoxidase from the agaric basidiomycete Agrocybe aegerita oxidizes aryl alcohols and aldehydes, *Appl. Environ. Microbiol.*, **2004**, *70* (8), 4575–4581.

(21) Peter, S.; Kinne, M.; Ullrich, R.; Kayser, G.; Hofrichter, M., Epoxidation of linear, branched and cyclic alkenes catalyzed by unspecific peroxygenase, *Enzyme Microb. Technol.*, **2013**, *52* (6-7), 370–376.

(22) Ullrich, R.; Dolge, C.; Kluge, M.; Hofrichter, M., Pyridine as novel substrate for regioselective oxygenation with aromatic peroxygenase from Agrocybe aegerita, *FEBS Lett.*, **2008**, *582* (29), 4100–4106.

(23) Kluge, M. G.; Ullrich, R.; Scheibner, K.; Hofrichter, M., Spectrophotometric assay for detection of aromatic hydroxylation catalyzed by fungal haloperoxidase–peroxygenase, *Appl. Microbiol. Biot.*, **2007**, *75* (6), 1473–1478.

(24) Ramirez-Escudero, M.; Molina-Espeja, P.; Gomez de Santos, P.; Hofrichter, M.; Sanz-Aparicio, J.; Alcalde, M., Structural insights into the substrate promiscuity of a laboratory-evolved peroxygenase, *ACS Chem. Biol.*, **2018**, *13* (12), 3259–3268.

(25) Hofrichter, M.; Kellner, H.; Pecyna, M. J.; Ullrich, R., Fungal unspecific peroxygenases: heme-thiolate proteins that combine peroxidase and cytochrome P450 properties, In *Monooxygenase, Peroxidase and Peroxygenase Properties and Mechanisms of Cytochrome P450*, **2015**, pp 341–368.

(26) Hofrichter, M.; Ullrich, R., Oxidations catalyzed by fungal peroxygenases, *Curr. Opin. Chem. Biol.*, **2014**, *19*, 116–125.

(27) Nevalainen, H., Fungal peroxygenases: A phylogenetically old superfamily of heme enzymes with promiscuity for oxygen transfer reactions, In *Grand Challenges in Fungal Biotechnology*, **2020**, pp 369–403.

(28) Faiza, M.; Huang, S.; Lan, D.; Wang, Y., New insights on unspecific peroxygenases: superfamily reclassification and evolution, *BMC Evol. Biol.*, **2019**, *19*, 1–19.

(29) Ni, Y.; Fernández-Fueyo, E.; Baraibar, A. G.; Ullrich, R.; Hofrichter, M.; Yanase, H.; Alcalde, M.; van Berkel, W. J.; Hollmann, F., Peroxygenase-catalyzed oxyfunctionalization reactions promoted by the complete oxidation of methanol, *Angew. Chem. Int. Ed.*, **2016**, *55* (2), 798–801.

(30) Fernández-Fueyo, E.; Ni, Y.; Baraibar, A. G.; Alcalde, M.; van Langen, L. M.; Hollmann, F., Towards preparative peroxygenase-catalyzed oxyfunctionalization reactions in organic media, *J. Mol.Catal. B: Enzym.*, **2016**, *134*, 347–352.

(31) Rauch, M. C.; Tieves, F.; Paul, C. E.; Arends, I. W.; Alcalde, M.; Hollmann, F., Peroxygenase-catalysed epoxidation of styrene derivatives in Neat Reaction Media, *ChemCatChem*, **2019**, *11* (18), 4519–4523.

(32) Karich, A.; Ullrich, R.; Scheibner, K.; Hofrichter, M., Fungal unspecific peroxygenases oxidize the majority of organic EPA priority pollutants, *Front. Microbiol.*, **2017**, *8*, 288117.

(33) Gröbe, G.; Ullrich, R.; Pecyna, M. J.; Kapturska, D.; Friedrich, S.; Hofrichter, M.; Scheibner, K., High-yield production of aromatic peroxygenase by the agaric fungus Marasmius rotula, *AMB Express*, **2011**, *1*, 1–11.

(34) Anh, D. H.; Ullrich, R.; Benndorf, D.; Svatos, A.; Muck, A.; Hofrichter, M., The coprophilous mushroom Coprinus radians secretes a haloperoxidase that catalyzes aromatic peroxygenation, *Appl. Environ. Microbiol.*, **2007**, *73* (17), 5477–5485.

(35) Babot, E. D.; del Río, J. C.; Kalum, L.; Martínez, A. T.; Gutiérrez, A., Oxyfunctionalization of aliphatic compounds by a recombinant peroxygenase from Coprinopsis cinerea, *Biotechnol.Bioeng.*, **2013**, *110* (9), 2323–2332.

(36) Molina-Espeja, P.; Garcia-Ruiz, E.; Gonzalez-Perez, D.; Ullrich, R.; Hofrichter, M.; Alcalde, M., Directed evolution of unspecific peroxygenase from Agrocybe aegerita, *Appl. Environ. Microbiol.*, **2014**, *80* (11), 3496–3507.

(37) Molina-Espeja, P.; Ma, S.; Mate, D. M.; Ludwig, R.; Alcalde, M., Tandem-yeast expression system for engineering and producing unspecific peroxygenase, *Enzyme Microb. Technol.*, **2015**, *73*, 29–33.

(38) Püllmann, P.; Knorrscheidt, A.; Münch, J.; Palme, P. R.; Hoehenwarter, W.; Marillonnet, S.; Alcalde, M.; Westermann, B.; Weissenborn, M. J., A modular two yeast species secretion

system for the production and preparative application of unspecific peroxygenases, *Commun. Biol.*, **2021**, *4* (1), 562.

(39) Püllmann, P.; Weissenborn, M. J., Improving the heterologous production of fungal peroxygenases through an episomal Pichia pastoris promoter and signal peptide shuffling system, *ACS Synth. Biol.*, **2021**, *10* (6), 1360–1372.

(40) Carro, J.; González-Benjumea, A.; Fernández-Fueyo, E.; Aranda, C.; Guallar, V.; Gutiérrez, A.; Martínez, A. T., Modulating fatty acid epoxidation vs hydroxylation in a fungal peroxygenase, *ACS Catal.*, **2019**, *9* (7), 6234–6242.

(41) Linde, D.; Olmedo, A.; González-Benjumea, A.; Estévez, M.; Renau-Mínguez, C.; Carro, J.; Fernández-Fueyo, E.; Gutiérrez, A.; Martínez, A. T., Two new unspecific peroxygenases from heterologous expression of fungal genes in Escherichia coli, *Appl. Environ. Microbiol.*, **2020**, *86* (7), e02899–19.

(42) Ullrich, R.; Poraj-Kobielska, M.; Scholze, S.; Halbout, C.; Sandvoss, M.; Pecyna, M. J.; Scheibner, K.; Hofrichter, M., Side chain removal from corticosteroids by unspecific peroxygenase, *J. Inorg. Biochem.*, **2018**, *183*, 84–93.

(43) Kiebist, J.; Schmidtke, K.-U.; Zimmermann, J.; Kellner, H.; Ullrich, R.; Zänder, D.; Hofrichter, M.; Scheibner, K., A Peroxygenase from Chaetomium globosum Catalyzes the Selective Oxygenation of Testosterone, *ChemBioChem*, **2017**, *18* (6), 563–569.

(44) Hobisch, M.; Holtmann, D.; de Santos, P. G.; Alcalde, M.; Hollmann, F.; Kara, S., Recent developments in the use of peroxygenases–Exploring their high potential in selective oxyfunctionalisations, *Biotechnol. Adv.*, **2021**, *51*, 107615.

(45) González-Benjumea, A.; Carro, J.; Renau-Mínguez, C.; Linde, D.; Fernández-Fueyo, E.; Gutiérrez, A.; Martínez, A. T., Fatty acid epoxidation by Collariella virescens peroxygenase and heme-channel variants, *Catal. Sci. Technol.*, **2020**, *10* (3), 717–725.

(46) Piontek, K.; Ullrich, R.; Liers, C.; Diederichs, K.; Plattner, D. A.; Hofrichter, M., Crystallization of a 45 kDa peroxygenase/peroxidase from the mushroom Agrocybe aegerita and structure determination by SAD utilizing only the haem iron, *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.*, **2010**, *66* (6), 693–698.

(47) Sundaramoorthy, M.; Terner, J.; Poulos, T. L., The crystal structure of chloroperoxidase: a heme peroxidase–cytochrome P450 functional hybrid, *Structure*, **1995**, *3* (12), 1367–1378.

(48) Wang, X.; Peter, S.; Kinne, M.; Hofrichter, M.; Groves, J. T., Detection and kinetic characterization of a highly reactive heme–thiolate peroxygenase compound I, *J. Am. Chem. Soc.*, **2012**, *134* (31), 12897–12900.

(49) Wang, X.; Wang, X., Driving Force for Oxygen Atom Transfer by Heme-Thiolate Enzymes, In *A Novel Heme-Thiolate Peroxygenase AaeAPO and Its Implications for CH Activation Chemistry*, **2016**, pp 75–89.

(50) Peter, S.; Kinne, M.; Wang, X.; Ullrich, R.; Kayser, G.; Groves, J. T.; Hofrichter, M., Selective hydroxylation of alkanes by an extracellular fungal peroxygenase, *FEBS J.*, **2011**, *278* (19), 3667–3675.

(51) Olmedo, A.; Aranda, C.; Del Río, J. C.; Kiebist, J.; Scheibner, K.; Martínez, A. T.; Gutiérrez, A., From alkanes to carboxylic acids: Terminal oxygenation by a fungal peroxygenase, *Angew. Chem. Int. Ed.*, **2016**, *128* (40), 12436–12439.

(52) Knorrscheidt, A.; Soler, J.; Hünecke, N.; Püllmann, P.; Garcia-Borràs, M.; Weissenborn, M. J., Accessing chemo-and regioselective benzylic and aromatic oxidations by protein engineering of an unspecific peroxygenase, *ACS Catal.*, **2021**, *11* (12), 7327–7338.

(53) Mitsuteru, N.; Yoshio, O., Improved synthesis of 16α-hydroxylated androgens: Intermediates of estriol formation in pregnancy, *Steroids*, **1978**, *32* (4), 519–527.

(54) Agematu, H.; Matsumoto, N.; Fujii, Y.; Kabumoto, H.; Doi, S.; Machida, K.; Ishikawa, J.; Arisawa, A., Hydroxylation of testosterone by bacterial cytochromes P450 using the Escherichia coli expression system, *Biosci., Biotechnol., Biochem.*, **2006**, *70* (1), 307–311.

(55) Murphy, R. C.; Johnson, K. M., Cholesterol, reactive oxygen species, and the formation of biologically active mediators, *J. Biol.Chem.*, **2008**, *283* (23), 15521–15525.

(56) Schäfer, B. *Naturstoffe der chemischen Industrie*; Springer-Verlag: 2020.

(57) Fritts, C.; Waldroup, P., Effect of source and level of vitamin D on live performance and bone development in growing broilers, *J. Appl. Poult. Res.*, **2003**, *12* (1), 45–52.

(58) Käppeli, S.; Fröhlich, E.; Gebhardt-Henrich, S. G.; Pfulg, A.; Schäublin, H.; Zweifel, R.; Wiedmer, H.; Stoffel, M., Effects of dietary supplementation with synthetic vitamin D3 and 25-hydroxycholecalciferol on blood calcium and phosphate levels and performance in laying hens, *Eur. Poult. Sci.*, **2011**, *75* (3), 179–184.

(59) Michalczuk, M.; Pietrzak, D.; Niemiec, J.; Mroczek, J., Effectiveness of vitamin D3 and calcidiol (25-OH-D3) application in feeding broiler chickens-production performance and meat quality, *J. Food Nutr. Sci.*, **2010**, *60* (2).

(60) Lucas, F.; Babot, E. D.; Cañellas, M.; del Río, J. C.; Kalum, L.; Ullrich, R.; Hofrichter, M.; Guallar, V.; Martínez, A. T.; Gutiérrez, A., Molecular determinants for selective C 25-hydroxylation of vitamins D 2 and D 3 by fungal peroxygenases, *Catal. Sci. Technol.*, **2016**, *6* (1), 288–295.

(61) Babot, E. D.; del Río, J. C.; Kalum, L.; Martínez, A. T.; Gutiérrez, A., Regioselective Hydroxylation in the Production of 25-Hydroxyvitamin D by Coprinopsis cinerea Peroxygenase, *ChemCatChem*, **2015**, *7* (2), 283–290.

(62) Martin-Diaz, J.; Paret, C.; García-Ruiz, E.; Molina-Espeja, P.; Alcalde, M., Shuffling the neutral drift of unspecific peroxygenase in Saccharomyces cerevisiae, *Appl. Environ. Microbiol.*, **2018**, *84* (15), e00808–18.

(63) Freakley, S. J.; Kochius, S.; van Marwijk, J.; Fenner, C.; Lewis, R. J.; Baldenius, K.; Marais, S. S.; Opperman, D. J.; Harrison, S. T.; Alcalde, M., et al., A chemo-enzymatic oxidation cascade to activate C–H bonds with in situ generated H2O2, *Nat. Commun.*, **2019**, *10* (1), 4178.

(64) Pesic, M.; Willot, S. J.-P.; Fernández-Fueyo, E.; Tieves, F.; Alcalde, M.; Hollmann, F., Multienzymatic in situ hydrogen peroxide generation cascade for peroxygenase-catalysed oxyfunctionalisation reactions, *Z. fur Naturforsch. - C J. Biosci.*, **2019**, *74* (3-4), 101–104.

(65) Tieves, F.; Willot, S. J.-P.; van Schie, M. M. C. H.; Rauch, M. C. R.; Younes, S. H. H.; Zhang, W.; Dong, J.; Gómez de Santos, P.; Robbins, J. M.; Bommarius, B., et al., Formate Oxidase (FOx) from Aspergillus oryzae: One Catalyst Enables Diverse H2O2-Dependent Biocatalytic Oxidation Reactions, *Angew. Chem., Int. Ed.*, **2019**, *58* (23), 7873–7877.

(66) Willot, S. J.-P.; Hoang, M. D.; Paul, C. E.; Alcalde, M.; Arends, I. W.; Bommarius, A. S.; Bommarius, B.; Hollmann, F., FOx News: Towards methanol-driven biocatalytic oxyfunctionalisation reactions, *ChemCatChem*, **2020**, *12* (10), 2713–2716.

(67) Willot, S. J.-P.; Fernández-Fueyo, E.; Tieves, F.; Pesic, M.; Alcalde, M.; Arends, I. W.; Park, C. B.; Hollmann, F., Expanding the spectrum of light-driven peroxygenase reactions, *ACS Catal.*, **2018**, *9* (2), 890–894.

(68) Zhang, W.; Liu, H.; Van Schie, M. M.; Hagedoorn, P.-L.; Alcalde, M.; Denkova, A. G.; Djanashvili, K.; Hollmann, F., Nuclear waste and biocatalysis: a sustainable liaison?, *ACS Catal.*, **2020**, *10* (23), 14195–14200.

(69) Churakova, E.; Kluge, M.; Ullrich, R.; Arends, I.; Hofrichter, M.; Hollmann, F., Specific photobiocatalytic oxyfunctionalization reactions, *Angew. Chem., Int. Ed.*, **2011**, *45* (123), 10904–10907.

(70) Knorrscheidt, A.; Püllmann, P.; Schell, E.; Homann, D.; Freier, E.; Weissenborn, M. J., Identification of novel unspecific peroxygenase chimeras and unusual YfeX axial heme ligand by a versatile high-throughput GC-MS approach, *ChemCatChem*, **2020**, *12* (19), 4788–4795.

(71) Knorrscheidt, A.; Soler, J.; Hünecke, N.; Püllmann, P.; Garcia-Borràs, M.; Weissenborn, M. J., Simultaneous screening of multiple substrates with an unspecific peroxygenase en-

abled modified alkane and alkene oxyfunctionalisations, *Catal. Sci. Technol.*, **2021**, *11* (18), 6058–6064.

(72) Monterrey, D. T.; Menés-Rubio, A.; Keser, M.; Gonzalez-Perez, D.; Alcalde, M., Unspecific peroxygenases: The pot of gold at the end of the oxyfunctionalization rainbow?, *Curr. Opin. Green Sustain. Chem.*, **2023**, *41*, 100786.

(73) Beltrán-Nogal, A.; Sánchez-Moreno, I.; Méndez-Sánchez, D.; de Santos, P. G.; Hollmann, F.; Alcalde, M., Surfing the wave of oxyfunctionalization chemistry by engineering fungal unspecific peroxygenases, *Curr. Opin. Struc. Biol.*, **2022**, *73*, 102342.

(74) Melling, B.; Mielke, T.; Whitwood, A. C.; O'Riordan, T. J.; Mulholland, N.; Cartwright, J.; Unsworth, W. P.; Grogan, G., Complementary specificity of unspecific peroxygenases enables access to diverse products from terpene oxygenation, *Chem Catal.*, **2024**, *4* (2).

(75) Ebrecht, A. C.; Mofokeng, T. M.; Hollmann, F.; Smit, M. S.; Opperman, D. J., Lactones from unspecific peroxygenase-catalyzed in-chain hydroxylation of saturated fatty acids, *Org. Lett.*, **2023**, *25* (27), 4990–4995.

(76) Han, X.; Chen, F.; Li, H.; Ge, R.; Shen, Q.; Duan, P.; Sheng, X.; Zhang, W., Reaction engineering blocks ether cleavage for synthesizing chiral cyclic hemiacetals catalyzed by unspecific peroxygenase, *Nat. Commun.*, **2024**, *15* (1), 1235.

(77) Li, W.; Davis, D. L.; Speina, K. J.; Monroe, C. B.; Moncrieffe, A. S.; Cao, Y.; Chen, C.-C.; Groves, J. T., Unspecific Peroxygenase Catalyzes Selective Remote-Site Functionalizations, *ChemCatChem*, **2025**, *17* (2), e202401285.

(78) Schmitz, F.; Koschorreck, K.; Hollmann, F.; Urlacher, V. B., Aromatic hydroxylation of substituted benzenes by an unspecific peroxygenase from Aspergillus brasiliensis, *React. Chem. Eng.*, **2023**, *8* (9), 2177–2186.

(79) Wibowo, M.; Ding, L., Chemistry and biology of natural azoxy compounds, *J. Nat. Prod.*, **2020**, *83* (11), 3482–3491.

(80) Li, H.; Huang, Y.; Chen, F.; Zeng, Z.; Hollmann, F.; Wu, X.; Zhang, X.; Duan, P.; Su, H.; Shi, J., et al., Unspecific peroxygenase enabled formation of azoxy compounds, *Nat. Commun.*, **2024**, *15* (1), 8312.

(81) Li, H.; Shen, Q.; Zhou, X.; Duan, P.; Hollmann, F.; Huang, Y.; Zhang, W., Peroxygenase-catalysed sulfoxidations in non-aqueous media, *ChemSusChem*, **2024**, *17* (6), e202301321.

(82) Estabrook, R. W., A passion for P450s (remembrances of the early history of research on cytochrome P450), *Drug Metab. Dispos.*, **2003**, *31* (12), 1461–1473.

(83) Domanski, T.; Halpert, J., Analysis of mammalian cytochrome P450 structure and function by site-directed mutagenesis, *Curr. Drug Metab.*, **2001**, *2* (2), 117–137.

(84) Nelson, D. R.; Ming, R.; Alam, M.; Schuler, M. A., Comparison of cytochrome P450 genes from six plant genomes, *Trop. Plant Biol.*, **2008**, *1*, 216–235.

(85) Chen, W.; Lee, M.-K.; Jefcoate, C.; Kim, S.-C.; Chen, F.; Yu, J.-H., Fungal cytochrome p450 monooxygenases: their distribution, structure, functions, family expansion, and evolutionary origin, *Genome Biol. Evol.*, **2014**, *6* (7), 1620–1634.

(86) Nelson, D. R., Cytochrome P450 diversity in the tree of life, *Biochim. Biophys. Acta - Proteins Proteom.*, **2018**, *1866* (1), 141–154.

(87) Schlichting, I.; Berendzen, J.; Chu, K.; Stock, A. M.; Maves, S. A.; Benson, D. E.; Sweet, R. M.; Ringe, D.; Petsko, G. A.; Sligar, S. G., The catalytic pathway of cytochrome P450cam at atomic resolution, *Science*, **2000**, *287* (5458), 1615–1622.

(88) Vermilion, J.; Ballou, D.; Massey, V.; Coon, M., Separate roles for FMN and FAD in catalysis by liver microsomal NADPH-cytochrome P-450 reductase. *J. Biol. Chem.*, **1981**, *256* (1), 266–277.

(89) Oprian, D. D.; Coon, M., Oxidation-reduction states of FMN and FAD in NADPH-cytochrome P-450 reductase during reduction by NADPH. *J. Biol. Chem.*, **1982**, *257* (15), 8935–8944.

(90) Schiffler, B.; Bernhardt, R., Bacterial (CYP101) and mitochondrial P450 systems—how comparable are they?, *Biochem. Bioph. Res. Co.*, **2003**, *312* (1), 223–228.

(91) Hanukoglu, I., Electron transfer proteins of cytochrome P450 systems, In *Adv. Mol. Cell Biol.* **1996**, pp 29–56.

(92) Harris, D.; Loew, G., Determinants of the spin state of the resting state of cytochrome P450cam, *J. Am. Chem. Soc.*, **1993**, *115* (19), 8775–8779.

(93) Fisher, M. T.; Sligar, S. G., Control of heme protein redox potential and reduction rate: linear free energy relation between potential and ferric spin state equilibrium, *J. Am. Chem. Soc.*, **1985**, *107* (17), 5018–5019.

(94) Meunier, B.; Bernadou, J., Active iron-oxo and iron-peroxo species in cytochromes P450 and peroxidases; oxo-hydroxo tautomerism with water-soluble metalloporphyrins, *Metal-Oxo and Metal-Peroxo Species in Catalytic Oxidations*, **2001**, 1–35.

(95) Kovacs, J. A., Synthetic analogues of cysteinate-ligated non-heme iron and non-corrinoid cobalt enzymes, *Chem. Rev.*, **2004**, *104* (2), 825–848.

(96) Surawatanawong, P.; Tye, J. W.; Hall, M. B., Density functional theory applied to a difference in pathways taken by the enzymes cytochrome P450 and superoxide reductase: spin states of ferric hydroperoxo intermediates and hydrogen bonds from water, *Inorg. Chem.*, **2010**, *49* (1), 188–198.

(97) Huang, X.; Groves, J. T., Oxygen activation and radical transformations in heme proteins and metalloporphyrins, *Chem. Rev.*, **2017**, *118* (5), 2491–2553.

(98) Kühnel, K.; Derat, E.; Terner, J.; Shaik, S.; Schlichting, I., Structure and quantum chemical characterization of chloroperoxidase compound 0, a common reaction intermediate of diverse heme enzymes, *Proc. Natl. Acad. Sci.*, **2007**, *104* (1), 99–104.

(99) Newcomb, M.; Zhang, R.; Chandrasena, R. E. P.; Halgrimson, J. A.; Horner, J. H.; Makris, T. M.; Sligar, S. G., Cytochrome P450 compound I, *J. Am. Chem. Soc.*, **2006**, *128* (14), 4580–4581.

(100) Rittle, J.; Green, M. T., Cytochrome P450 compound I: capture, characterization, and CH bond activation kinetics, *Science*, **2010**, *330* (6006), 933–937.

(101) Woggon, W.-D., Cytochrome P450: Significance, reaction mechanisms and active site analogues, In *Bioorganic Chemistry: Models and Applications*, **2005**, pp 39–96.

(102) Prasad, S.; Mitra, S., Substrate modulates compound I formation in peroxide shunt pathway of Pseudomonas putida cytochrome P450cam, *Biochem. Bioph. Res. Co.*, **2004**, *314* (2), 610–614.

(103) Urlacher, V. B.; Girhard, M., Cytochrome P450 monooxygenases in biotechnology and synthetic biology, *Trends Biotechnol.*, **2019**, *37* (8), 882–897.

(104) Ortiz de Montellano, P. R., Hydrocarbon hydroxylation by cytochrome P450 enzymes, *Chem. Rev.*, **2010**, *110* (2), 932–948.

(105) Zanger, U. M.; Schwab, M., Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation, *Pharmacol. Therapeut.*, **2013**, *138* (1), 103–141.

(106) Guengerich, F. P.; Munro, A. W., Unusual cytochrome P450 enzymes and reactions, *J. Biol. Chem.*, **2013**, *288* (24), 17065–17073.

(107) Weissenborn, M. J.; Koenigs, R. M., Iron-porphyrin catalyzed carbene transfer reactions–an evolution from biomimetic catalysis towards chemistry-inspired non-natural reactivities of enzymes, *ChemCatChem*, **2020**, *12* (8), 2171–2179.

(108) Brandenberg, O. F.; Fasan, R.; Arnold, F. H., Exploiting and engineering hemoproteins for abiological carbene and nitrene transfer reactions, *Curr. Opin. Biotech.*, **2017**, *47*, 102–111.

(109) Coelho, P. S.; Brustad, E. M.; Kannan, A.; Arnold, F. H., Olefin cyclopropanation via carbene transfer catalyzed by engineered cytochrome P450 enzymes, *Science*, **2013**, *339* (6117), 307–310.

(110) Chen, M. M.; Coelho, P. S.; Arnold, F. H., Utilizing Terminal Oxidants to Achieve P450-Catalyzed Oxidation of Methane, *Adv. Synth. Catal.*, **2012**, *354* (6), 964–968.

(111) Renault, H.; Bassard, J.-E.; Hamberger, B.; Werck-Reichhart, D., Cytochrome P450-mediated metabolic engineering: current progress and future challenges, *Curr. Opin. Plant Biol.*, **2014**, *19*, 27–34.

(112) Salazar, O.; Cirino, P. C.; Arnold, F. H., Thermostabilization of a cytochrome P450 peroxygenase, *ChemBioChem*, **2003**, *4* (9), 891–893.

(113) Seng Wong, T.; Arnold, F. H.; Schwaneberg, U., Laboratory evolution of cytochrome P450 BM-3 monooxygenase for organic cosolvents, *Biotechnol. Bioeng.*, **2004**, *85* (3), 351–358.

(114) Zhang, X.; Peng, Y.; Zhao, J.; Li, Q.; Yu, X.; Acevedo-Rocha, C. G.; Li, A., Bacterial cytochrome P450-catalyzed regio-and stereoselective steroid hydroxylation enabled by directed evolution and rational design, *Bioresour. Bioprocess.*, **2020**, *7*, 1–18.

(115) Lundemo, M.; Notonier, S.; Striedner, G.; Hauer, B.; Woodley, J., Process limitations of a whole-cell P450 catalyzed reaction using a CYP153A-CPR fusion construct expressed in Escherichia coli, *Appl. Microbiol. Biot.*, **2016**, *100*, 1197–1208.

(116) Chen, J.; Kong, F.; Ma, N.; Zhao, P.; Liu, C.; Wang, X.; Cong, Z., Peroxide-driven hydroxylation of small alkanes catalyzed by an artificial P450BM3 peroxygenase system, *ACS Catal.*, **2019**, *9* (8), 7350–7355.

(117) Siriphongphaew, A.; Pisnupong, P.; Wongkongkatep, J.; Inprakhon, P.; Vangnai, A. S.; Honda, K.; Ohtake, H.; Kato, J.; Ogawa, J.; Shimizu, S., et al., Development of a whole-cell biocatalyst co-expressing P450 monooxygenase and glucose dehydrogenase for synthesis of epoxyhexane, *Appl. Microbiol. Biot.*, **2012**, *95*, 357–367.

(118) Sung, C.; Jung, E.; Choi, K.-Y.; Bae, J.-h.; Kim, M.; Kim, J.; Kim, E.-J.; Kim, P. I.; Kim, B.-G., The production of ω-hydroxy palmitic acid using fatty acid metabolism and cofactor optimization in Escherichia coli, *Appl. Microbiol. Biot.*, **2015**, *99*, 6667–6676.

(119) Ryan, J. D.; Fish, R. H.; Clark, D. S., Engineering cytochrome P450 enzymes for improved activity towards biomimetic 1, 4-NADH cofactors, *ChemBioChem*, **2008**, *9* (16), 2579–2582.

(120) Löw, S. A.; Löw, I. M.; Weissenborn, M. J.; Hauer, B., Enhanced ene-reductase activity through alteration of artificial nicotinamide cofactor substituents, *ChemCatChem*, **2016**, *8* (5), 911–915.

(121) Sadeghi, S. J.; Fantuzzi, A.; Gilardi, G., Breakthrough in P450 bioelectrochemistry and future perspectives, *Biochim. Biophys. Acta - Proteins Proteom.*, **2011**, *1814* (1), 237–248.

(122) Llaudet, E. C.; Darimont, D.; Samba, R.; Matiychyn, I.; Stelzle, M.; Weissenborn, M. J.; Hauer, B., Expanding an Efficient, Electrically Driven and CNT-Tagged P450 System into the Third Dimension: A Nanowired CNT-Containing and Enzyme-Stabilising 3 D Sol–Gel Electrode, *ChemBioChem*, **2016**, *17* (14), 1367–1373.

(123) Jensen, K.; Johnston, J. B.; de Montellano, P. R. O.; Møller, B. L., Photosystem I from plants as a bacterial cytochrome P450 surrogate electron donor: terminal hydroxylation of branched hydrocarbon chains, *Biotechnol. Lett.*, **2012**, *34*, 239–245.

(124) Nielsen, A. Z.; Ziersen, B.; Jensen, K.; Lassen, L. M.; Olsen, C. E.; Møller, B. L.; Jensen, P. E., Redirecting photosynthetic reducing power toward bioactive natural product synthesis, *ACS Synth. Biol.*, **2013**, *2* (6), 308–315.

(125) Hoffmann, S. M.; Weissenborn, M. J.; Gricman, Ł.; Notonier, S.; Pleiss, J.; Hauer, B., The impact of linker length on P450 fusion constructs: activity, stability and coupling, *ChemCatChem*, **2016**, *8* (8), 1591–1597.

(126) Munro, A. W.; Girvan, H. M.; McLean, K. J., Cytochrome P450–redox partner fusion enzymes, *Biochim. Biophys. Acta - Gen. Subj.*, **2007**, *1770* (3), 345–359.

(127) Hardwick, J. P., Cytochrome P450 omega hydroxylase (CYP4) function in fatty acid metabolism and metabolic diseases, *Biochem. Pharmacol.*, **2008**, *75* (12), 2263–2275.

(128) Hammerer, L.; Winkler, C. K.; Kroutil, W., Regioselective biocatalytic hydroxylation of fatty acids by cytochrome P450s, *Catal. Lett.*, **2018**, *148*, 787–812.

(129) Funhoff, E. G.; Bauer, U.; García-Rubio, I.; Witholt, B.; van Beilen, J. B., CYP153A6, a soluble P450 oxygenase catalyzing terminal-alkane hydroxylation, *J. Bacteriol.*, **2006**, *188* (14), 5220–5227.

(130) Hoffmann, S. M.; Danesh-Azari, H.-R.; Spandolf, C.; Weissenborn, M. J.; Grogan, G.; Hauer, B., Structure-guided redesign of CYP153AM. aq for the improved terminal hydroxylation of fatty acids, *ChemCatChem*, **2016**, *8* (20), 3234–3239.

(131) Schörken, U.; Kempers, P., Lipid biotechnology: Industrially relevant production processes, *Eur. J. Lipid Sci. Tech.*, **2009**, *111* (7), 627–645.

(132) Mills, B. M.; Talati, M. M.; Winter, G. A. Bio-Dodecanedioic Acid (DDDA) Production, Senior Design Reports (CBE), University of Pennsylvania, Accessed: [2020-05-18], 2018.

(133) Weiss, A., Selective microbial oxidations in industry: Oxidations of alkanes, fatty acids, heterocyclic compounds, aromatic compounds and glycerol using native or recombinant microorganisms, In *Modern Biooxidation: Enzymes, Reactions and Applications*, **2007**, pp 193–209.

(134) Sugimoto, H.; Shinkyo, R.; Hayashi, K.; Yoneda, S.; Yamada, M.; Kamakura, M.; Ikushiro, S.-i.; Shiro, Y.; Sakaki, T., Crystal structure of CYP105A1 (P450SU-1) in complex with 1α, 25-dihydroxyvitamin D3, *Biochemistry*, **2008**, *47* (13), 4017–4027.

(135) Hayashi, K.; Sugimoto, H.; Shinkyo, R.; Yamada, M.; Ikeda, S.; Ikushiro, S.; Kamakura, M.; Shiro, Y.; Sakaki, T., Structure-based design of a highly active vitamin D hydroxylase from Streptomyces griseolus CYP105A1, *Biochemistry*, **2008**, *47* (46), 11964–11972.

(136) Hayashi, K.; Yasuda, K.; Sugimoto, H.; Ikushiro, S.; Kamakura, M.; Kittaka, A.; Horst, R. L.; Chen, T. C.; Ohta, M.; Shiro, Y., et al., Three-step hydroxylation of vitamin D3 by a genetically engineered CYP105A1: enzymes and catalysis, *FEBS J.*, **2010**, *277* (19), 3999–4009.

(137) Zhu, G.-D.; Okamura, W. H., Synthesis of vitamin D (calciferol), *Chem. Rev.*, **1995**, *95* (6), 1877–1952.

(138) Peters, M. W.; Meinhold, P.; Glieder, A.; Arnold, F. H., Regio-and enantioselective alkane hydroxylation with engineered cytochromes P450 BM-3, *J. Am. Chem. Soc.*, **2003**, *125* (44), 13442–13450.

(139) Meinhold, P.; Peters, M. W.; Hartwick, A.; Hernandez, A. R.; Arnold, F. H., Engineering cytochrome P450 BM3 for terminal alkane hydroxylation, *Adv. Synth. Catal.*, **2006**, *348* (6), 763–772.

(140) Bell, S. G.; Orton, E.; Boyd, H.; Stevenson, J.-A.; Riddle, A.; Campbell, S.; Wong, L.-L., Engineering cytochrome P450cam into an alkane hydroxylase, *Dalton Trans.*, **2003** (, 11), 2133–2140.

(141) Xu, F.; Bell, S. G.; Lednik, J.; Insley, A.; Rao, Z.; Wong, L.-L., The heme monooxygenase cytochrome P450cam can be engineered to oxidize ethane to ethanol, *Angew. Chem.*, **2005**, *117* (26), 4097–4100.

(142) Meinhold, P.; Peters, M. W.; Chen, M. M.; Takahashi, K.; Arnold, F. H., Direct conversion of ethane to ethanol by engineered cytochrome P450 BM3, *ChemBioChem*, **2005**, *6* (10), 1765–1768.

(143) Fasan, R.; Chen, M. M.; Crook, N. C.; Arnold, F. H., Engineered alkane-hydroxylating cytochrome P450BM3 exhibiting nativelike catalytic properties, *Angew. Chem. Int. Ed.*, **2007**, *46* (44), 8414–8418.

(144) Lewis, J. C.; Mantovani, S. M.; Fu, Y.; Snow, C. D.; Komor, R. S.; Wong, C.-H.; Arnold, F. H., Combinatorial alanine substitution enables rapid optimization of cytochrome P450BM3 for selective hydroxylation of large substrates, *ChemBioChem*, **2010**, *11* (18), 2502.

(145) Ruijssenaars, H. J.; Sperling, E. M.; Wiegerinck, P. H.; Brands, F. T.; Wery, J.; de Bont, J. A., Testosterone 15β-hydroxylation by solvent tolerant Pseudomonas putida S12, *J. Biotechnol.*, **2007**, *131* (2), 205–208.

(146) Lisurek, M.; Simgen, B.; Antes, I.; Bernhardt, R., Theoretical and experimental evaluation of a CYP106A2 low homology model and production of mutants with changed activity and selectivity of hydroxylation, *ChemBioChem*, **2008**, *9* (9), 1439–1449.

(147) Harlow, G. R.; Halpert, J. R., Alanine-scanning mutagenesis of a putative substrate recognition site in human cytochrome P450 3A4: role of residues 210 and 211 in flavonoid activation and substrate specificity, *J. Biol. Chem.*, **1997**, *272* (9), 5396–5402.

(148) Khatri, Y.; Jóźwik, I. K.; Ringle, M.; Ionescu, I. A.; Litzenburger, M.; Hutter, M. C.; Thunnissen, A.-M. W.; Bernhardt, R., Structure-based engineering of steroidogenic CYP260A1 for stereo-and regioselective hydroxylation of progesterone, *ACS Chem. Biol.*, **2018**, *13* (4), 1021–1028.

(149) Kille, S.; Zilly, F. E.; Acevedo, J. P.; Reetz, M. T., Regio-and stereoselectivity of P450-catalysed hydroxylation of steroids controlled by laboratory evolution, *Nat. Chem.*, **2011**, *3* (9), 738–743.

(150) Acevedo-Rocha, C. G.; Gamble, C. G.; Lonsdale, R.; Li, A.; Nett, N.; Hoebenreich, S.; Lingnau, J. B.; Wirtz, C.; Fares, C.; Hinrichs, H., et al., P450-catalyzed regio-and diastereoselective steroid hydroxylation: efficient directed evolution enabled by mutability landscaping, *ACS Catal.*, **2018**, *8* (4), 3395–3410.

(151) Whitehouse, C. J.; Bell, S. G.; Wong, L.-L., P450 BM3 (CYP102A1): connecting the dots, *Chem. Soc. Rev.*, **2012**, *41* (3), 1218–1260.

(152) Paddon, C. J.; Westfall, P. J.; Pitera, D. J.; Benjamin, K.; Fisher, K.; McPhee, D.; Leavell, M.; Tai, A.; Main, A.; Eng, D., et al., High-level semi-synthetic production of the potent antimalarial artemisinin, *Nature*, **2013**, *496* (7446), 528–532.

(153) Zhang, K.; Shafer, B. M.; Demars, M. D.; Stern, H. A.; Fasan, R., Controlled oxidation of remote sp3 C–H bonds in artemisinin via P450 catalysts with fine-tuned regio-and stereoselectivity, *J. Am. Chem. Soc.*, **2012**, *134* (45), 18695–18704.

(154) Reisky, L.; Büchsenschütz, H. C.; Engel, J.; Song, T.; Schweder, T.; Hehemann, J.-H.; Bornscheuer, U. T., Oxidative demethylation of algal carbohydrates by cytochrome P450 monooxygenases, *Nat. Chem. Biol.*, **2018**, *14* (4), 342–344.

(155) Hehemann, J.-H.; Kelly, A. G.; Pudlo, N. A.; Martens, E. C.; Boraston, A. B., Bacteria of the human gut microbiome catabolize red seaweed glycans with carbohydrate-active enzyme updates from extrinsic microbes, *Proc. Natl. Acad. Sci.*, **2012**, *109* (48), 19786–19791.

(156) Wargacki, A. J.; Leonard, E.; Win, M. N.; Regitsky, D. D.; Santos, C. N. S.; Kim, P. B.; Cooper, S. R.; Raisner, R. M.; Herman, A.; Sivitz, A. B., et al., An engineered microbial platform for direct biofuel production from brown macroalgae, *Science*, **2012**, *335* (6066), 308–313.

(157) Çelik, A.; Roberts, G. A.; White, J. H.; Chapman, S. K.; Turner, N. J.; Flitsch, S. L., Probing the substrate specificity of the catalytically self-sufficient cytochrome P450 RhF from a Rhodococcus sp. *Chem. Commun.*, **2006** (, 43), 4492–4494.

(158) Lewis, J. C.; Bastian, S.; Bennett, C. S.; Fu, Y.; Mitsuda, Y.; Chen, M. M.; Greenberg, W. A.; Wong, C.-H.; Arnold, F. H., Chemoenzymatic elaboration of monosaccharides using engineered cytochrome P450BM3 demethylases, *Proc. Natl. Acad. Sci.*, **2009**, *106* (39), 16550–16555.

(159) Trudeau, D. L.; Tawfik, D. S., Protein engineers turned evolutionists—the quest for the optimal starting point, *Curr. Opin. Biotech.*, **2019**, *60*, 46–52.

(160) Kamble, A.; Srinivasan, S.; Singh, H., In-silico bioprospecting: finding better enzymes, *Mol. Biotechnol.*, **2019**, *61*, 53–59.

(161) Robinson, S. L.; Piel, J.; Sunagawa, S., A roadmap for metagenomic enzyme discovery, *Nat. Prod. Rep.*, **2021**, *38* (11), 1994–2023.

(162) Zhu, B.; Wang, D.; Wei, N., Enzyme discovery and engineering for sustainable plastic recycling, *Trends Biotechnol.*, **2022**, *40* (1), 22–37.

(163) Buller, R.; Lutz, S.; Kazlauskas, R.; Snajdrova, R.; Moore, J.; Bornscheuer, U., From nature to industry: Harnessing enzymes for biocatalysis, *Science*, **2023**, *382* (6673), eadh8615.

(164) Kries, H.; Blomberg, R.; Hilvert, D., De novo enzymes by computational design, *Curr. Opin. Chem. Biol.*, **2013**, *17* (2), 221–228.

(165) Kortemme, T., De novo protein design—From new structures to programmable functions, *Cell*, **2024**, *187* (3), 526–544.

(166) Bunzel, H. A.; Anderson, J. R.; Mulholland, A. J., Designing better enzymes: Insights from directed evolution, *Curr. Opin. Struc. Biol.*, **2021**, *67*, 212–218.

(167) Yeh, A. H.-W.; Norn, C.; Kipnis, Y.; Tischer, D.; Pellock, S. J.; Evans, D.; Ma, P.; Lee, G. R.; Zhang, J. Z.; Anishchenko, I., et al., De novo design of luciferases using deep learning, *Nature*, **2023**, *614* (7949), 774–780.

(168) Watson, J. L.; Juergens, D.; Bennett, N. R.; Trippe, B. L.; Yim, J.; Eisenach, H. E.; Ahern, W.; Borst, A. J.; Ragotte, R. J.; Milles, L. F., et al., De novo design of protein structure and function with RFdiffusion, *Nature*, **2023**, *620* (7976), 1089–1100.

(169) Pan, X.; Kortemme, T., Recent advances in de novo protein design: Principles, methods, and applications, *J. Biol. Chem.*, **2021**, *296*.

(170) Reetz, M. T.; Kahakeaw, D.; Lohmer, R., Addressing the numbers problem in directed evolution, *ChemBioChem*, **2008**, *9* (11), 1797–1804.

(171) Mair, P.; Gielen, F.; Hollfelder, F., Exploring sequence space in search of functional enzymes using microfluidic droplets, *Curr. Opin. Chem. Biol.*, **2017**, *37*, 137–144.

(172) Körfer, G.; Pitzler, C.; Vojcic, L.; Martinez, R.; Schwaneberg, U., In vitro flow cytometry-based screening platform for cellulase engineering, *Sci. Rep.*, **2016**, *6* (1), 26128.

(173) Markel, U.; Essani, K. D.; Besirlioglu, V.; Schiffels, J.; Streit, W. R.; Schwaneberg, U., Advances in ultrahigh-throughput screening for directed enzyme evolution, *Chem. Soc. Rev.*, **2020**, *49* (1), 233–262.

(174) Badran, A. H.; Liu, D. R., In vivo continuous directed evolution, *Curr. Opin. Chem. Biol.*, **2015**, *24*, 1–10.

(175) d'Oelsnitz, S.; Ellington, A., Continuous directed evolution for strain and protein engineering, *Curr. Opin. Biotech.*, **2018**, *53*, 158–163.

(176) Martin-Diaz, J.; Molina-Espeja, P.; Hofrichter, M.; Hollmann, F.; Alcalde, M., Directed evolution of unspecific peroxygenase in organic solvents, *Biotechnol. Bioeng.*, **2021**, *118* (8), 3002–3014.

(177) Zhang, R. K.; Chen, K.; Huang, X.; Wohlschlager, L.; Renata, H.; Arnold, F. H., Enzymatic assembly of carbon–carbon bonds via iron-catalysed sp 3 C–H functionalization, *Nature*, **2019**, *565* (7737), 67–72.

(178) Sun, Z.; Lonsdale, R.; Wu, L.; Li, G.; Li, A.; Wang, J.; Zhou, J.; Reetz, M. T., Structure-guided triple-code saturation mutagenesis: efficient tuning of the stereoselectivity of an epoxide hydrolase, *ACS Catal.*, **2016**, *6* (3), 1590–1597.

(179) De Visser, J. A. G.; Krug, J., Empirical fitness landscapes and the predictability of evolution, *Nat. Rev. Genet.*, **2014**, *15* (7), 480–490.

(180) Hartman, E. C.; Tullman-Ercek, D., Learning from protein fitness landscapes: a review of mutability, epistasis, and evolution, *Curr. Opin. Syst. Biol.*, **2019**, *14*, 25–31.

(181) Wittmund, M.; Cadet, F.; Davari, M. D., Learning epistasis and residue coevolution patterns: Current trends and future perspectives for advancing enzyme engineering, *ACS Catal.*, **2022**, *12* (22), 14243–14263.

(182) Reetz, M. T.; Bocola, M.; Carballeira, J. D.; Zha, D.; Vogel, A., Expanding the range of substrate acceptance of enzymes: combinatorial active-site saturation test, *Angew. Chem. Int. Ed.*, **2005**, *44* (27), 4192–4196.

(183) Reetz, M. T.; Carballeira, J. D., Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes, *Nat. Protoc.*, **2007**, *2* (4), 891–903.

(184) Reetz, M. T., The importance of additive and non-additive mutational effects in protein engineering. *Angew. Chem. Int. Ed.*, **2013**, *52* (10).

(185) Hutchison, C. A.; Phillips, S.; Edgell, M. H.; Gillam, S.; Jahnke, P.; Smith, M., Mutagenesis at a specific position in a DNA sequence. *J. Biol. Chem.*, **1978**, *253* (18), 6551–6560.

(186) Dalbadie-McFarland, G.; Cohen, L.; Riggs, A.; Morin, C.; Itakura, K.; Richards, J., Oligonucleotide-directed mutagenesis as a general and powerful method for studies of protein function. *Proc. Natl. Acad. Sci.*, **1982**, *79* (21), 6409–6413.

(187) Yu, H.; Huang, H., Engineering proteins for thermostability through rigidifying flexible sites, *Biotechnol. Adv.*, **2014**, *32* (2), 308–315.

(188) Liu, T.; Wang, Y.; Luo, X.; Li, J.; Reed, S. A.; Xiao, H.; Young, T. S.; Schultz, P. G., Enhancing protein stability with extended disulfide bonds, *Proc. Natl. Acad. Sci.*, **2016**, *113* (21), 5910–5915.

(189) Dombkowski, A. A.; Sultana, K. Z.; Craig, D. B., Protein disulfide engineering, *FEBS Lett.*, **2014**, *588* (2), 206–212.

(190) Cui, H.; Eltoukhy, L.; Zhang, L.; Markel, U.; Jaeger, K.-E.; Davari, M. D.; Schwaneberg, U., Less unfavorable salt bridges on the enzyme surface result in more organic cosolvent resistance, *Angew. Chem. Int. Ed.*, **2021**, *60* (20), 11448–11456.

(191) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A., et al., Highly accurate protein structure prediction with AlphaFold, *Nature*, **2021**, *596* (7873), 583–589.

(192) Stephens, Z. D.; Lee, S. Y.; Faghri, F.; Campbell, R. H.; Zhai, C.; Efron, M. J.; Iyer, R.; Schatz, M. C.; Sinha, S.; Robinson, G. E., Big data: astronomical or genomical?, *PLoS Biol.*, **2015**, *13* (7), e1002195.

(193) Magliery, T. J., Protein stability: computation, sequence statistics, and new experimental methods, *Curr. Opin. Struc. Biol.*, **2015**, *33*, 161–168.

(194) Porebski, B. T.; Buckle, A. M., Consensus protein design, *Protein Eng. Des. Sel.*, **2016**, *29* (7), 245–251.

(195) Jones, B. S.; Ross, C. M.; Foley, G.; Pozhydaieva, N.; Sharratt, J. W.; Kress, N.; Seibt, L. S.; Thomson, R. E.; Gumulya, Y.; Hayes, M. A., et al., Engineering Biocatalysts for the C- H Activation of Fatty Acids by Ancestral Sequence Reconstruction, *Angew. Chem. Int. Ed.*, **2024**, *63* (18), e202314869.

(196) Prakinee, K.; Phaisan, S.; Kongjaroon, S.; Chaiyen, P., Ancestral Sequence Reconstruction for Designing Biocatalysts and Investigating their Functional Mechanisms, *JACS Au*, **2024**, *4* (12), 4571–4591.

(197) Spence, M. A.; Kaczmarski, J. A.; Saunders, J. W.; Jackson, C. J., Ancestral sequence reconstruction for protein engineers, *Curr. Opin. Struc. Biol.*, **2021**, *69*, 131–141.

(198) Gumulya, Y.; Baek, J.; Wun, S.; Thomson, R.; Harris, K.; Hunter, D.; Behrendorff, J.; Kulig, J.; Zheng, S.; Wu, X., et al., Engineering highly functional thermostable proteins using ancestral sequence reconstruction, *Nat. Catal.*, **2018** (, 1), 878–888.

(199) Kiefer, F.; Arnold, K.; Künzli, M.; Bordoli, L.; Schwede, T., The SWISS-MODEL Repository and associated resources, *Nucleic Acids Res.*, **2009**, *37* (suppl_1), D387–D392.

(200) Muhammed, M. T.; Aki-Yalcin, E., Homology modeling in drug discovery: Overview, current applications, and future perspectives, *Chem. Biol. Drug Des.*, **2019**, *93* (1), 12–20.

(201) Baek, M.; DiMaio, F.; Anishchenko, I.; Dauparas, J.; Ovchinnikov, S.; Lee, G. R.; Wang, J.; Cong, Q.; Kinch, L. N.; Schaeffer, R. D., et al., Accurate prediction of protein structures and interactions using a three-track neural network, *Science*, **2021**, *373* (6557), 871–876.

(202) Fan, J.; Fu, A.; Zhang, L., Progress in molecular docking, *Quant. Biol.*, **2019**, *7*, 83–89.

(203) Forli, S.; Huey, R.; Pique, M. E.; Sanner, M. F.; Goodsell, D. S.; Olson, A. J., Computational protein–ligand docking and virtual drug screening with the AutoDock suite, *Nat. Protoc.*, **2016**, *11* (5), 905–919.

(204) Hollingsworth, S. A.; Dror, R. O., Molecular dynamics simulation for all, *Neuron*, **2018**, *99* (6), 1129–1143.

(205) Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W., et al. AMBER 10, University of California, 2008.

(206) Senn, H. M.; Thiel, W., QM/MM studies of enzymes, *Curr. Opin. Chem. Biol.*, **2007**, *11* (2), 182–187.

(207) Hansen, N.; Van Gunsteren, W. F., Practical aspects of free-energy calculations: a review, *J. Chem. Theory Comput.*, **2014**, *10* (7), 2632–2647.

(208) Chipot, C.; Pohorille, A. *Free energy calculations*; Springer: 2007; Vol. 86.

(209) Richter, F.; Leaver-Fay, A.; Khare, S. D.; Bjelic, S.; Baker, D., De novo enzyme design using Rosetta3, *PLoS ONE*, **2011**, *6* (5), e19230.

(210) Huang, P.-S.; Boyken, S. E.; Baker, D., The coming of age of de novo protein design, *Nature*, **2016**, *537* (7620), 320–327.

(211) Rohl, C. A.; Strauss, C. E.; Misura, K. M.; Baker, D., Protein structure prediction using Rosetta, In *Methods in Enzymology*, **2004**, pp 66–93.

(212) Zhang, Y.; Ma, C.; Dischert, W.; Soucaille, P.; Zeng, A.-P., Engineering of phosphoserine aminotransferase increases the conversion of L-homoserine to 4-hydroxy-2-ketobutyrate in a glycerol-independent pathway of 1, 3-propanediol production from glucose, *Biotechnol. J.*, **2019**, *14* (9), 1900003.

(213) Wijma, H. J.; Floor, R. J.; Bjelic, S.; Marrink, S. J.; Baker, D.; Janssen, D. B., Enantioselective enzymes by computational design and in silico screening, *Angew. Chem. Int. Ed.*, **2015**, *54* (12), 3726–3730.

(214) Delgado-Arciniega, E.; Wijma, H. J.; Hummel, C.; Janssen, D. B., Computationally supported inversion of ketoreductase stereoselectivity, *ChemBioChem*, **2023**, *24* (9), e202300032.

(215) Li, D.; Wu, Q.; Reetz, M. T., Focused rational iterative site-specific mutagenesis (FRISM), In *Methods in Enzymology*, **2020**, pp 225–242.

(216) Amrein, B. A.; Steffen-Munsberg, F.; Szeler, I.; Purg, M.; Kulkarni, Y.; Kamerlin, S. C. L., CADEE: computer-aided directed evolution of enzymes, *IUCrJ.*, **2017**, *4* (1), 50–64.

(217) Shao, Q.; Jiang, Y.; Yang, Z. J., EnzyHTP: a high-throughput computational platform for enzyme modeling, *J. Chem. Inf. Model.*, **2022**, *62* (3), 647–655.

(218) Stourac, J.; Vavra, O.; Kokkonen, P.; Filipovic, J.; Pinto, G.; Brezovsky, J.; Damborsky, J.; Bednar, D., Caver Web 1.0: identification of tunnels and channels in proteins and analysis of ligand transport, *Nucleic Acids Res.*, **2019**, *47* (W1), W414–W422.

(219) Chovancova, E.; Pavelka, A.; Benes, P.; Strnad, O.; Brezovsky, J.; Kozlikova, B.; Gora, A.; Sustr, V.; Klvana, M.; Medek, P., et al., CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures, *PLoS Comput. Biol.*, **2012**.

(220) Zhao, P.; Kong, F.; Jiang, Y.; Qin, X.; Tian, X.; Cong, Z., Enabling peroxygenase activity in cytochrome P450 monooxygenases by engineering hydrogen peroxide tunnels, *J. Am. Chem. Soc.*, **2023**, *145* (9), 5506–5511.

(221) Zhou, J.; Huang, M., Navigating the landscape of enzyme design: from molecular simulations to machine learning, *Chem. Soc. Rev.*, **2024**, *53* (16), 8202–8239.

(222) Sun, R.; Wu, D.; Chen, P.; Zheng, P., Cutting-edge computational approaches in enzyme design and activity enhancement, *Biochem. Eng. J.*, **2024**, 109510.

(223) Xu, H., The slow but steady rise of binding free energy calculations in drug discovery, *J. Comput.-Aided Mol. Des.*, **2023**, *37* (2), 67–74.

(224) Goldenzweig, A.; Fleishman, S. J., Principles of protein stability and their application in computational design, *Annu. Rev. Biochem.*, **2018**, *87* (1), 105–129.

(225) Weinstein, J. J.; Goldenzweig, A.; Hoch, S.; Fleishman, S. J., PROSS 2: a new server for the design of stable and highly expressed protein variants, *Method. Biochem. Anal.*, **2021**, *37* (1), 123–125.

(226) Goldenzweig, A.; Goldsmith, M.; Hill, S. E.; Gertman, O.; Laurino, P.; Ashani, Y.; Dym, O.; Unger, T.; Albeck, S.; Prilusky, J., et al., Automated structure-and sequence-based design of proteins for high bacterial expression and stability, *Mol. Cell*, **2016**, *63* (2), 337–346.

(227) Malladi, S. K.; Schreiber, D.; Pramanick, I.; Sridevi, M. A.; Goldenzweig, A.; Dutta, S.; Fleishman, S. J.; Varadarajan, R., One-step sequence and structure-guided optimization of HIV-1 envelope gp140, *Curr. Res. Struct. Biol.*, **2020**, *2*, 45–55.

(228) Khersonsky, O.; Lipsh, R.; Avizemer, Z.; Ashani, Y.; Goldsmith, M.; Leader, H.; Dym, O.; Rogotner, S.; Trudeau, D. L.; Prilusky, J., et al., Automated design of efficient and functionally diverse enzyme repertoires, *Mol. Cell*, **2018**, *72* (1), 178–186.

(229) Niklas, S.; Wan; M"unch, L.; Weissenborn, J.; Martin, J., Secretion and directed evolution of unspecific peroxygenases in S. cerevisiae, In *Methods in Enzymology***, 2023**, pp 267–306.

(230) Tonin, F.; Tieves, F.; Willot, S.; Van Troost, A.; Van Oosten, R.; Breestraat, S.; Van Pelt, S.; Alcalde, M.; Hollmann, F., Pilot-scale production of peroxygenase from Agrocybe aegerita, *Org. Process Res. Dev.*, **2021**, *25* (6), 1414–1418.

(231) Horst, A.; Bormann, S.; Meyer, J.; Steinhagen, M.; Ludwig, R.; Drews, A.; Ansorge-Schumacher, M.; Holtmann, D., Electro-enzymatic hydroxylation of ethylbenzene by the evolved unspecific peroxygenase of Agrocybe aegerita, *J. Mol. Catal. B: Enzym.*, **2016**, *133*, S137–S142.

(232) Alcántara, A. R.; Domínguez de María, P.; Littlechild, J. A.; Schürmann, M.; Sheldon, R. A.; Wohlgemuth, R., Biocatalysis as key to sustainable industrial chemistry, *ChemSusChem*, **2022**, *15* (9), e202102709.

(233) Adrio, J. L.; Demain, A. L., Microbial enzymes: tools for biotechnological processes, *Biomolecules*, **2014**, *4* (1), 117–139.

(234) Expert Market Research Global Enzymes Market Report, Accessed: 2025-08-20, 2023.

(235) IMARC Group Enzymes Market Size, Share, Trends, and Forecast, Accessed: 2025-08-20, 2023.

(236) Grand View Research Enzymes Market Size, Share & Trends Analysis Report By Type, Product, Source, Accessed: 2025-08-20, 2024.

(237) May, O., Industrial Enzyme Applications – Overview and Historic Perspective, In *Industrial Enzyme Applications***, 2019**, pp 1–24.

(238) Raveendran, S.; Parameswaran, B.; Ummalyma, S. B.; Abraham, A.; Mathew, A. K.; Madhavan, A.; Rebello, S.; Pandey, A., Applications of microbial enzymes in food industry, *Food Technol. Biotech.*, **2018**, *56* (1), 16.

(239) Fasim, A.; More, V. S.; More, S. S., Large-scale production of enzymes for biotechnology uses, *Curr. Opin. Biotech.*, **2021**, *69*, 68–76.

(240) Victorino da Silva Amatto, I.; Gonsales da Rosa-Garzon, N.; Antônio de Oliveira Simões, F.; Santiago, F.; Pereira da Silva Leite, N.; Raspante Martins, J.; Cabral, H., Enzyme engineering and its industrial applications, *Biotechnol. Appl. Bioc.*, **2022**, *69* (2), 389–409.

(241) Ebner, K.; Pfeifenberger, L. J.; Rinnofner, C.; Schusterbauer, V.; Glieder, A.; Winkler, M., Discovery and heterologous expression of unspecific peroxygenases, *Catalysts*, **2023**, *13* (1), 206.

(242) Conesa, A.; van de Velde, F.; van Rantwijk, F.; Sheldon, R. A.; van den Hondel, C. A.; Punt, P. J., Expression of the Caldariomyces fumagoChloroperoxidase in Aspergillus niger and Characterization of the Recombinant Enzyme, *J. Biol. Chem.*, **2001**, *276* (21), 17635–17640.

(243) Bormann, S.; Kellner, H.; Hermes, J.; Herzog, R.; Ullrich, R.; Liers, C.; Ulber, R.; Hofrichter, M.; Holtmann, D., Broadening the biocatalytic toolbox—Screening and expression of new unspecific peroxygenases, *Antioxidants*, **2022**, *11* (2), 223.

(244) Vieira Gomes, A. M.; Souza Carmo, T.; Silva Carvalho, L.; Mendonça Bahia, F.; Parachin, N. S., Comparison of yeasts as hosts for recombinant protein production, *Microorganisms*, **2018**, *6* (2), 38.

(245) Gonzalez-Perez, D.; Garcia-Ruiz, E.; Alcalde, M., Saccharomyces cerevisiae in directed evolution: an efficient tool to improve enzymes, *Bioengineered*, **2012**, *3* (3), 174–179.

(246) Ding, Y.-W.; Lu, C.-Z.; Zheng, Y.; Ma, H.-Z.; Jin, J.; Jia, B.; Yuan, Y.-J., Directed evolution of the fusion enzyme for improving astaxanthin biosynthesis in Saccharomyces cerevisiae, *Synth. Syst. Biotechnol.*, **2023**, *8* (1), 46–53.

(247) Viña-Gonzalez, J.; Elbl, K.; Ponte, X.; Valero, F.; Alcalde, M., Functional expression of aryl-alcohol oxidase in Saccharomyces cerevisiae and Pichia pastoris by directed evolution, *Biotechnol. Bioeng.*, **2018**, *115* (7), 1666–1674.

(248) Mateljak, I.; Rice, A.; Yang, K.; Tron, T.; Alcalde, M., The generation of thermostable fungal laccase chimeras by SCHEMA-RASPP structure-guided recombination in vivo, *ACS Synth. Biol.*, **2019**, *8* (4), 833–843.

(249) González-Rodríguez, S.; Trueba-Santiso, A.; Lu-Chau, T. A.; Moreira, M. T.; Eibes, G., Valorization of bioethanol by-products to produce unspecific peroxygenase with Agrocybe aegerita: Technological and proteomic perspectives, *New Biotechnol.*, **2023**, *76*, 63–71.

(250) Gomez de Santos, P.; González-Benjumea, A.; Fernandez-Garcia, A.; Aranda, C.; Wu, Y.; But, A.; Molina-Espeja, P.; Maté, D. M.; Gonzalez-Perez, D.; Zhang, W., et al., Engineering

a highly regioselective fungal peroxygenase for the synthesis of hydroxy fatty acids, *Angew. Chem.*, **2023**, *135* (9), e202217372.

(251) Hilberath, T.; Van Oosten, R.; Victoria, J.; Brasselet, H.; Alcalde, M.; Woodley, J. M.; Hollmann, F., Toward kilogram-scale peroxygenase-catalyzed oxyfunctionalization of cyclohexane, *Org. Process Res. Dev.*, **2023**, *27* (7), 1384–1389.

(252) Meyer, L.-E.; Hauge, B. F.; Kvorning, T. M.; De Santis, P.; Kara, S., Continuous oxyfunctionalizations catalyzed by unspecific peroxygenase, *Catal. Sci. Technol.*, **2022**, *12* (21), 6473–6485.

(253) Abramson, J.; Adler, J.; Dunger, J.; Evans, R.; Green, T.; Pritzel, A.; Ronneberger, O.; Willmore, L.; Ballard, A. J.; Bambrick, J., et al., Accurate structure prediction of biomolecular interactions with AlphaFold 3, *Nature*, **2024**, *630* (8016), 493–500.

(254) Hekkelman, M. L.; de Vries, I.; Joosten, R. P.; Perrakis, A., AlphaFill: enriching AlphaFold models with ligands and cofactors, *Nat. Methods*, **2023**, *20* (2), 205–213.

(255) Teufl, M.; Zajc, C. U.; Traxlmayr, M. W., Engineering strategies to overcome the stability–function trade-off in proteins, *ACS Synth. Biol.*, **2022**, *11* (3), 1030–1039.

(256) Hou, Q.; Rooman, M.; Pucci, F., Enzyme stability-activity trade-off: new insights from protein stability weaknesses and evolutionary conservation, *J. Chem. Theory Comput.*, **2023**, *19* (12), 3664–3671.

(257) Vanella, R.; Küng, C.; Schoepfer, A. A.; Doffini, V.; Ren, J.; Nash, M. A., Understanding activity-stability tradeoffs in biocatalysts by enzyme proximity sequencing, *Nat. Commun.*, **2024**, *15* (1), 1807.

(258) Chen, K.; Arnold, F. H., Engineering cytochrome P450s for enantioselective cyclopropenation of internal alkynes, *J. Am. Chem. Soc.*, **2020**, *142* (15), 6891–6895.

(259) Wang, Z.-Y.; Ma, X.-Y.; Wu, Y.; Liu, Y.; Lin, G.-X.; Liu, X.-Q.; Zhang, C.; Chen, P.; Zheng, Y.; Jia, Z.-J., Engineering Unspecific Peroxygenases for Enantioselective α-Hydroxylation of β-Ketoesters, *Angew. Chem.*, **2025**, e202509359.

(260) Xu, G.; Crotti, M.; Saravanan, T.; Kataja, K. M.; Poelarends, G. J., Enantiocomplementary Epoxidation Reactions Catalyzed by an Engineered Cofactor-Independent Non-natural Peroxygenase, *Angew. Chem.*, **2020**, *132* (26), 10460–10464.

(261) Swoboda, A.; Pfeifenberger, L. J.; Duhović, Z.; Bürgler, M.; Oroz-Guinea, I.; Bangert, K.; Weißensteiner, F.; Parigger, L.; Ebner, K.; Glieder, A., et al., Enantioselective High-Throughput Assay Showcased for the Identification of (R)-as well as (S)-Selective Unspecific Peroxygenases for C- H Oxidation, *Angew. Chem.*, **2023**, *135* (46), e202312721.

(262) Zhao, P.; Chen, J.; Ma, N.; Chen, J.; Qin, X.; Liu, C.; Yao, F.; Yao, L.; Jin, L.; Cong, Z., Enabling highly (R)-enantioselective epoxidation of styrene by engineering unique non-natural P450 peroxygenases, *Chem. Sci.*, **2021**, *12* (18), 6307–6314.

(263) McDonald, A. D.; Higgins, P. M.; Buller, A. R., Substrate multiplexed protein engineering facilitates promiscuous biocatalytic synthesis, *Nat. Commun.*, **2022**, *13* (1), 5242.

(264) St-Jacques, A. D.; Eyahpaise, M.-E. C.; Chica, R. A., Computational design of multisubstrate enzyme specificity, *ACS Catal.*, **2019**, *9* (6), 5480–5485.

(265) Gómez de Santos, P.; Mateljak, I.; Hoang, M. D.; Fleishman, S. J.; Hollmann, F.; Alcalde, M., Repertoire of computationally designed peroxygenases for enantiodivergent C–H oxyfunctionalization reactions, *J. Am. Chem. Soc.*, **2023**, *145* (6), 3443–3453.

(266) Tawfik, D. S.; Gruic-Sovulj, I., How evolution shapes enzyme selectivity–lessons from aminoacyl-tRNA synthetases and other amino acid utilizing enzymes, *FEBS J.*, **2020**, *287* (7), 1284–1305.

(267) Nam, H.; Lewis, N. E.; Lerman, J. A.; Lee, D.-H.; Chang, R. L.; Kim, D.; Palsson, B. O., Network context and selection in the evolution to enzyme specificity, *Science*, **2012**, *337* (6098), 1101–1104.

(268) Wrenbeck, E. E.; Azouz, L. R.; Whitehead, T. A., Single-mutation fitness landscapes for an enzyme on multiple substrates reveal specificity is globally encoded, *Nat. Commun.*, **2017**, *8* (1), 15695.

(269) Zhang, W.; Dourado, D. F.; Fernandes, P. A.; Ramos, M. J.; Mannervik, B., Multidimensional epistasis and fitness landscapes in enzyme evolution, *Biochem. J.*, **2012**, *445* (1), 39–46.

(270) Sandhu, M.; Chen, J. Z.; Matthews, D. S.; Spence, M. A.; Pulsford, S. B.; Gall, B.; Kaczmarski, J. A.; Nichols, J.; Tokuriki, N.; Jackson, C. J., Computational and Experimental Exploration of Protein Fitness Landscapes: Navigating Smooth and Rugged Terrains, *Biochemistry*, **2025**, *64* (8), 1673–1684.

(271) Wittmann, B. J.; Yue, Y.; Arnold, F. H., Informed training set design enables efficient machine learning-assisted directed protein evolution, *Cell Systems*, **2021**, *12* (11), 1026–1045.

(272) Munch, J.; Soler, J.; Hunecke, N.; Homann, D.; Garcia-Borras, M.; Weissenborn, M. J., Computational-aided engineering of a selective unspecific peroxygenase toward enantiodivergent β-ionone hydroxylation, *ACS Catal.*, **2023**, *13* (13), 8963–8972.

(273) Stimple, S. D.; Smith, M. D.; Tessier, P. M., Directed evolution methods for overcoming trade-offs between protein activity and stability, *Alche J.*, **2020**, *66* (3), e16814.

(274) Romero, P. A.; Krause, A.; Arnold, F. H., Navigating the protein fitness landscape with Gaussian processes, *Proc. Natl. Acad. Sci.*, **2013**, *110* (3), E193–E201.

(275) Hopf, T. A.; Ingraham, J. B.; Poelwijk, F. J.; Schärfe, C. P.; Springer, M.; Sander, C.; Marks, D. S., Mutation effects predicted from sequence co-variation, *Nat. Biotechnol.*, **2017**, *35* (2), 128–135.

(276) Wang, T.; Xiang, G.; He, S.; Su, L.; Wang, Y.; Yan, X.; Lu, H., DeepEnzyme: a robust deep learning model for improved enzyme turnover number prediction by utilizing features of protein 3D-structures, *Brief. Bioinform.*, **2024**, *25* (5).

(277) Yang, J.; Li, F.-Z.; Arnold, F. H., Opportunities and challenges for machine learning-assisted enzyme engineering, *ACS Cent. Sci.*, **2024**, *10* (2), 226–241.

(278) Feehan, R.; Montezano, D.; Slusky, J. S., Machine learning for enzyme engineering, selection and design, *Protein Eng. Des. Sel.*, **2021**, *34*, gzab019.

(279) Markus, B.; Andreas, K.; Arkadij, K.; Stefan, L.; Gustav, O.; Elina, S.; Radka, S., Accelerating biocatalysis discovery with machine learning: a paradigm shift in enzyme engineering, discovery, and design, *ACS Catal.*, **2023**, *13* (21), 14454–14469.

(280) Ding, K.; Chin, M.; Zhao, Y.; Huang, W.; Mai, B. K.; Wang, H.; Liu, P.; Yang, Y.; Luo, Y., Machine learning-guided co-optimization of fitness and diversity facilitates combinatorial library design in enzyme engineering, *Nat. Commun.*, **2024**, *15* (1), 6392.

(281) Vanella, R.; Kovacevic, G.; Doffini, V.; de Santaella, J. F.; Nash, M. A., High-throughput screening, next generation sequencing and machine learning: advanced methods in enzyme engineering, *Chem. Commun.*, **2022**, *58* (15), 2455–2467.

(282) Sharifani, K.; Amini, M., Machine learning and deep learning: A review of methods and applications, *World Information Technology and Engineering Journal*, **2023**, *10* (07), 3897–3904.

(283) Karplus, M.; McCammon, J. A., Molecular dynamics simulations of biomolecules, *Nat. Struct. Biol.*, **2002**, *9* (9), 646–652.

(284) Dror, R. O.; Dirks, R. M.; Grossman, J.; Xu, H.; Shaw, D. E., Biomolecular simulation: a computational microscope for molecular biology, *Annu. Rev. Biophys.*, **2012**, *41*, 429–452.

(285) Teh, S.-Y.; Lin, R.; Hung, L.-H.; Lee, A. P., Droplet microfluidics, *Lab Chip*, **2008**, *8* (2), 198–220.

(286) Yan, W.; Li, X.; Zhao, D.; Xie, M.; Li, T.; Qian, L.; Ye, C.; Shi, T.; Wu, L.; Wang, Y., Advanced strategies in high-throughput droplet screening for enzyme engineering, *Biosens. Bioelectron.*, **2024**, *248*, 115972.

(287) Arnold, F. H., Directed evolution: bringing new chemistry to life, *Angew. Chem. Int. Ed.*, **2017**, *57* (16), 4143.

(288) Broom, A.; Trainor, K.; Jacobi, Z.; Meiering, E. M., Computational modeling of protein stability: quantitative analysis reveals solutions to pervasive problems, *Structure*, **2020**, *28* (6), 717–726.

(289) Kroll, A.; Ranjan, S.; Engqvist, M. K.; Lercher, M. J., A general model to predict small molecule substrates of enzymes based on machine and deep learning, *Nat. Commun.*, **2023**, *14* (1), 2787.

(290) Kouba, P.; Kohout, P.; Haddadi, F.; Bushuiev, A.; Samusevich, R.; Sedlar, J.; Damborsky, J.; Pluskal, T.; Sivic, J.; Mazurenko, S., Machine learning-guided protein engineering, *ACS Catal.*, **2023**, *13* (21), 13863–13895.