PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE

SCHOOL OF ENGINEERING

# IDENTIFICATION AND CHARACTERIZATION OF FLAVOPROTEIN MONOOXYGENASES FOR BIOCATALYSIS

## ALEJANDRO A. GRAN SCHEUCH

Thesis submitted to the Office of Graduate Studies in partial fulfillment of the requirements for the Degree of Doctor in Engineering Sciences

Advisor:
**LORETO PARRA ATALA**
**MARCO W. FRAAIJE**

Santiago de Chile, January, 2021

PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE

SCHOOL OF ENGINEERING

# IDENTIFICATION AND CHARACTERIZATION OF FLAVOPROTEIN MONOOXYGENASES FOR BIOCATALYSIS

## ALEJANDRO A. GRAN SCHEUCH

Members of the Committee:

**JUAN A. ASENJO**

**MÓNICA VÁSQUEZ PÉREZ**

**CÉSAR RAMIREZ SARMIENTOS**

**WILLEM J.H. VAN BERKEL**

**JUAN DE DIOS ORTUZAR**

Thesis submitted to the Office of Graduate Studies in partial fulfillment of the requirements for the Degree Doctor in Engineering Sciences

Santiago de Chile, January, 2021

# Identification and characterization of flavoprotein monooxygenases for biocatalysis

**Alejandro Alberto Gran Scheuch**

# Identification and characterization of flavoprotein monooxygenases for biocatalysis

**PhD thesis**

to obtain the degree of PhD at the
University of Groningen
on the authority of the
Rector Magnificus Prof. C. Wijmenga
and in accordance with
the decision by the College of Deans
and
to obtain the degree of PhD at the
Pontificia Universidad Católica de Chile
on the authority of the
Rector Prof. I. Sánchez Diaz
and in accordance with
the decision by the the Office of Graduate Studies

Double PhD degree

This thesis will be defended in public on

Friday 8 January 2021 at 14.30 hours

by

**Alejandro Alberto Gran Scheuch**
born on 26 December 1989
in Las Condes, Chile

**Supervisors:**
Prof. Dr. M.W. Fraaije
Prof. Dr. D.B. Janssen

**Co-supervisor:**
Dr. L. Parra

**Assessment Committee:**
Prof. Dr. J.A. Asenjo
Prof. Dr. M. Vásquez
Prof. Dr. W.J.H. van Berkel
Prof. Dr. A.J.M. Driessen

# TABLE OF CONTENTS

# Baeyer-Villiger monooxygenases, tunable oxidative biocatalysts

Maximilian J. L. J. Fürst, **Alejandro Gran-Scheuch**, Friso S. Aalbers and Marco W. Fraaije

# ABSTRACT

Pollution, accidents, and misinformation have earned the pharmaceutical and chemical industry a poor public reputation, despite their undisputable importance to society. Biotechnological advances hold the promise to enable a future of drastically reduced environmental impact and rigorously more efficient production routes at the same time. This is exemplified in the Baeyer-Villiger reaction, which offers a simple synthetic route to oxidize ketones to esters, but application is hampered by the requirement of hazardous and dangerous reagents. As an attractive alternative, flavin-containing Baeyer-Villiger monooxygenases (BVMOs) have been investigated for their potential as biocatalysts for a long time, and many variants have been characterized. After a general look at the state of biotechnology, we here summarize the literature on biochemical characterizations, mechanistic and structural investigations, as well as enzyme engineering efforts in BVMOs. With a focus on recent developments, we critically outline the advances towards tuning these enzymes suitable for industrial applications.

**Keywords**: Baeyer-Villiger, ketone oxidation, peroxyflavin, cyclohexanone monooxygenase, phenylacetone monooxygenase, biocatalysis, protein engineering

# INTRODUCTION

"The field of organic chemistry is exhausted."[1] This notion, which many scientists later judged a fallacy,[2] was not an isolated opinion in the late 19th century[3] from when the quote stems. It is ascribed to chemist Adolf von Baeyer and was supposedly in response to the success in synthesizing glucose,[4] achieved by his earlier student, Emil Fischer. While Fischer was said to share von Baeyer's confidence,[3] their potential rush to judgment did not prevent either of them to later be awarded the Nobel Prize. In the wake of ever more discoveries being made, scientists today largely refrain from such drastically exclusivistic statements and rather call organic chemistry a 'mature science'.[5]

In hindsight, the time of von Baeyer's controversial statement can in fact be considered as the early days of organic synthesis. Chemistry only started to transform from an analytic to a synthetic discipline after 1828,[6] when Wöhler's urea synthesis was the first proof that organic compounds do not require a 'vital force'.[7] Similarly to this paradigm shift in chemistry nearly 200 years ago, biology is currently at a turning point.[6, 8] Although bread making and beer-brewing can be considered biotechnological processes invented thousands of years ago, the deliberate creation of synthetic biological systems only succeeded in the late 20th century. As much of modern research, biotechnology is an interdisciplinary area,[5] though, a particularly strong overlap with organic synthesis occurs in the field of biocatalysis. One of the main arguments for using enzymes for chemical transformations is that even if inventions in organic chemistry will never exhaust—its major feedstock soon will. Considering the continuing depletion of the world's fossil fuel reserves, a major contemporary challenge represents the switch to synthetic routes starting from alternative building blocks. In the light of the chemical industry and their supplier's historically disastrous impact on the environment,[9] a second challenge is the transition to what has been termed "green chemistry":[10] the choice of building blocks from sustainable sources and the avoidance of hazardous substances. Moreover, with the chemical industry being the single most energy intensive industry sector worldwide,[11] strategies to increase efficiency of chemical processes are urgently needed. Unfortunately, however, such considerations find only reluctant implementation in practice. Despite an increased public pressure due to the poor reputation of the chemical industry,[12] the market economy still nearly irrevocably ensures the design of industrial processes by economical considerations.[13] In research, delaying factors might include the hesitancy to rethink traditional approaches and the fact that environmental considerations are often inconspicuous on lab-scale, or out of focus due to the limited scientific prestige.[12-13] In the meantime, biocatalytic transformations emerged as a profoundly different alternative. Besides the prospect of inherently green catalysts, a hallmark of biocatalysis is product selectivity, as enzymatic reactions arguably allow total control over chemo-, regio-, and enantioselectivity. This renders biocatalysis especially useful for the preparation of pharmaceuticals, where isomeric impurity can

have dramatic physiological consequences.[14] One of the biggest assets of enzymes is the prospective of their targeted functional evolvability.[15-16] Ever more sophisticated molecular biological methods for DNA manipulation allow easy access to large numbers of enzyme variants, which can be screened for desired activities. Despite being one of the oldest techniques, random mutagenesis libraries continues to be an extremely successful enzyme engineering approach.[15] On the other hand, more rational approaches guided by structural and biochemical data in combination with computational predictions have gained popularity.[17] Although still impractical in most scenarios, the complete *de novo* design of enzymes has been demonstrated and likely will become a key technology in the future.[18]

Although often seen as a limitation, the usually found restriction of enzymes to aqueous systems and ambient temperatures is also advantageous: these processes not only abide by the principles of green chemistry; the consistency in process conditions also facilitates the design of cascade reactions, which circumvents the need to isolate intermediate products. Cascades can be designed as *in vitro* processes, in which chemoenzymatic strategies may combine the power of chemo- and biocatalysis.[19] With whole cells as catalysts being the economically most attractive approach,[20] another highly promising procedure is to establish cascades fully *in vivo*. Recent advances in genetic manipulation techniques greatly accelerated metabolic engineering approaches, allowing the introduction of foreign metabolic pathways into recombinant microbial hosts. These pathways may be of natural origin, partially adapted, or designed entirely *de novo*. Recent examples of the recombinant production of natural products such as opiods[21-22] or cannabinoids[23] attracted considerable attention not only in the scientific community. Artificial metabolic routes designed in a "bioretrosynthetic"[24] fashion, on the other hand, allow diverse applications ranging from novel $CO_2$ fixation strategies[25] to the production of synthetic compounds such as the anti-malarial drug artemisinin.[26] With research in this area of biotechnology rapidly developing, it has been suggested to constitute a new field, called synthetic metabolism.[27]

## THE BAEYER-VILLIGER REACTION OF PEROXIDES AND MONOOXYGENASES

Presumably, considerations of green chemistry were far from the mind of the before-mentioned Adolf von Baeyer, when 110 years ago, he and his disciple Victor Villiger were experimenting with potassium monopersulfate. In the honor of their discovery that this and other peroxides can oxidize ketones to esters, we now call this the Baeyer-Villiger reaction. Although it is a widely known method in organic chemistry nowadays,[28-29] several unsolved difficulties reduce its attractiveness and thus applicability. Especially on large scale, a remaining problem is the shock-sensitivity and explosiveness of many

peroxides.[30] Commonly applied peracids are prepared from their corresponding acids using concentrated hydrogen peroxide. As these solutions in high concentrations are prone to ignition and other forms of violent decomposition,[31] they have largely been withdrawn from the market.[32] Reactions with peroxides and peracids furthermore lead to stoichiometric amounts of hazardous waste products. More promise lies in recent achievements of reactions using directly hydrogen peroxide as the oxidant,[33] making use of metal[34] or organocatalysts.[35] However, such processes also require waste treatment and the catalysts need to be prepared in additional, often complex synthetic routes. In comparison to other oxygenation reactions, examples of asymmetric Baeyer-Villiger oxidations were noted to be scarce and to show limited selectivities, reactivities and scopes.[33]

Due to these reasons, biocatalysis offers a particularly promising alternative and has attracted considerable attention. So-called Baeyer-Villiger monooxygenases (BVMOs) use the free, abundant, and green oxidant $O_2$, and only generate water as byproduct. BVMOs were discovered in the late 1960s by Forney and Markovetz, who were interested in the microbial catabolism of naturally occurring, long-chain methyl ketones. They noticed that the products generated from these compounds by a Pseudomonad were incompatible with terminal methyl oxidation, which was previously assumed to be the only degradation pathway.[36] Subsequently, they were able to identify the responsible enzymatic reaction as a Baeyer-Villiger transformation, dependent on NADPH and molecular oxygen.[37] In parallel, Trudgill and coworkers were investigating microorganisms able to grow on non-naturally occurring aliphatics. They identified an oxygen- and NADPH-dependent enzyme from *Acinetobacter calcoaceticus* NCIMB 9871 involved in the microbial metabolism of fossil fuel-derived cyclohexane and suggested that it catalyzes the conversion of cyclohexanone to $\varepsilon$-caprolactone.[38] They confirmed their findings by isolating the protein and established that the enzyme contains a flavin adenine dinucleotide cofactor as prosthetic group.[39] This cyclohexanone monooxygenase (AcCHMO) quickly attracted attention because of its broad substrate scope and because caprolactone was already well known as a precursor to nylon 6.[40-41]

## STRUCTURES

Over the decades, AcCHMO has come to be the number one prototype BVMO, despite the failure to obtain its structure. Only in 2019, a mutant could finally be crystallized;[42] however, it remains to be seen whether its structure can serve as a good enough approximation to wild type, considering that it contains ten active-site substitutions. 15 years earlier, the first BVMO crystal structure was solved for phenylacetone monooxygenase (PAMO) from *Thermobifida fusca* (**Figure 1a**),[43] causing this thermostable enzyme to become a structural prototype. The structure shed light on a two-residue insertion displayed by PAMO, which

was found to be located in the active site and subsequently called "the bulge" (**Figure 1b**). Eight other BVMOs and various mutant structures followed (**Table 1**), totaling to 38 structures at the time of writing. Mechanistic insights have mostly been gained by structural studies on CHMO from *Rhodococcus* sp. HI-31 (RhCHMO) and PAMO. Overall, the structures of BVMOs are surprising similar, despite sequence identities of often less than 40%. With the exception of PAMO, many BVMOs are often rather unstable; however, no obvious structural features could be identified as the origin of this stability. A study that compared PAMO and AcCHMO's tolerance towards cosolvents—a feature frequently shown to be related to thermostability[44]—suggested PAMO's increased number of ionic bridges would cause the lower solvent susceptibility, as it could prevent damage to the secondary and tertiary structure.[45] The same reasoning was given for the higher robustness of a recently crystalized CHMO from *Thermocrispum municipale* (TmCHMO).[46] BVMOs display a multi-domain architecture consisting of an FAD-binding, an NADP-binding and a helical domain. The latter distinguishes BVMOs from other class B flavoprotein monooxygenase families and causes a partial shielding of the active site and the formation of a tunnel towards it. Some BVMO subgroups contain N-terminal extensions of varying length. The structure of such an extension was established in PockeMO, where it forms a long helix and several loops that wrap around the enzyme.[47] This enzyme is more thermostable than most BVMOs, but it is unknown whether the extension plays a role in that. Such a function was suggested for 4-hydroxyacetophenone monooxygenase (HAPMO), where deletion of the extension was not tolerated when exceeding a few amino acids.[48] Removal of only nine amino acids already impaired stability and furthermore decreased the enzyme's tendency to dimerize. Besides FAD, which is found in all BVMO crystal structures, the nicotinamide cofactor is also found in many structures, in accordance with its tight binding to the enzymes.[49] A certain structural mobility of cofactors and loops in BVMOs has been observed and the debate on its role in catalysis has recently been reviewed.[50] The determination of various BVMO structures has been instrumental for the investigation of their catalytic mechanism.

**1**

**Table 1. Available BVMO crystal structures**

| Name | Acronym | Source strain | Uniprot ID | PDB entries | Important residues[a] | | | | Ref. |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | D | R | $\Omega$ | Bulge | |
| cyclohexanone monooxygenase | AcCHMO | Acinetobacter calcoaceticus NCIMB9871 | Q9R2F5 | 6A37[b] | 57 | 327 | W 490 | P–F 431–432 | 42 |
| Aspergillus flavus monooxygenase 838 | Af838MO | Aspergillus flavus NRRL3357 | B8N653 | 5J7X | 63 | 337 | W 502 | PTAF 441–444 | 51 |
| cyclohexanone monooxygenase | RhCHMO | Rhodococcus sp. HI-31 | C0STX7 | 3GWD, 3GWF, 3UCL[c], 4RG3[c], 4RG4[c] | 59 | 329 | W 492 | P–F 433–434 | 52-54 |
| cyclohexanone monooxygenase | RpCHMO | Rhodococcus sp. Phi1 | Q84H73 | 6ERA[b], 6ER9 | 60 | 330 | W 493 | P–F 434–435 | 55 |
| cyclohexanone monooxygenase | TmCHMO | Thermocrispum municipale DSM 44069 | A0A1L1QK39 | 5M10[c], 5M0Z, 6GQI[c] | 59 | 329 | W 492 | P–F 433–434 | 46, 56 |
| 2-oxo-$\Delta^3$–4,5,5-trimethylcyclo-pentenylacetyl-coenzyme A monooxygenase | OTEMO | Pseudomonas putida ATCC 17453 | H3JQW0 | 3UOV, 3UOX, 3UOY, 3UOZ, 3UP4, 3UP5 | 59 | 337 | W 501 | GSTF 440–443 | 57 |
| phenylacetone monooxygenase | PAMO | Thermobifida fusca YX | Q47PU3 | 1W4X, 2YLR, 2YLS, 2YLT[c], 2YLW[bc], 2YLX[bc], 2YLZ[b], 2YM1[b], 2YM2[b], 4C74, 4C77[b], 4D03[b], 4D04[b], 4OVI | 66 | 337 | W 501 | PSAL 440–443 | 43, 58–59 |
| Parvibaculum lavamentivoran monooxygenase | PlBVMO | Parvibaculum lavamentivorans | A7HU16 | 6JDK | 67 | 340 | W 504 | PSGF 443–446 | 60 |
| polycyclic ketone monooxygenase | PockeMO | Thermothelomyces thermophila ATCC 42464 | G2QA95 | 5MQ6 | 133 | 426 | Y 600 | S–Q 536–537 | 47 |
| steroid monooxygenase | STMO | Rhodococcus rhodochrous | O50641 | 4AOS, 4AOX, 4AP1[b],4AP3[b] | 71 | 342 | W 506 | PSVL 445–448 | 61 |

[a]D: active-site aspartate, R: active-site arginine, $\Omega$: active-site aromatic residue, bulge: active site insertion loop
[b]Mutated variant
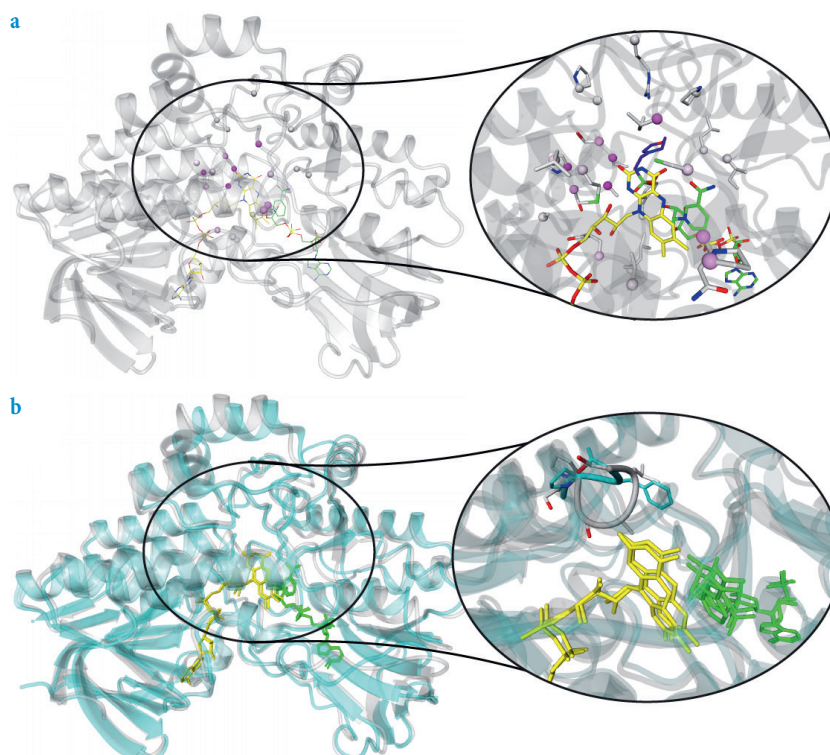[c]Crystallized in complex with an active-site ligand

**Figure 1. Structures of BVMOs.** A) Crystal structure of PAMO shown as ribbons. FAD, NADP⁺ and an active-site ligand are shown as sticks with yellow, green, and dark purple carbons, respectively. C-alpha carbons of residues targeted for engineering are indicated by a sphere. The sphere's color is graded grey to magenta, eflecting the number of reported mutants targeting that site. B) Superimposition of CHMO and PAMO and close-up view of the bulge, a two-residue insertion displayed by PAMO.

## MECHANISM OF THE BAEYER-VILLIGER REACTION

BVMO catalysis (**Scheme 1**) is initiated by NADPH binding and subsequent flavin reduction, after which the nicotinamide cofactor adopts a stable position.[52, 59] Because the stereochemistry of the transferred hydride is in disagreement with the nicotinamide orientation in the stable position, a potential conformational change of NADPH during the reduction step is currently under discussion.[50] Flavoproteins allow detailed mechanistic studies due to the characteristic absorption spectra traversed by the flavin cofactor during the various states of catalysis (**Scheme 2**). In BVMOs, a stable peroxyflavin was identified to be the catalytically active species.[62] Formed by the radical reaction[63] of two electron-reduced FAD with molecular oxygen, this spectroscopically observable flavin intermediate was already known from the flavin-dependent aromatic hydroxylases[64] and luciferases.[65]
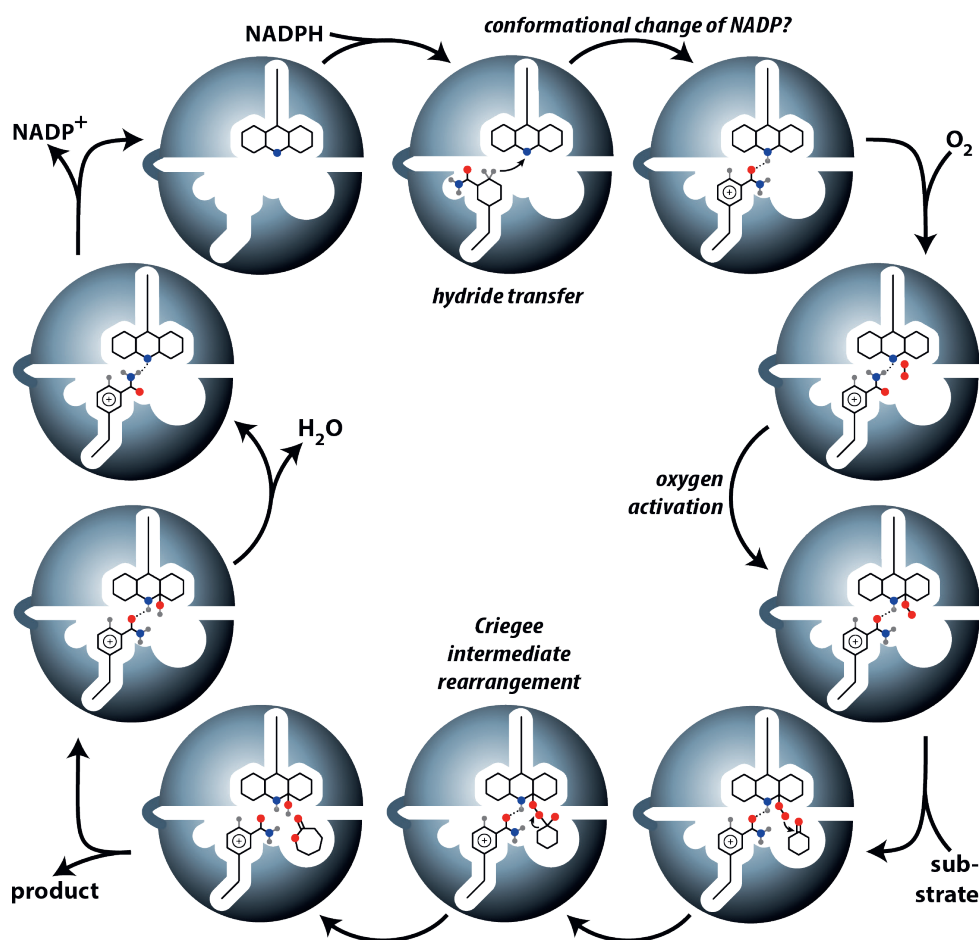
**Scheme 1. Overall catalytic cycle of BVMOs involving various redox states of the flavin and nicotinamide cofactors.** Important atoms are marked by red (oxygen), blue (nitrogen) or grey (hydrogen) circles.

The finding was perhaps rather unsurprising, considering that the chemical Baeyer-Villiger reaction is also afforded by peroxides. However, while with few exceptions,[29] the chemical reaction is acid catalyzed, thus entailing a protonated peroxide, the catalytic flavin species requires a deprotonated peroxy group.[66] While quickly decaying in solution,[67] some BVMOs stabilize this reactive species for several minutes in the absence of a substrate, before its decomposition forms hydrogen peroxide in the "uncoupling" side reaction known to occur in all monooxygenases.[68-71] The exact factors flavoenzymes exert to influence the longevity of both the protonated and unprotonated peroxyflavin are largely unknown, despite reported lifetimes ranging from milliseconds in some oxidases[72] to several minutes or even hours in FMOs[73-74] and luciferases.[75] In BVMOs and other class B

monooxygenases, NADP[+] was, however, found to be critical for intermediate stabilization, as a manifold increased peroxyflavin decay was observed in the absence of the cofactor.[62, 66, 76] Crystal structures and quantum mechanics calculations[77] indicate that the NADP[+] amide oxygen establishes a crucial hydrogen bond to the hydrogen of the flavin's N5 (**Scheme 3**). It is assumed that this stabilization prevents uncoupling by thwarting the otherwise quickly occurring proton transfer to the peroxy group and subsequent $H_2O_2$ elimination.[78] An active-site arginine, whose mutation abolishes Baeyer-Villiger activity,[79] was shown to be essential for the formation, but not for stabilization of the peroxyflavin.[76] The arginine ensures, however, peroxyflavin deprotonation, supported by a nearby aspartate that increases the arginine's nucleophilicity (**Scheme 3**).[77] If a suitable ketone substrate is available, the next canonical step is the nucleophilic attack on the carbonyl group. In BVMOs, the proper positioning of the substrate is thought to be aided by a hydrogen bond between the 2' OH group of the NADP[+] ribose and the carbonyl oxygen (**Scheme 3**).[77] The chemical Baeyer-Villiger reaction was already for a long time assumed to proceed via an intermediate whose nature initially caused some debate. Isotopic labeling experiments[80] eventually gave conclusive evidence for the pathway suggested by Rudolf Criegee,[81] in whose honor the tetrahedral intermediate was subsequently named. Although not directly observable, several computational studies support this mechanism.[82-85] Very recently, experimental evidence was provided from a stereoelectronic trap for the intermediate, using synthetic endocyclic peroxylactones.[86] In BVMOs, a flavin Criegee intermediate was also never observed, but in the absence of any counter-evidence it is generally accepted that here the flavin and substrate also form an addition product, and computational studies support this theory.[77, 87] The product then results from a concerted subsequent migration step, in which the weak O–O bond is heterolytically cleaved, while a new C–O bond is formed. The rearrangement proceeds with retention of configuration[88-89] and is often rate-determining for the chemical reaction, although both experimental[29] and theoretical[82, 90] evidence indicate that the kinetics can change depending on the substituents, pH, and solvent. The regioisomeric outcome of the reaction is generally predictable and governed by a combination of influencing parameters. Firstly, due to the positive charge developing on the migrating carbon in the transition state, the more electronegative carbon, which is better able to accommodate this charge, is more apt to migrate.[91] Thus, carbons with electron donating substituents and those allowing resonance stabilization migrate better than methyl groups and electron withdrawing substituents.[29] Secondly, the C–C bond migrates preferentially when it is *anti*-periplanar to the peroxy O–O bond (**Scheme 2**), a condition known as the primary stereoelectronic effect.[92] Its influence on determining migration is apparently more significant than the migratory aptitude. This was concluded from the observation that a less substituted bond migrates when forced into an *anti*-periplanar conformation in a restrained bicyclic Criegee intermediate.[93] A secondary stereoelectronic effect has also been postulated, requiring that one of the lone electron pairs of the hydroxyl group in the intermediate also needs to be *anti*-periplanar to the

peroxy O–O bond (**Scheme 2**).[94] This effect only manifests in certain substrates, where substituents can sterically hinder the hydroxyl group rotation and presumably plays no role in enzyme catalysis, where the hydroxyl group is assumed to be deprotonated.[77] Lastly, the arrangement can be influenced by steric effects.[95-96] These may furthermore already affect the addition step, where the nucleophilic attack must occur from a favorable angle.[29, 97] Steric control becomes most obvious in the enzymatic reaction, where intermolecular steric restraints can enforce an otherwise electronically prohibited pathway. It is for that reason that BVMO catalysis allows the synthesis of products, which are not accessible by chemical means (**Figure 2**).
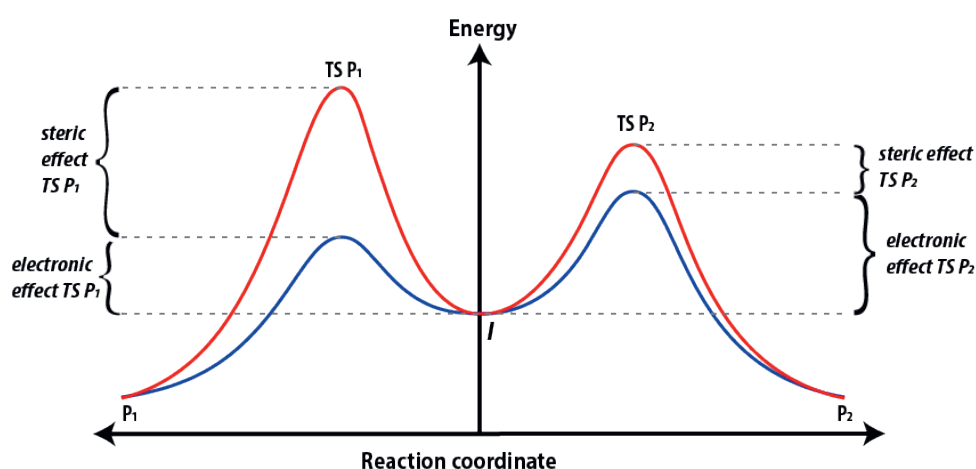


**Figure 2. Simplified energy diagram depicting the electronic and steric effects affecting regioselectivity in BVMO reactions.** In the Baeyer-Villiger reaction, an intermediate ($I$) is formed, which can undergo two varying migration pathways (**Scheme 2**), leading to two possible products ($P_1$ and $P_2$). In chemical catalysis, the predominant factors can collectively be called electronic effects, and the difference they exhibit on the energy of the two possible transition states, usually dictates the regioselectivity of the product (blue line). In enzyme catalysis, steric effects of active-site residues exhibit an additional force contributing to the overall energy of the transition states which can override the electronically favored pathway (red line).

While the peroxide-catalyzed reaction finishes under formation of the corresponding acid, the flavin can pick up a proton to form a hydroxyflavin, whose spontaneous dehydration reconstitutes the oxidized flavin.[67] It was suggested that this step is accelerated by a deprotonated active site residue with a p$K_a$ of 7.3,[76] in line with the faster decay of this species at higher pH and the decreased overall reaction rates at low pH.[66, 76] Before the enzyme can restart a new catalytic cycle, the oxidized nicotinamide cofactor needs to be ejected, and this step (or an associated conformational change) was found to be limiting to the overall reaction rate in CHMO.[66] In PAMO, the slowest catalytic step was not unambiguously identifiable, but may correspond to a conformational change prior to NADP+ release.[76]

These findings entail two important and possibly conflicting conclusions: firstly, the two most detailed available studies on the mechanism of BVMO catalysis suggest that the enzymatic reaction is limited by a rate-determining step that is not involved in the chemical part of catalysis and therefore possibly substrate-independent. If this was generally the case, it could provide an explanation for the rather narrow range of maximal turnover rates observed for BVMOs with various substrates. Thus, reaction rates that are orders of magnitude higher than the currently know ones cannot be expected for any enzyme-substrate combination. However, this assumption is put in perspective by the second conclusion, which stems from the fact that (at least) the rate-determining step of catalysis appears to be non-identical in CHMO and PAMO. If the two prototype enzymes differ in this crucial aspect, one cannot rule out that even other mechanisms dictate catalysis in other BVMOs. A generalization, therefore, may not be possible, and is furthermore impeded by the mechanistic variations in the chemical part of the reaction specified above, which always have to be considered to play a role on top of the enzymatic peculiarities.
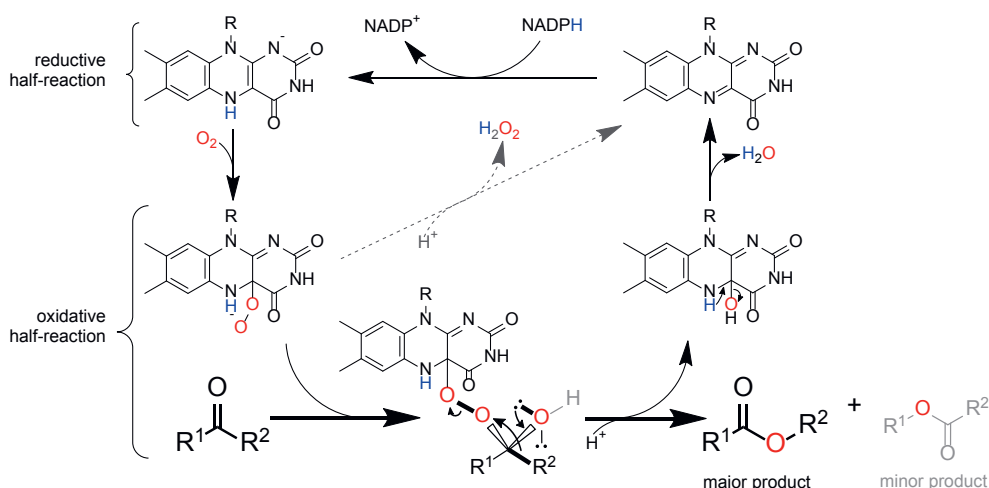


**Scheme 2.** Reaction mechanism of BVMOs. The flavin catalytic cycle consists of two half-reactions and ketone oxidation is catalyzed by a peroxyflavin, unless hydrogen peroxide loss causes an uncoupled NADPH oxidation (grey dashed arrow). The transformation from a ketone to an ester traverses through a regioselectivity-determining intermediate. Bond migration is dependent on the *anti*-periplanar alignment (indicated by thick bonds) of the migrating bond with the peroxy bond and one of the lone pairs on the former carbonyl oxygen. While protonated in the chemical Baeyer-Villiger reaction, this oxygen is, however, thought to be deprotonated in enzyme flavin intermediate (indicated in grey).

# PROMISCUOUS CATALYTIC ACTIVITIES
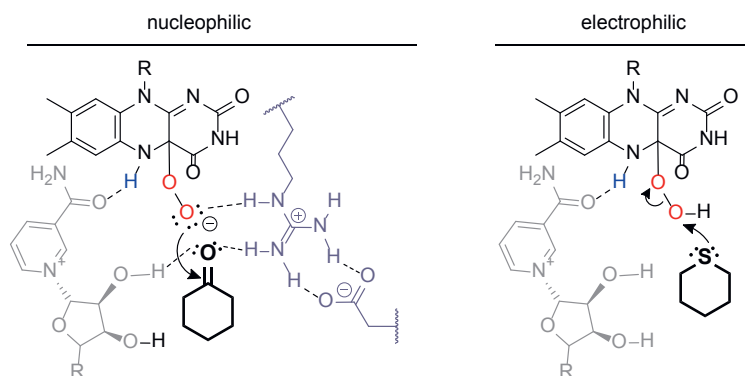
**Scheme 3.** Proposed mechanism for enzyme catalyzed oxidations. In the canonical, nucleophilic mechanism, the peroxyflavin attacks the substrate carbonyl. An active site aspartate increases the basicity of a neighboring arginine, which thus ensures deprotonation of the peroxyflavin. The arginine also activates the substrate ketone, supported by the 2' OH of the ribose of NADP[+]. In contrast, in the electrophilic mechanism a supposed hydroperoxyflavin reacts with the lone pair of a nucleophilic heteroatom.

In addition to the canonical ketone oxidation, BVMOs also are able to perform a range of promiscuous catalytic activities (**Scheme 4**). Well-established and mechanistically analogous to the canonical reaction are BVMO oxidations of aldehydes,[98-103] including furans.[104] This reaction yields acids upon hydrogen migration, or otherwise (often unstable) formates. Although reactions with unsaturated ketones supposedly should also proceed identical in mechanism, most BVMOs show no reactivity with these weaker electrophiles. The transformation is also challenging chemically, where side reactions such as epoxidations frequently occur, and otherwise invariably enol esters are formed, i.e. oxygen insertion occurs towards the double bond.[105] Recently, two bacterial BVMOs were reported that can convert several cyclic α,β-unsaturated ketones.[106] Interestingly, the two enzymes reacted regiodivergent in some cases, which allowed the selective synthesis of both ene- and enol lactones. Although the crystal structure of the preferentially enol ester-forming enzyme—a BVMO from *Parvibaculum lavamentivorans*—has recently been solved, a structural explanation for its unusual reactivity has yet to be provided.[60] Only two other unsaturated ketones were reported to be accepted by BVMOs before: a substituted cyclopentenone, converted to the corresponding ene lactone by CPMO,[107] and pulegone, a cyclohexanone derivative with a double bond outside the ring on the alpha carbon, for which activity was reported with monoterpene ketone monooxygenase (MMKMO),[108] and cyclopentadecanone monooxygenase (CPDMO).[55] The three enzymes involved in camphor degradation in *Pseudomonas putida*—2,5-diketocamphane 1,2-monooxygenase (2,5-DKCMO), 3,6-diketocamphane 1,6-monooxygenase (3,6-DKCMO) and OTEMO[109]— were also reported to convert several cyclopentenones and cyclohexenones. The results

were questioned by the Alphand group, however,[106] although OTEMO's natural substrate is assumed to be a cyclopentenone derivative.[109-110] Conversion of a linear α,β-unsaturated ketone to the ene ester has been shown for the Baeyer-Villiger reaction-catalyzing human FMO5.[111] Oxidation of esters, which bear an even less electrophilic carbonyl, has been reported for a single BVMO, which is able to catalyze first the ketone oxidation and subsequently further converts the ester to its carbonate.[112]

Similarly to peroxides,[113] BVMOs were early found to promiscuously catalyze heteroatom oxidations as well.[98, 114] Sulfoxidations are particularly well studied and many enzymes produced sulfoxides with high enantioselectivity.[103, 115-125] Several existing patents describing the use of BVMOs for selective sulfoxidations emphasize the commercial potential.[126-128] Other reactions include oxidations of amines,[40, 102, 129-130] boron,[98, 131-132] and selenium.[98, 133-134] A single report of phosphite ester and iodine oxidation yet awaits further exploration,[98] as do the few reports of epoxidations catalyzed by BVMOs.[135-136] An entirely different approach to induce promiscuous catalytic activity is the use of BVMOs under anaerobic conditions to prevent peroxyflavin formation. Recent results with AcCHMO suggest that the so-stabilized reduced flavin can catalyze reductions, allowing tunable, stereoselective ketoreductase-like reactions.[137]
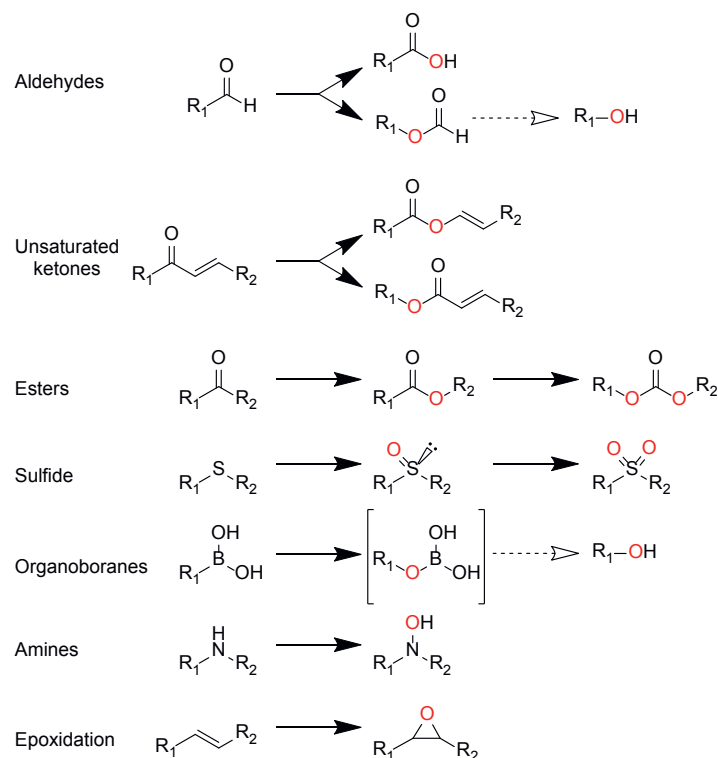


**Scheme 4.** Non-canonical oxidation reactions catalyzed by BVMOs. Solid arrows represent enzymatic catalysis; a dashed arrow indicates spontaneous reaction.

In contrast to the nucleophilic species required for the Baeyer-Villiger reaction, S-, N-, Se-, P-, and I- oxygenation require an electrophilic, protonated peroxyflavin. In line with the mechanism found for class A flavoprotein monooxygenases,[138] this hydroperoxyflavin was suggested to form in BVMOs and an apparent $pK_a$ for the formation was determined to be 8.4 in CHMO.[66] However, as the protonated species in CHMO was not able to perform sulfoxidations, the results are not fully conclusive and it was suggested that some protein conformational change is involved.[139] For PAMO, sulfoxidation enantioselectivity seems to depend on the protonation state of the peroxyflavin and the crucial[76, 79] active site arginine;[140] and mutation of the arginine abolished both Baeyer-Villiger as well as sulfoxidation activity.[76] One study with CHMO, however, seemed to indicate that its heteroatom oxidation activity does not depend on the arginine, as the mutation to alanine or glycine yielded variants with retained S- and N- oxidation activity.[141] In this scenario, the loss of arginine could have two counteracting effects: as quantum mechanics studies suggest that a nearby aspartate protonates the arginine and this stabilizes the negatively charged, deprotonated peroxyflavin,[77] arginine mutation could favor hydroperoxyflavin formation and thus the electrophilic mechanism. On the other hand, arginine loss decreases the overall reaction rate as the residue also promotes the reductive half-reaction and the rate of (hydro)peroxyflavin formation.[76, 142] Interestingly, the substitution of a highly conserved aromatic residue with arginine was found in two independent studies that screened for variants with increased sulfoxidation activity.[42, 127] In most BVMOs, this residue is a tryptophan that hydrogen bonds to the 3' OH of the NADP ribose. Considering the enzyme's tolerance of other aromatic residues at this position,[143] this interaction is likely not influencing the electronics at the 2' OH, which critically hydrogen bonds to the substrate carbonyl to activate it for nucleophilic attack (**Scheme 3**).[77] Rather, a mutation to arginine could push the positively charged coenzyme, possibly causing a disruption of the hydrogen bond to the substrate. Instead, the group might come closer to the peroxyflavin and cause its protonation; this mechanism would favor the electrophilic route and seems to be the mode of action in the closely related N-hydroxylating monooxygenases.[144]

## VARIETY OF BVMOS

In the quest of discovering useful biocatalysts, many studies aimed to identify enzymes displaying features such as high regio- or enantioselectivity, showing a broad or a synthetically interesting substrate scope, lacking substrate or product inhibition and having high stability in typical process conditions. The classic ways to obtain novel efficient biocatalyst are mutagenesis on well-known catalysts and the exploitation of genome sequence databases, which are a rich and largely untapped resource for enzymes with attractive biocatalytic characteristics and novel chemistries.

## BVMO classification

Considerable research has been performed on BVMOs using comparative sequence analysis. Using a curated, representative sequence set, one study suggested that a BVMO gene was already present in the last universal common ancestor.[145] This study also found that there is no conclusive evidence that phylogenetic BVMO subgroups share biocatalytic properties, although this frequently has been and continues to be suggested in literature.[47, 146-147] In the last decades, many BVMOs, both prokaryotic and eukaryotic, have been described and approximately a hundred representatives were cloned and recombinantly expressed. In many cases, the natural role of those BVMOs could not be identified. In other cases, BVMOs were shown to be involved in the biosynthesis of secondary metabolites such as toxins,[148-152] or antibiotics.[153] While these enzymes often seem to be rather substrate specific, several BVMOs from catabolic pathways, involved e.g. in the degradation of cyclic aliphatics,[38, 154-156] can convert a larger range of substrates. Together with the structurally very similar N-hydroxylating- and flavin-containing monooxygenases, BVMOs have been classified as belonging to the class B of flavoprotein monooxygenases.[49] Recently, another sister group has been added—YUCCAs,[157] which are plant enzymes involved in auxin biosynthesis and shown to catalyze a Baeyer-Villiger-like reaction.[158] Some FMOs, including the human isoform 5,[111] were also found to catalyze Baeyer-Villiger reactions, [159] and it was suggested that these enzymes form a particular subgroup, classified as class II FMOs.[160] Their relaxed coenzyme specificity[161] enables interesting application opportunities.[162] Structurally largely unrelated are a few Baeyer-Villiger reaction-catalyzing enzymes found in class A[163] and C flavoprotein monooxygenases,[164] which otherwise comprise the aromatic hydroxylases and luciferases, respectively[49] (**Table 2**). Cytochrome P450 monooxygenases, of which some can catalyze Baeyer-Villiger reactions,[165-166] are entirely unrelated and employ heme cofactors instead of flavins.

**Table 2.** Classification of Baeyer-Villiger biocatalysts.

| Group | Flavoprotein subclass | Hydride donor | Prosthetic group | Components[a] | Prototype protein |
|---|---|---|---|---|---|
| Type I BVMOs | B | NADPH | FAD | α | TfPAMO[43] |
| Type II BVMOs | C | NADH | FMN (substrate) | α + β | 3,6-DKCMO[167] |
| Type O BVMOs | A | NADPH | FAD | α | MtmOIV[168] |
| Type I FMOs | B | NADPH | FAD | α | HsFMO5[111] |
| Type II FMOs | B | NAD(P)H | FAD | α | RjFMO-E[160] |

[a]α: Encoded by a single gene, α + β: Encoded by multiple genes (monooxygenase and a reductase component)

Many Baeyer-Villiger monooxygenases have been discovered and characterized by genome mining.[153, 169-172] Instead of trying to be comprehensive, this review will focus on some examples we believe are worthwhile to examine deeper (**Figure 3, Table 3**). From

**Figure 3. Cladogram analysis of BVMOs examples.** The color of the clade represents the flavoprotein group to which the respective BVMOs belong (cyan for type I BVMOs, yellow for type II BVMOs, orange for type O BVMOs, green for type I FMOs and red for type II FMOs). A star indicates the availability of crystal structures, in green for wild type and white for mutant. The bar chart shows the melting temperature. The outside rings represent the acceptance for different ketone substrates. Note that this only represents substrates that have been tested, while the actual scope might be (much) larger. The species and codes are listed in Table S1.

these proteins, most are type I BVMOs, which are encoded in a single gene and belong to the class B flavoprotein monooxygenases.[49] Several residues in BVMOs are highly conserved and useful for the identification of type I BVMOs. There are two specific sequences described: FxGxxxHxxxW[P/D][173] and [A/G]GxWxxxx[F/Y]P[G/M]xxxD.[172] A modification to the short BVMO fingerprint was suggested (FxGxxxHTxxW[P/D])[174], however, this consensus proved to be only partially conserved in a more divergent dataset of sequences.[145] These motifs are flanked by two Rossmann fold domains harboring a

GxGxx[G/A] motif required for tight binding of the two cofactors. In some cases, minor deviations from the consensus for the nucleotide binding sequence have been reported (MoxY, CPDMO).[155, 175] Although the exact functional role of the fingerprint residues is not completely clear, the long consensus sequence entails the conserved active-site aspartate, while the short fingerprint is related to the linker connecting the FAD and NADP-binding domains.[43, 59] As a common feature, type I BVMOs share the strict dependence on FAD as a tightly bound prosthetic group and NADPH as electron donor, with the exception of MekA from *Pseudomonas veronii* MEK700, which seems to accept either NADH or NADPH.[176]

The preferred host for producing recombinant BVMOs, has been *Escherichia coli*, which does not contain a native homolog itself. BVMOs can also be directly applied in whole-cell conversions, as demonstrated in many reports focusing on valuable bioconversions (see section 'Biotechnological application'), but more detailed characterizations such as kinetic studies often use purified enzymes. Although some homologs show very high expression levels, *E. coli* may not be able to provide the cofactors in the necessary quantities,[177] thereby negatively affecting stability.[178] This effect is assumed to be even more critical when BVMOs are to be applied in *in vivo* cascades with other redox enzymes.[179] An additionally complicating factor in whole cell conversions is oxygen supply, which limits the reaction at high biomass concentrations.[180] When BVMO homologs with interesting biocatalytic properties were found to express poorly, several approaches to improve functional expression and stability were explored. Besides optimization of the expression conditions (cultivation temperature and time, induction method) and the more and more common use of synthetic genes with host-optimized codons, fusion approaches with soluble tags are popular countermeasures. One study also co-expressed molecular chaperones with a BVMO from *P. putida*, and found that optimal results rely on their distinct expression levels.[181]
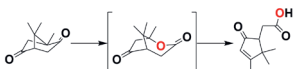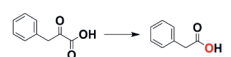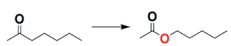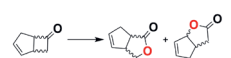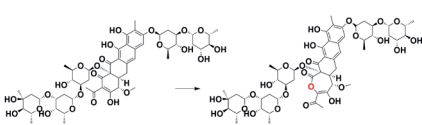
## Eukaryotic Type I BVMOs

Baeyer-Villiger oxidations have frequently been demonstrated in physiological studies.[196-200] BVMO genes were described as scarce in microorganism,[173] though in fact they exhibit an uneven genomic distribution.[201] While bacterial BVMOs are most abundantly found in actinomycetes, there is also a high prevalence in some filamentous fungi. Particularly, BVMOs were found in Basidiomycota, Zygomycota, and the Ascomycota, where they are especially abundant in the *Aspergillus* genus.[145-146] Until recently, most of the research with isolated enzymes investigated prokaryotic BVMOs—possibly due to the easiness to work without the splice components of eukaryotes or to avoid problems with rare codons. One of the first type I BVMOs obtained from a fungus was steroid monooxygenase from *Cylindrocarpon radicicola* ATCC 11011 (STMO), which was purified from cells grown in the presence of progesterone.[202] Although several fungi with Baeyer-Villiger activities

were described, it was only in 2012 when the first recombinant fungal BVMO was expressed by the group of Bornscheuer.[186] This enzyme comes from the same ascomycete as STMO. This fungus is also described to metabolize cyclohexanone as a carbon source and this ability was linked to the presence of a second BVMO, identified as cycloalkanone monooxygenase (CAMO). CAMO shows 45% sequence identity with AcCHMO and exhibits a broad substrate scope, among which cycloaliphatic and bicycloaliphatic ketones showed the highest activites. However, its thermostability is quite poor, and with 28 °C, the temperature for 50% residual activity after 5 minutes of incubation is considerably lower than that of AcCHMO (36 °C).[203] BVMOAf1 from the fungus *Aspergillus fumigatus* Af293 was described one year later.[204] This BVMO stems from a pathogenic organism that is known to be a source of biocompounds such as helvolic acid and fumagillin, in whose biosynthesis the enzyme could be involved. Its activity was found to be relatively low, with maximal rates of catalysis around 0.5 s$^{-1}$; but high enantioselectivities in the oxidations of thioanisole, benzyl ethyl sulfide and bicyclo[3.2.0]hept-2-en-6-one were observed. This enzyme exhibits relatively high thermostability: while the highest activity was recorded at 50 °C, the $T_m$ was 41 °C. In addition, after 1 h incubation in buffers containing 5% of various cosolvents, its activity remained without significant loss. Four other enzymes were discovered from *A. flavus* NRRL3357 (BVMOAfl210, 456, 619 and 838).[205] From those, BVMOAfl838 displayed high conversion of aliphatic ketones; but was unable to convert most of the cyclic ketones tested. BVMOAfl838 later was the first reported crystal structure of a fungal type I BVMO.[51] The enzyme showed an optimal temperature of approximately 40 °C, but was rapidly inactivated at that temperature, displaying a half-life of only 20 min. The structure could not be co-crystallized either with the nicotinamide cofactor or with the substrate and showed a global fold similar to other described BVMOs. Near to the supposed substrate entry channel is a mobile loop that presents a lysine (K511). This residue was suggested to be proximal to the 2'-phosphate of NADPH, and the K511A mutant exhibited a higher uncoupling. Later, more BVMOs from *Aspergillus* were characterized: BVMOAfl706 and BVMOAfl334 (~45% amino acid sequence identity), which converted a range of cyclic and substituted cyclic ketones, and showed highest conversions and $k_{cat}$ values of 4.3 s$^{-1}$ and 2 s$^{-1}$ for cyclohexanone, respectively.[206] Interestingly, no substrate inhibition was observed for BVMOAfl706 with cyclohexanone using concentrations up to 30 mM. In contrast, AcCHMO, shows a $K_i$ of approximately 35 mM[207-208] Subsequently, a study tried to exploit BVMOAfl706 in a cascade reaction for the lactonization of cyclohexanone, but the enzyme seemed to be responsible for the formation of an undesired side product.[209] The last fungal example is polycyclic ketone monooxygenase (PockeMO) from the thermophilic fungus *Thermothelomyces thermophila*, which was discovered and crystallized.[47] This fungus is known to efficiently degrade cellulose and derivatives from plant biomass. This enzyme presented high enantioselectivity for bicyclo[3.2.0]hept-2-en-6-one, and displayed an unusually broad activity on several polycyclic molecules, hence its name. PockeMO exhibited highest activity at 50 °C and a $T_m$ of 47 °C. As metabolically

**Table 3.** Prototype reactions of Baeyer-Villiger monooxygenases.

| Name | Substrate | Prototype reaction | $k_{cat}$ [s$^{-1}$] | $K_m$ [µM] | Ref |
|---|---|---|---|---|---|
| AcCHMO | cyclohexanone | | 6.0-39 | 3-9 | 46, 182-183 |
| GoACMO | acetone | | 1.4 | 170 | 184 |
| BVMO4 | 2-phenyl-propionaldehyde | | n.d. | n.d. | 185 |
| BVMOAfl838 | 3-octanone | | 6.6 | 170 | 51 |
| CAMO | cyclobutanone | | 6.8 | 7 | 186 |
| CmBVMO | 2-dodecanone | | 0.4 | 4 | 187 |
| CPDMO | cyclopentadecanone | | 4.2 | 6.0 | 155 |
| HAPMO | 4-hydroxy-acetophenone | | 10-12 | 9-40 | 103, 188 |
| ObBVMO | 4-methylcyclohex-2-en-1-one | | n.d. | n.d. | 106 |
| PlBVMO | 4-methylcyclohex-2-en-1-one | | n.d. | n.d. | 106 |
| OTEMO | 2-oxo-$\Delta^3$-4,5,5-trimethyl-cyclo pentenylacetyl-CoA | | 4.8 | 18 | 57 |
| PockeMO | bicyclo[3.2.0] hept-2-en-6-one | | 3.3 | 400 | 47 |
| RpBVMO | methyl levulinate | | 1.5 | 350 | 189 |
| SAPMO | 4-sulfoacetophenone | | 2.9 | 60 | 190 |
| STMO | progesterone | | 0.7 | 85 | 191-192 |
| TfPAMO | phenylacetone | | 1.9-3 | 60-80 | 169, 193 |
| TmCHMO | cyclohexanone | | 2.0 | <1 | 46 |
| PtlE | 1-deoxy-11-oxopentalenate | | n.d. | n.d. | 153 |
| 2,5-DKCMO | 2,5-diketocamphane | | n.d. | n.d. | 164 |

| | | | | | |
|---|---|---|---|---|---|
| 3,6-DKCMO | 3,6-diketocamphane | | n.d. | n.d. | [164] |
| AtYUC6 | phenylpyruvate | | 0.31 | 43 | [158] |
| HsFMO5 | 2-heptanone | | n.d. | n.d. | [194] |
| RjFMO-E | bicyclo[3.2.0] hept-2-en-6-one | | 2.0-4.3 | 3 | [161-162] |
| MtmOIV | premithramycin B | | 0.7 | 70 | [195] |

observed for fungi[196] and as was described for STMO[202] and CPDMO,[210] PockeMO is able to regioselectively catalyze the D-ring oxidation of steroid substrates producing the normal lactone. Later, de Gonzalo analyzed the applicability of PockeMO for the synthesis of optically active sulfoxides and showed full conversion of thioanisole into the (*R*)-sulfoxide with excellent selectivity, while for other alkyl phenyl sulfides a decreased activity and selectivity was observed.[117]

Thermostable enzymes have also been found in photosynthetic organisms: CmBVMO from the red algae *Cyanidioschyzon merolae* and PpBVMO from the moss *Physcomitrella patens*.[187] They showed high thermostability; in particular CmBVMO, which displayed a $T_m$ of 56 °C, whereas PpBVMO's $T_m$ was 44 °C. Although their activity was comparatively low, with $k_{cat}$ values in the 0.1–0.3 s$^{-1}$ range, they could achieve modest conversions of cyclohexanone.

## Prokaryotic Type I BVMOs

Among the many bacterial type I BVMOs described in the last years, there are several homologues of AcCHMO, as one goal was to identify a similar, but more stable biocatalyst. One particular example is TmCHMO, which shows 57% sequence identity with AcCHMO.[46] This enzyme stems from *Thermocrispum municipale* DSM 44069, a thermophilic microorganism isolated from municipal waste compost. TmCHMO was described to efficiently convert a variety of aliphatic, cyclic and aromatic ketones, and was also able to oxidize prochiral sulfides. Interestingly, TmCHMO exhibits a $T_m$ of 48 °C and presents stability against high temperatures and the presence of cosolvents. However, as AcCHMOs, this robust enzyme showed inhibition with high substrate concentrations.[208, 211] Another newly described BVMO is BVMO4, identified from the genome of *Dietzia* sp. D5. This enzyme phylogenetically clusters with cyclopentadecanone monooxygenase (CPDMO).[212] BVMO4 displayed a broad substrate scope accepting different ketones

and sulfides, but showed low activity. Although BVMO4 converted alicyclic and aliphatic ketones only moderately, it was also studied for its activity with phenyl group-containing and long aliphatic aldehydes. With respect to the latter, BVMO4 showed high regioselectivity with for example octanal, decanal, and 3-phenylpropionaldehyde, and preferentially synthesized the respective carboxylic acid over the formyl ester. Albeit with rather poor selectivities, this was the only reported BVMO able to convert a 2-substituted aldehyde to the respective acid, which is a precursor of ibuprofen and derivatives.[185] An effort to improve the activity of BVMO4 with cyclohexanone by site saturation mutagenesis over 12 described hot spots was reported.[213] Its activity was successfully increased against cyclic ketones and the oxidation of cyclohexanone was improved. A thorough biochemical characterization was described for a BVMO active on small substrates, acetone monooxygenase (ACMO) from the propane-metabolizing organism *Gordonia* sp. TY-5.[184] ACMO converts small ketones such as acetone and butanone with $k_{cat}$ values between 1.4–4.0 s$^{-1}$; but shows only modest stability, loosing over 60% of the activity after 1 h incubation at 25 °C in buffer. This enzyme displayed a weaker affinity for bulkier substrates and NADPH. The latter was suggested to be caused by a diminished electrostatic interaction between the 2'-phosphate of the coenzyme and the protein due to a substitution of a usually conserved lysine[79] by histidine. Additionally, a monooxygenase from *Leptospira biflexa* that was phylogenetically distant from other well-characterized BVMOs was described by the group of Rial in 2017.[214] LbBVMO showed a broad substrate scope for acyclic, aromatic, cyclic and fused ketones, and allowed the highly regioselective conversion of aliphatic and aromatic ketones. For *Rhodoccocus jostii* RHA1, 22 BVMOs were found in the genome, which showed a diverse scope when tested against a large set of potential substrates including different ketones and sulfides.[147, 172] From these enzymes, at least two are quite promiscuous regarding their substrate scope (RjBVMO4 and RjBVMO24), accepting the majority of the 25 tested compounds.

Furthermore, there a few well described BVMOs from Pseudomonads, like HAPMO and OTEMO, from *P. fluorescens* ACB and *P. putida* NCIMB 10007, respectively.[110, 188] The former has 30% sequence identity with AcCHMO and was studied for the oxidation of a wide range of acetophenones, such as 4-hydroxyacetophenone, 4-aminoacetophenone and 4-hydroxypropiophenone. For these substrates, HAPMO has $k_{cat}$ values between 10 and 12 s$^{-1}$. This enzyme has also been reported to catalyze the oxidation of fluorobenzaldehydes, aryl ketones and sulfides.[100, 118, 215] OTEMO, on the other hand, is involved in the metabolic pathway of camphor and was described to oxidize the cyclopentanone derivative 2-oxo-$\Delta^3$-4,5,5-trimethyl-cyclopentenylacetyl-CoA. While it exhibits a rate of 4.8 s$^{-1}$ for its natural substrate, the free acid shows a rate 30 times lower than for the CoA ester.[57] OTEMO has been mostly studied for the conversions of substituted cyclohexanones, bicyclic ketones and terpenones.[57, 109, 216] Another BVMO from *Pseudomonas* is PpKT2440BVMO from *P. putida* KT2440.[217] This enzyme showed acceptance for aliphatic ketones, but exhibited

low conversions for cyclic and aryl ketones. The highest levels of oxidation were reported for 2-, 3- and 4-decanone (93-99% conversion using resting cells). Later, this enzyme was engineered for the whole cell biotransformation of ricinoleic acid into a precursor of polyamide-11 (nylon-11), achieving conversions of 85% and a product concentration up to 130 mM.[218-219]

The latest example is a BVMO from *Rhodoccocus pyridinivorans* DSM 44555.[189] RpBVMO exhibited affinity for aliphatic methyl ketones and highest activity on 2-hexanone ($k_{cat}$ = 2 s$^{-1}$). RpBVMO was able to regioselectively convert hexanones, octanones and methyl levulinate. The latter is a 2-ketone derived from renewable levulinic acid gained from biomass. Interestingly, the biocatalyst was used to fully convert 200 mM of this substrate to methyl 3-acetoxypropionate with a space-time yield of 5.4 g L$^{-1}$ h$^{-1}$. The hydrolyzed product, 3-hydroxypropionate is a platform chemical used as sugar-derived building block for biodegradable polymer polyester synthesis and is an important intermediate in the non-petrol based production of a variety of bulk chemicals.[220]

## Type II BVMOs

Type II BVMOs are categorized as class C flavoprotein monooxygenases, which are two-component monooxygenases. During the catalytic cycle, one component reduces FMN using NADH or NADPH as hydride donor. The flavin is then transferred by free diffusion to the second component, which uses reduced FMN as co-substrate for oxygen activation.[221] This is biochemically interesting because the free reduced FMN could lead to non-selective reactions with molecular oxygen inside the cell.[222] This group is less studied than type I BVMOs, perhaps due to the higher convenience to work with only one component. In addition to the early confusion with the actual prosthetic group, it was previously mistakenly believed to be a flavoprotein using tightly bound FMN as a coenzyme and that was reduced *in situ* in the active site by NADH.[223] There are some examples of type II BVMOs related to the metabolic pathway of the racemic monoterpene camphor. In particular, the enzymes involved are 2,5- and 3,6-diketocamphane monooxygenases (2,5-DKCMO and 3,6-DKCMO). These proteins are encoded on the linear inducible CAM plasmid from *P. putida* ATCC 17453 and were named after their natural substrates.[223] The presence of two isoforms in the same plasmid was described for 2,5-DKCMO, one being localized 23 kb downstream and encoded on the opposite strand.[164] It was suggested that the high sequence identity between them is the result of a gene duplication event and a sequence divergence in the case of 3,6-DKCMO. 2,5-DKCMO and 3,6-DKCMO oxidize the third metabolic step of the catabolism of *rac*-camphor and are specific towards one enantiomer. They specifically act on 2,5- and 3,6-diketocamphene, respectively. In recombinant cells expressing the oxygenating subunit of 2,5- or 3,6-DKCMO, activity without a recombinant FMN reductase component was noticed,

which was explained by the activity of native reductases from the host.[109] Later, Fred, a homodimeric reductase encoded in the chromosomal DNA of *P. putida* was suggested to be the *bona fide* reductase component for the three DKCMOs (3,6- and 2 isoforms of 2,5-DKCMO).[164] The complexes were tested against several substrates, exhibiting exclusive specificity for the natural substrate. Later, the structure of the oxygenating component of 3,6-DKCMO was solved in complex with FMN, and showed a fold most similar to the bacterial luciferase-like superfamily.[167] The structure was somewhat controversial due to experimental discrepancies.[222, 224] Other members of the type II BVMOs are luciferases from *Photobacterium phosphoreum* NCIMB 844 and *Vibrio fischeri* ATCC 7744.[225] These two-component bacterial luciferases catalyze Baeyer-Villiger reactions of 2-tridecanone, monocyclic and bicyclic ketones. In addition, it was suggested that an NADPH-dependent 6-oxocineole monooxygenase of *Rhodococcus* sp. C1 could also be part of this class.[226]

## Type O BVMOs

The best-studied BVMO of the type O—for atypical or 'odd' BVMOs—is MtmOIV from the soil actinomycete *Streptomyces argillaceus* ATCC 12956. The enzyme is a homodimer involved in the biosynthesis of mithramycin, an aureolic acid-like polyketide studied as an anticancer drug and calcium-lowering agent.[163, 227] This enzyme does not have significant sequence identity with other well-described BVMOs, does not display the consensus motifs for type I BVMOs, and bears no structural resemblance with type I or type II BVMOs. This monooxygenase catalyzes the Baeyer-Villiger oxidation of one of the four rings of premithramycin B, forming the lactone, which is later converted to mithramycin DK. As other BVMOs, MtmoIV uses NADPH and FAD as hydride donor and prosthetic group, respectively. The enzyme belongs to the class A flavoprotein monooxygenases, and it has been suggested that their reaction requires a peroxyflavin intermediate for nucleophilic attack, even though class A flavoprotein monooxygenases classically form a hydroperoxyflavin and proceed through an electrophilic attack.[49] The crystal structure was solved in complex with FAD, and the active site contains an arginine residue (R52) over the isoalloxazine ring, which presumably stabilizes the negatively charged peroxyflavin and Criegee intermediates.[168] While classic BVMOs contain a positively charged arginine on the *re* side of the flavin, MtmOIV's R52 is on the *si* side. Structurally, MtmOIV is most similar to *para*-hydroxybenzoate hydroxylase as well as the glucocorticoid receptors (GR2) subclass of FAD-dependent enzymes.[168, 228] Unsurprisingly, as MtmOIV catalyzes the oxidation of a bulky tetracyclic polyketide with deoxysugar modifications, it has a large binding pocket for the substrate, which may interact mostly by van der Waals and hydrophobic interactions.[195] Concerning the kinetic parameters, this enzyme displays relatively low activities in the presence of the natural substrate. Despite this, MtmOIV is interesting to investigate, as it might be a useful biocatalyst for the oxidation of analogues of premithramycin B and allow a synthetic route to new drugs.

## Flavin-containing monooxygenases

Flavin-containing monooxygenases (FMOs), like type I BVMOs are part of the class B flavoprotein monooxygenases, and are described to catalyze the oxidation of 'soft' nucleophilic heteroatoms in a broad spectrum of substrates.[229] FMOs are single-component enzymes, contain FAD as a prosthetic group, and have a preference for NADPH over NADH as FAD-reducing coenzyme.[49] For FMOs, two types have been described: type I FMOs are identified by the motif FxGxxxHxxx[Y/F], which is similar to the short consensus motif of type I BVMOs. Mammals, including humans, express five transmembrane FMO isoforms in a developmental-, sex-, and tissue-specific manner.[230] These enzymes are involved in the metabolism of xenobiotics such as drugs, pesticides, and certain dietary components.[111] While this group is described to oxidize mainly nitrogen and sulfur atoms, exceptions to this rule have been identified early on: for example, isoform FMO1 from pig liver was able to catalyze the Baeyer-Villiger oxidation of salicylaldehyde to pyrocatechol.[231] In addition, the human isoform FMO5, which expresses mostly in the small intestine, the kidney and the lung and has been described to exhibit poor activities on classic FMO substrates, is also able to catalyze Baeyer-Villiger oxidations. The enzyme was recombinantly expressed, and converted preferentially aliphatic ketones, but also aldehydes and cyclic ketones with varying regioselectivity.[111] Consequently, it was proposed that HsFMO5 could act as a possibly undescribed detoxification route in human metabolism. In this regard, it is remarkable that the enzyme can convert for example nabumetone and pentoxifylline (two ω-substituted 2-ketone drugs) and also a metabolite of E7016— a potential anticancer agent.[194] On the other hand, HsFMO5 was also described to have a high uncoupling rate, constituting for 60% of the activity. This phenomenon was ascribed to a low C4α-(hydro) peroxyflavin stabilization due to a weaker interaction with NADP⁺. Another group among the type I FMOs is formed by the YUCCAs,[145] which have a key role in the physiology of monocots and dicots plants. These enzymes catalyzes a rate-limiting step in *de novo* auxin biosynthesis, an essential growth hormone and development regulator.[157, 232] Eleven of the 29 putative FMOs in *Arabidopsis thaliana* belong to the YUCCA family, and one of them, AtYUC6, was described to catalyze the decarboxylation of indole-3-pyruvate to the auxin indole-3-acetate.[158] A sequence similarity network shows that YUCCAs are more related to FMOs than to BVMOs, even though the predicted mechanism is more related to the latter. As in the reaction of BVMOs, catalysis proceeds through a Criegee intermediate with a nucleophilic attack by the C4α-(hydro)peroxyflavin followed by a decarboxylation step producing the auxin. For AtYUC6, as for HsFMO5, a short-lived C4α-(hydro)peroxyflavin intermediate was measured.[111, 158] Additionally, a few enzymes that constitute the novel subclass of type II FMOs have been discovered in recent years. As the type I FMOs, this group can catalyze both heteroatom oxidations, as well as Baeyer-Villiger oxidations. Unlike the type I BVMOs, these enzymes cannot be identified by the long fingerprint sequence, but contain two Rossman fold motifs and exhibit the type I FMO motif FxGxxxHxxx[Y/F][K/R] with a few substitutions: a histidine instead of [Y/F]

and aspartate, proline, valine or glycine instead of [K/R].[160, 233] It was reported that these enzymes are promiscuous for the hydride donor, accepting either NADH or NADPH. This feature is attractive since the change of specificity for the cofactor of NADPH-dependent BVMOs is not a trivial task, as has been seen in studies of BVMO variants generated to identify residues related to the specificity for NADPH and the improvement of NADH catalytic efficiency.[79, 234-235] At present, there are some attempts to investigate this new group in more detail. Enzymes from *Pseudomonas stutzeri* NF13 (PsFMO), *Cellvibrio* sp. BR (CFMO) and *Stenotrophomonas maltophilia* PML168 (SmFMO) were studied. Although the kinetic parameters, conversion yields, enantioselectivities and substrate scope turned out to be poor, SmFMO displayed similar activities either with NADH or NADPH. For SmFMO the $K_m$ for the prototypic substrate bicyclo[3.2.0]hept-2-en-6-one was 40 times lower with NADH than with NADPH, and the conversion of the substrate was also considerably higher (90% vs 15%, respectively).[159] SmFMO was co-crystallized in complex with FAD, and it was suggested that the promiscuity is linked to the replacement of Arg234 and Thr235 as occurring in MaFMO—a related type I FMO from *Methylophaga aminisulfidivorans* —by a glutamine and a histidine (Gln193 and His194). However, the double mutant did not radically affect the cofactor specificity in SmFMO but the single mutant H194T caused a switch in cofactor preference from NADH to NADPH (mostly by reducing the $K_{m\,NADPH}$).[236] This effect was suggested to be related to the interaction of T235 with the ribose 2'-phosphate oxygen in MaFMO. Later, two novel proteins were found with variations of MaFMO's R234 and T235: CFMO and PsFMO, which share 58% and 61% sequence identity with SmFMO, respectively. [237] These enzymes were also described to accept NADH as a cofactor, but were mostly studied for asymmetric sulfoxidations. Another subgroup of type II FMOs, which features sequence alterations like an extension in the N-terminus, showed higher conversions and broader substrate scope for ketones. These include the FMOs from *R. jostii* RHA1, RjFMO-E, F and G,[160] and PsFMO-A, B and C from *Pimelobacter* sp. Bb-B.[233] RjFMO-E, F and G were found to be able to convert the classic substrate bicyclo[3.2.0]hept-2-en-6-one and cyclobutanones, but displayed only modest enantioselectivities and performed poorly in catalyzing the oxidation of phenylacetone. RjFMO-E displayed a higher affinity for NADPH, but also the affinity for NADH is in the micromolar range. Interestingly, the $k_{cat}$ for bicyclo[3.2.0]hept-2-en-6-one with NADH is higher than that with NADPH (4.3 vs. 2.7 s$^{-1}$) and almost 80 times higher than the reported $k_{cat}$ for SmFMO.[159-160] Finally, PsFMO A-C, three enzymes from a hydrocarbon-degrading bacterium were studied. These proteins show a sequence identity of 29-35% with RjFMO-E. PsFMO-A displayed the widest substrate scope and like the FMOs from *R. jostii* the highest activities were obtained with the ketones camphor and bicyclo[3.2.0]hept-2-en-6-one. High conversions were observed, but the enantioselectivities were only high for the normal lactone (>99% *ee* for the normal lactone and 57% *ee* for the abnormal lactone). More studies are expected for this class of enzymes, as their cofactor promiscuity constitutes a big potential in future biocatalysis. It remains

unknown whether or not NADH can fulfill the dual catalytic role described for NADPH in classical BVMOs—as hydride donor and stabilizer of the hydroperoxide flavin.

## BIOTECHNOLOGICAL APPLICATION

### Obstacles

The application of BVMOs is partially characterized as troublesome due to a number of important limiting factors, including: enzyme expression,[219] enzyme stability,[238] NADPH-dependence,[243-244] oxygen-dependence,[180] and substrate and product inhibition.[208] However, depending on the specific BVMO, there will be specific obstacles, e.g. some BVMOs have good expression, yet poor stability, or vice versa. In this subsection, we will discuss each of these limitations, and refer to studies that have addressed them. First of all, the application of BVMOs can be carried out in four different forms: with isolated enzymes, with immobilized enzymes, with crude/cell-free extract, or with whole cells. Most commonly, application-oriented reactions applied whole cells, as they provide a number of advantages: (1) improved stability of the enzymes due to the cellular environment,[20] (2) no addition of NADP(H) is needed, (3) co-expression of other enzymes can facilitate cofactor recycling or cascade reactions, (4) no cell lysis and enzyme purification steps are needed, and (5) it allows for continual expression of the enzyme(s). However, there are also some disadvantages with whole cells, such as: (1) mass balance issues and product removal,[245] (2) problematic oxygen supply to the cells,[180] (3) plasmid stability with requirement of antibiotic,[246-247] and (4) limited transport of substrates/products in and out of the cell.[248] In addition, a study on a cascade reaction *in vivo*, where a kinetic model was used to analyze performance, revealed that cofactor concentrations in the cell were limiting the reaction rate.[179] Possibly, this challenge could be addressed through metabolic engineering, or the use of a different host. Still, each of the ways to apply BVMOs has trade-offs, and it will be case-specific whether one is more suitable than the other. A recent mini-review addresses some of these aspects that are relevant for the development of a biocatalytic (industrial) process.[249]

### Industrial demand, TTN and stability

Most studies on BVMOs describe reactions on small lab-scale. Yet, to meet the demands of an industrial process, the limiting factors presented above need to be addressed. Specifically, to produce low-priced compounds, such as building blocks for polymers, a ratio of 2000-10000 g product / g (immobilized) enzyme (also referred to as 'biocatalyst loading') should be met in order to be an economically viable process.[20] To illustrate, assuming a 100 g mol$^{-1}$ product and a 50 kDa enzyme, 20–100 mol product / g enzyme, the demanded ratio translates to $1 \cdot 10^6$ – $5 \cdot 10^6$ total turnovers (TTN) per enzyme. Due to these numbers many BVMOs are still excluded from industrial application, unless the

target product is of high value, as is the case for pharmaceuticals, or if effort is invested to improve the biocatalyst and the process. In particular, improvement of the stability of the biocatalyst is needed in many cases, as is discussed in the subsection of 'Enzyme engineering' of this review. Moreover, an important issue for both the metric 'total turnover number' (TTN) and 'stability' is their inconsistent use in studies so far. Many studies refrain from determining the TTNs, and stability is often described in different ways, such as melting point and half-life, which makes it difficult to compare data. In addition, the stability of an enzyme in an industrial setting could be different compared to the lab setting. What could help the field of BVMO biocatalysis in general is if studies provide data for these characteristics, or at least a clear description of biocatalyst loading, because this metric gives an impression of the efficiency and stability of the biocatalyst, whether it is whole cells or isolated enzymes.

## NADPH recycling

While expression and stability are more related to enzyme engineering, the efficient use of NADPH is primarily determined by the way the BVMO is applied. The dependence of BVMOs on NADPH is an important challenge, since the cofactor is expensive: around 1800\$ / 5 g, compared to 280\$ / 5 g for NADH.[250] Therefore, it is necessary to minimize the amount of NADPH that is used. One way to address this challenge is to devise a set of reactions that are in redox balance. This can be achieved by combining an oxidation reaction, in which the reduced cofactor is formed, with the BVMO reaction, in which the reduced cofactor is oxidized again.[251] A typical example of such a redox-neutral reaction is the combination of a BVMO with an enzyme that can catalyze the oxidation of a sacrificial substrate, e.g. glucose dehydrogenase.[243-244] Alternatively, whole cells can be used for internal cofactor regeneration,[252-253] as well as non-enzymatic ways.[254] Nicotinamide cofactor regeneration strategies have been extensively reviewed elsewhere.[244] Another type of a redox-balanced pair of reactions are cascade reactions, where the product of the first reaction is the substrate for the second reaction.[255] One advantage of cascade reactions is that the isolation of intermediate products is not needed, as can be the case in other synthesis routes. In general, there are quite well-studied solutions to the challenge of cofactor recycling, with examples given in the next subsection.

## Examples of cascade strategies and novel BVMO applications

Over the past decades, some progress was made in optimizing large-scale reactions, employing strategies such as biphasic systems,[263] whole cell conversions,[252] and enzyme immobilization.[261, 264-265] Reviews focusing on biocatalysis with BVMOs from prior years are referred to for a broader overview.[201,266-268] Alongside these developments, several groups have explored different reactions and combinations of reactions with BVMOs, of which we present an overview, focused on studies from recent years. In particular, these combinations of reactions include cascades, as well as chemoenzymatic routes. To facilitate

cofactor recycling, an elegant strategy is to use a cascade reaction. For BVMOs, a frequently looked-at example is the cascade reaction with CHMO and an alcohol dehydrogenase (ADH), starting from cyclohexanol (**Scheme 5**, blue box). The alcohol oxidation generates NADPH and cyclohexanone, which is then oxidized by CHMO to ε-caprolactone. Several groups investigated and developed this cascade reaction.[208, 356, 269] Initially, problems were encountered concerning substrate and product inhibition. Higher levels of conversions could be achieved by keeping the substrate concentration low, through slow feeding, and removal of the lactone product by a subsequent polymerization/hydrolysis using a lipase such as CAL-A (**Scheme 5**).[270] This biocatalytic route was recently applied in whole cells that co-express CHMO and ADH on a 0.5 L scale, feeding of cyclohexanol, and addition of a lipase for hydrolysis of caprolactone to 6-hydroxyhexanoic acid.[271] After optimization, the process at 0.5 L scale could reach a product titer of 20 g $L^{-1}$, with an isolated yield of 81% of 6-hydroxyhexanoic acid. To address the cofactor balance, a different kind of cascade reaction was developed by Hollmann and Kara.[239] With the production of lactones in mind, an alcohol oxidation reaction of a linear diol was run in parallel in one pot with a Baeyer-Villiger reaction on a cyclic ketone catalyzed by CHMO (**Scheme 5**, red box).[239] As alcohol oxidation by an alcohol dehydrogenase (ADH) depends on NAD(P)$^+$ and produces NAD(P)H, combining this reaction with a BVMO or FMO reaction brings a redox balance. When one alcohol group of a linear diol becomes oxidized, the molecule undergoes cyclization to the hemiacetal or lactol. This lactol can be oxidized again to form a lactone (**Scheme 5**, red box). However, since the ADH generates two molecules of NAD(P)H in the conversion of one diol to one lactone, the substrate concentrations should be 2:1 of FMO substrate to ADH substrate.

This approach was termed convergent cascade, since two different substrates converge to the same product; the lactone. The Kara group later made an analogous combination, to produce γ-butyrolactones using RjFMO-E (**Scheme 6b**).[162] An interesting aspect of that study is that the FMO that was used to perform the Baeyer-Villiger reactions could accept NADH, making it a more feasible process compared to an NADPH-dependent reaction, considering the higher cost of NADPH[250] and its inferior stability.[275] A related strategy is to create a fusion of a cofactor recycling enzyme with a BVMO. This approach enables co-expression of both enzymes (as a fusion enzyme), and simplifies purification, whole cell conversions, and co-immobilization. Enzyme fusions with BVMOs have been reviewed recently[276] and thus will only briefly be discussed here. One recent study looked at fusions of three cofactor-regenerating enzymes with TmCHMO: glucose dehydrogenase (GDH), phosphite dehydrogenase and formate dehydrogenase (FDH).[277]
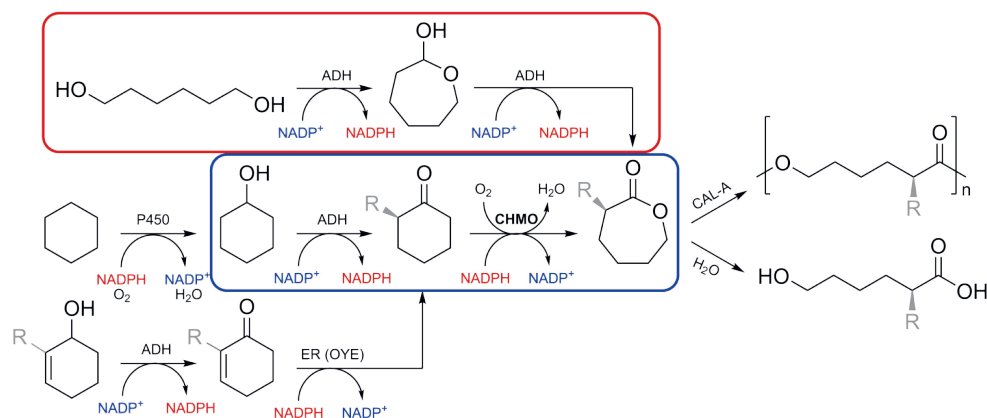
**Scheme 5.** Overview of biocatalyst combinations for cascades involving cyclohexanol and CHMO. Blue box: redox-neutral cascade from cyclohexanol to ε-caprolactone. Red box: ADH conversion of 1,6-hexanediol, which can be combined with cyclohexanone conversion by CHMO to recycle NADPH.[239] A cascade starting from cyclohexane involving a P450 monooxygenase was described.[272] Unsaturated cyclic alcohols or unsaturated cyclic ketones can be used with ene-reductase (ER) cascades, to make chiral lactones.[273-274]

These were compared in conversions and tested with various substrates and cosolvents, including a deep eutectic solvent (DES). Recently, following up on the convergent cascade (**Scheme 6b**),[162] fusions of the ADH and FMO were created to produce γ-butyrolactone in an unusual setup, using organic solvent.[278] Studies in the past have found that enzymes can actually be more stable and active in organic solvent, though the use was limited to lipases and esterases.[279] However, the ADH-FMO reaction is more challenging as it relies on NADH, which is why the authors chose to fuse the two enzymes. Cell-free extract from cells expressing the enzyme fusion was lyophilized, and subsequently added to organic media with 5% (v/v) water, to which the two substrates (diol and cyclic ketone) were added.[278] Although the yield was limited (27%), the fusion enzyme was able to perform the cascade reaction in this micro-aqueous media, and outperformed the combination of the separate enzymes. Moreover, no external NADH was added, which is appealing in terms of applications.

The approach of enzyme fusion is also very suitable for multi-enzyme cascade reactions. In some cases, the fusion outperforms the combination of separate enzymes, which is linked to an effect of the proximity of the enzymes called substrate channeling.[280-282] In 2013, Jeon et al. developed fusions of ADHs with BVMOs to convert hydroxy fatty acids into esters, in whole cells expressing the fusion enzyme (**Scheme 6a**).[262] The authors could demonstrate that the fused enzyme had a higher level of conversion for the cascade reaction. A similar pair of ADH with TmCHMO was fused to produce ε-caprolactone from cyclohexanol.[283] Although the fusion was more productive than the separate enzymes, substrate feeding and product removal through a lipase were needed to obtain full conversions, as was described previously (Scheme 5).[270]
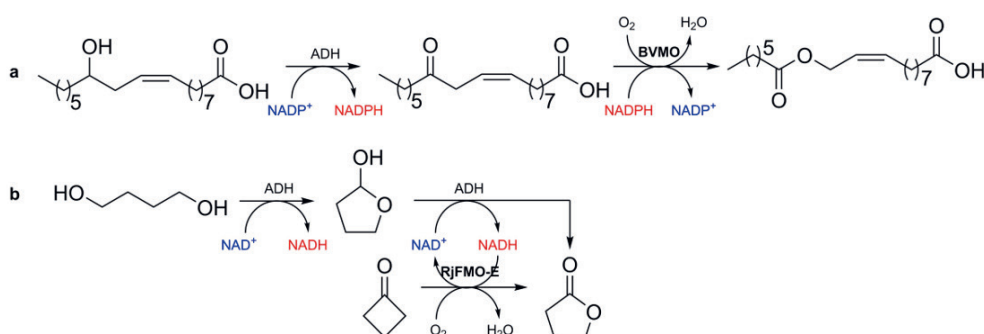
**Scheme 6.** Examples of cascade reactions involving BVMOs. a) Conversion of hydroxylated long-chain fatty acids to produce esters.[181, 218-219, 258-260] The cascade could also start from an unsaturated precursor with a hydratase to make the hydroxyl group,[258] or with a P450 to perform hydroxylation. b) Convergent cascade analogous to the reaction displayed in scheme 5, red box.[239] This particular cascade relies on NADH, through the use of RjFMO-E.[162] The same reaction was also used with fused enzymes in organic solvent.[278]

The development of strategies, like enzyme fusion, use of cosolvents, and cascade reactions have shown to be meaningful steps on a path toward biotechnological application. However, it is a path that still needs further exploration in order to meet the demands of an industrial process. The studies from the recent years show the variety of products that can be accessed through BVMOs. Given the limited turnover numbers that are reached in these studies, we conclude two things: (1) with the current state of BVMOs, any industrial application can only be economically feasible if the products are of high value (such as esomeprazole)[242], and/or through thorough optimization of the biocatalyst and process. (2) For the application of BVMOs for bulk chemicals (e.g. monomers) there are some examples,[257, 262] though more work needs to be done with respect to biocatalyst loading (in other words: operational stability and activity). So far, some studies have moved in the direction of biotechnological application, and shown to apply BVMOs for the synthesis of various compounds. As BVMOs become more suited, reliable, and recognized for biocatalytic application, it is likely that more groups and companies will look to harness the utility that these biocatalysts can provide. Though, to realize scaled-up applications, joint efforts will be needed that bring together different expertise, ranging from enzyme engineering to process development, to effectively tackle the specific challenges.

## CONCLUSIONS AND FUTURE DIRECTIONS

Biotechnology is at an exciting crossroad where ever more discoveries lead to the developments of applications in the various sub-disciplines that have (e)merged. Biocatalysis is maturing to a serious alternative to classical chemical transformations and this hopefully can contribute to a greener industry and new products at the same time.

Baeyer-Villiger monooxygenases are intriguing catalysts for a demanding reaction that allow unrivaled control of the reaction selectivity. Numerous variants have been described that feature activities suitable for countless reactions of synthetic value. Nevertheless, for flavoenzymologists, there are still some questions and challenges to unravel. Limitations, such as cofactor dependency, limited stability and undesired specificities are clearly identified and active research is making progress in overcoming these. However, an extended knowledge will be valuable for stability engineering, where seemingly distant mutations can sometimes abolish activity.[240, 241] And although the stability of BVMOs has been tackled, it can be doubted that this is enough to reach a broad application. However, with so many thermo- and hyperthermostable enzymes known from other enzyme families,[284] it seems fair to speculate that it is only a matter of time until a BVMO representative will be discovered as well. Furthermore, the puzzle of coenzyme promiscuity needs to be further cleared, which will also be essential in enabling the engineering of true independency on the dephosphorylated cofactor. Lastly, the stability of the peroxyflavin should be better investigated, as uncertainties about variations in the mode of uncoupling exist.[178] While the influencing factors are largely unknown and of academic interest, improvements in oxygenation coupling will also make biotechnological BVMO reactions more reliable, efficient—and thus—realistic.

## AIM AND OUTLINE OF THE THESIS

This thesis aimed to be a contribution to the flavoenzymology field by evaluating diverse biocatalytic features of several flavoproteins, with an emphasis on BVMOs. **Chapter 1** provides an overview on the catalytic properties of BVMOs.

In **Chapter 2**, an experimental protocol is described that allows the generation of a small library of fusion proteins with varying linkers. This was used to evaluate the effect of the length of a glycine-rich linker in a monooxygenase-dehydrogenase fusion: an alcohol dehydrogenase - cyclohexanone monooxygenase fusion.

**Chapter 3** describes a comprehensive study on the production of reduced oxygen species by flavoprotein oxidases and monooxygenases. Insight in which reduced oxygen species (hydrogen peroxide and superoxide) are formed and what conditions can influence their formation, is highly relevant when considering flavoenzymes for applications or understanding their physiological role. The production of hydrogen peroxide and superoxide radical was measured at different operational conditions for: phenylacetone monooxygenase, its C65D mutant (acting as a NADPH oxidase), eugenol oxidase and 5-hydroxymethylfurfural oxidase.

1

The second part of this thesis aimed at the discovery of novel flavoprotein monooxygenases. In **chapter 4,** in collaboration with the institute of microbiology of the ETH (Zürich), we described the characterization of Lmb-C and Oock. Both bacterial monooxygenases are part of biosynthetic routes towards specific secondary metabolites: lobatamide A and oocydin, respectively. In **chapter 5**, the flavin-containing monooxygenases (FMOs) HdFMO and CbFMO were identified and characterized. HdFMO is a type I FMO from *Hypsibius dujardini*. CbFMO was identified from a metagenomic environmental sample related to a *Chloroflexi* bacterium genome, and it was phylogenetically clustered to type II FMOs. Although these FMOs were related to well-described robust or promiscuous enzymes, their biocatalytic potential was rather modest. Finally, in the last chapter, we discovered by genome mining two type I BVMOs (**chapter 6**). The proteomic search was performed in the genome of the actinobacterium *Streptomyces leeuwenhoekii* C34, isolated from the Atacama Desert (Chile). The BVMOs (Sle_13190 and Sle_62070) were expressed fused to phosphite dehydrogenase (as cofactor regeneration system), and they were further characterized. Sle_62070 was found to be relatively thermostable and highly active. Specifically, this biocatalyst exhibited a broad acceptance for cyclic ketones and high regio- and enantioselectivity.

Overall, this thesis delivered new enzymes for the toolbox collection of known flavoproteins monooxygenases and insights into their molecular functioning as oxidative biocatalysts.

# SUPPORTING INFORMATION CHAPTER I

**Table S1.** List of the BVMOs used in the cladogram analysis (Figure 2)

| BVMO | Name | Organism | Uniprot ID |
|---|---|---|---|
| 2,5-DKCMO | 2,5-diketocamphane 1,2-MO 1 | *Pseudomonas putida* | Q6STM1 |
| 3,6-DKCMO | 3,6-diketocamphane 1,6-MO | *Pseudomonas putida* | D7UER1 |
| AcCHMO | Cyclohexanone 1,2-MO | *Acinetobacter calcoaceticus* | P12015 |
| AYUC6 | YUUCA flavin-containing MO | *Arabidopsis thaliana* | Q8VZ59 |
| BVMOAf1 | Baeyer-Villiger MO | *Aspergillus fumigatus* | Q4WAZ0 |
| BVMOAfl706 | Baeyer-Villiger MO | *Aspergillus flavus* | B8NCF3 |
| BVMOAfl838 | Baeyer-Villiger MO | *Aspergillus flavus* | B8N653 |
| CFMO | Flavin-containing MO | *Cellvibrio sp. BR* | I3IEE4 |
| CmBVMO | Baeyer-Villiger MO | *Cyanidioschyzon merolae* | M1VDM5 |
| CoCPMO | Cyclopentanone 1,2-MO | *Comamonas sp.* | Q8GAW0 |
| CrCAMO | Cycloalkanone MO | *Cylindrocarpon radicicola* | G8H1L8 |
| DiBVMO3 | Baeyer-Villiger MO | *Dietzia sp. D5* | A0A166N9M8 |
| DiBVMO4 | Baeyer-Villiger MO | *Dietzia sp. D5* | U5S003 |
| EthA | Flavin-containing MO | *Mycobacterium tuberculosis* | P9WNF9 |
| GoACMO | Acetone MO | *Gordonia sp. TY-5* | A0A3G5BIW4 |
| SsFMO1 | Flavin-containing MO 1 | *Sus scrofa* | P16549 |
| HsFMO2 | Flavin-containing MO 2 | *Homo sapiens* | Q99518 |
| HsFMO3 | Flavin-containing MO 3 | *Homo sapiens* | P31513 |
| HsFMO4 | Flavin-containing MO 4 | *Homo sapiens* | P31512 |
| HsFMO5 | Flavin-containing MO 5 | *Homo sapiens* | P49326 |
| LbBVMO | Baeyer-Villiger MO | *Leptospira biflexa* | B0SRK0 |
| PvMEKMO | Methyl ethyl ketone MO | *Pseudomonas veronii* | Q0MRG6 |
| SaMtmOIV | Mithramycin MO | *Streptomyces argillaceus* | Q194P4 |
| ObBVMO | Baeyer-Villiger MO | *Pseudooceanicola batsensis* | A3U3H1 |
| PfHAPMO | 4-Hydroxyacetophenone MO | *Pseudomonas fluorescens* | Q93TJ5 |
| PlBVMO | Baeyer-Villiger MO | *Parvibaculum lavamentivorans* | A7HU16 |
| TtPockeMO | Polycyclic ketone MO | *Thermothelomyces thermophila* | G2QA95 |
| PpOTEMO | 2-oxo-$\Delta^3$-4,5,5-trimethylcyclopentenylacetyl-CoA MO | *Pseudomonas putida* | H3JQW0 |
| PsCPDMO | Cyclopentadecanone MO | *Pseudomonas sp. HI-70* | T2HVF7 |
| PsFMO | Flavin-containing MO | *Pseudomonas stutzeri* | M2V3J0 |
| PsFMO_A | Flavin-containing MO | *Pimelobacter sp. Bb-B* | A0A0A1DMS0 |
| SaPtlE | Neopentalenolactone D synthase | *Streptomyces avermitilis* | Q82IY8 |
| RhCHMO | Cyclohexanone MO | *Rhodococcus sp. HI-31* | C0STX7 |

| | | | |
|---|---|---|---|
| RjBVMO24 | Baeyer-Villiger MO | *Rhodoccous jostii* | Q0S5T2 |
| RjBVMO4 | Baeyer-Villiger MO | *Rhodoccous jostii* | Q0SC70 |
| RjFMO_E | Flavin-containing MO | *Rhodoccous jostii* | Q0SIH9 |
| RjFMO_F | Flavin-containing MO | *Rhodoccous jostii* | Q0S8V1 |
| RjFMO_G | Flavin-containing MO | *Rhodoccous jostii* | Q0S4R1 |
| RpBVMO | Baeyer-Villiger MO | *Rhodococcus pyridinivorans* | A0A495NH77 |
| RpCHMO | Cyclohexanone MO | *Rhodococcus sp. Phi1* | Q84H73 |
| RrSTMO | Steroid MO | *Rhococcus rhodochrous* | O50641 |
| CtSAPMO | 4-Sulfoacetophenone MO | *Comamonas testosteroni* | B7X4D9 |
| SlBVMO | Baeyer-Villiger MO | *Streptomyces leeuwenhoekii* | A0A0F7W6X7 |
| SmFMO | Flavin-containing MO | *Stenotrophomonas maltophilia* | B2FRL2 |
| TfPAMO | Phenylacetone MO | *Thermobifida fusca* | Q47PU3 |
| TmCHMO | Cyclohexanone MO | *Thermocrispum municipale* | A0A1L1QK39 |

**1**

## REFERENCES

1. Rupe, H., *Adolf von Baeyer als Lehrer und Forscher. Erinnerungen aus seinem Privatlaboratorium.* Verlag von Ferdinand Enke: Stuttgart, 1932.

2. Walden, P. In *Geschichte der Organischen Chemie Seit 1880*, Walden, P., Ed. Springer: Berlin, Heidelberg, 1941; pp 892-906.

3. Stevens, T. S. Karl Johann Freudenberg. 29 January 1886-3 April 1983. *Biogr. Mem. Fellows Royal Soc.* 1984, 30, 169-189.

4. Fischer, E. Synthese des Traubenzuckers. *Ber. Dtsch. Chem. Ges.* 1890, 23 (1), 799-805.

5. Seebach, D. Organic synthesis—where now? *Angew. Chem. Int. Ed. Engl.* 1990, 29 (11), 1320-1367.

6. Yeh, B. J.; Lim, W. A. Synthetic biology: lessons from the history of synthetic organic chemistry. *Nat. Chem. Biol.* 2007, 3, 521-525.

7. Asimov, I., *A Short History of Chemistry.* Anchor Books: Garden City, New York, USA, 1965; p 263.

8. Erb, T. J. Back to the future: Why we need enzymology to build a synthetic metabolism of the future. *Beilstein J. Org. Chem.* 2019, 15 (1), 551-557.

9. United States Environmental Protection Agency, 2013 Toxics Release Inventory National Analysis. United States Environmental Protection Agency, Ed. Office of Environmental Information: Washington, DC, 2015; pp 1-78. https://www.epa.gov/toxics-release-inventory-tri-program/2013-tri-national-analysis-materials (accessed 10/09/2019)

10. Anastas, P.; Eghbali, N. Green chemistry: principles and practice. *Chem. Soc. Rev.* 2010, 39 (1), 301-312.

11. United States Energy Information Administration, International Energy Outlook. United States Department of Energy, Ed. United States Energy Information Administration: Washington DC, 2010; pp 1-328. https://www.eia.gov/outlooks/archive/ieo10/ (accessed 10/09/2019)

12. Tapon, F.; Sarabura, M. The greening of corporate strategy in the chemical industry: Two steps forward, one step back. *Strat. Change* 1995, 4 (6), 307-321.

13. Sanderson, K. Chemistry: It's not easy being green. *Nature News* 2011, 469 (7328), 18-20.

14. Williams, K.; Lee, E. Importance of Drug Enantiomers in Clinical Pharmacology. *Drugs* 1985, 30 (4), 333-354.

15. Arnold, F. H. Directed Evolution: Bringing New Chemistry to Life. *Angew. Chem. Int. Ed. Engl.* 2018, 57 (16), 4143-4148.

16. Reetz, M. T., *Directed evolution of selective enzymes: catalysts for organic chemistry and biotechnology.* John Wiley & Sons: Weinheim, 2016.

17. Wijma, H. J.; Janssen, D. B. Computational design gains momentum in enzyme catalysis engineering. *FEBS J.* 2013, 280 (13), 2948-2960.

18. Huang, P.-S.; Boyken, S. E.; Baker, D. The coming of age of de novo protein design. *Nature* 2016, 537 (7620), 320-327.

19. Simons, C.; Hanefeld, U.; Arends, I. W.; Maschmeyer, T.; Sheldon, R. A. Towards catalytic cascade reactions: asymmetric synthesis using combined chemo-enzymatic catalysts. *Top. Catal.* 2006, 40 (1), 35-44.

20. Tufvesson, P.; Lima-Ramos, J.; Nordblad, M.; Woodley, J. M. Guidelines and Cost Analysis for Catalyst Production in Biocatalytic Processes. *Org. Process Res. Dev.* 2011, 15 (1), 266-274.

21. Nakagawa, A.; Matsumura, E.; Koyanagi, T.; Katayama, T.; Kawano, N.; Yoshimatsu, K.; Yamamoto, K.; Kumagai, H.; Sato, F.; Minami, H. Total biosynthesis of opiates by stepwise fermentation using engineered *Escherichia coli. Nat. Commun.* 2016, 7, 10390-10397.

22. Galanie, S.; Thodey, K.; Trenchard, I. J.; Filsinger Interrante, M.; Smolke, C. D. Complete biosynthesis of opioids in yeast. *Science* 2015, 349 (6252), 1095-1100.

23. Luo, X.; Reiter, M. A.; d'Espaux, L.; Wong, J.; Denby, C. M.; Lechner, A.; Zhang, Y.; Grzybowski, A. T.; Harth, S.; Lin, W.; Lee, H.; Yu, C.; Shin, J.; Deng, K.; Benites, V. T.; Wang, G.; Baidoo, E. E. K.; Chen, Y.; Dev, I.; Petzold, C. J.; Keasling, J. D. Complete biosynthesis of cannabinoids and their unnatural analogues in yeast. *Nature* 2019, 567 (7746), 123-126.

24. Turner, N. J.; O'reilly, E. Biocatalytic retrosynthesis. *Nat. Chem. Biol.* 2013, 9 (5), 285-288.

25. Schwander, T.; Schada von Borzyskowski, L.; Burgener, S.; Cortina, N. S.; Erb, T. J. A synthetic pathway for the fixation of carbon dioxide *in vitro. Science* 2016, 354 (6314), 900-904.

26.   Paddon, C. J.; Westfall, P. J.; Pitera, D. J.; Benjamin, K.; Fisher, K.; McPhee, D.; Leavell, M. D.; Tai, A.; Main, A.; Eng, D.; Polichuk, D. R.; Teoh, K. H.; Reed, D. W.; Treynor, T.; Lenihan, J.; Jiang, H.; Fleck, M.; Bajad, S.; Dang, G.; Dengrove, D.; Diola, D.; Dorin, G.; Ellens, K. W.; Fickes, S.; Galazzo, J.; Gaucher, S. P.; Geistlinger, T.; Henry, R.; Hepp, M.; Horning, T.; Iqbal, T.; Kizer, L.; Lieu, B.; Melis, D.; Moss, N.; Regentin, R.; Secrest, S.; Tsuruta, H.; Vazquez, R.; Westblade, L. F.; Xu, L.; Yu, M.; Zhang, Y.; Zhao, L.; Lievense, J.; Covello, P. S.; Keasling, J. D.; Reiling, K. K.; Renninger, N. S.; Newman, J. D. High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* 2013, 496, 528-532.

27.   Erb, T. J.; Jones, P. R.; Bar-Even, A. Synthetic metabolism: metabolic engineering meets enzyme design. *Curr. Opin. Chem. Biol.* 2017, 37, 56-62.

28.   Krow, G. The Baeyer–Villiger Oxidation of Ketones and Aldehydes. *Org. React.* 1993, 43, 251-798.

29.   Renz, M.; Meunier, B. 100 Years of Baeyer–Villiger Oxidations. *Eur. J. Org. Chem.* 1999, (4), 737-750.

30.   Oxley, J.; Smith, J. In *Peroxide Explosives*, Detection and Disposal of Improvised Explosives, Dordrecht, Schubert, C. H. S., Kuznetsov, A., Eds. Springer Netherlands: Dordrecht, 2006; pp 113-121.

31.   Clark, D. E. Peroxides and peroxide-forming compounds. *Chem. Health Saf.* 2001, 8 (5), 12-22.

32.   Strukul, G. Transition metal catalysis in the Baeyer–Villiger oxidation of ketones. *Angew. Chem. Int. Ed. Engl.* 1998, 37 (9), 1198-1209.

33.   Bryliakov, K. P. Catalytic asymmetric oxygenations with the environmentally benign oxidants $H_2O_2$ and $O_2$. *Chem. Rev.* 2017, 117 (17), 11406-11459.

34.   Strukul, G.; Scarso , A. In *Applied Homogeneous Catalysis with Organometallic Compounds*, Cornils, B., Herrmann, W. A., Beller, M., Paciello, R., Eds. 2017; pp 1485-1508.

35.   Featherston, A. L.; Shugrue, C. R.; Mercado, B. Q.; Miller, S. J. Phosphothreonine (pThr)-based multifunctional peptide catalysis for asymmetric baeyer–villiger oxidations of cyclobutanones. *ACS Catal.* 2019, 9 (1), 242-252.

36.   Forney, F. W.; Markovetz, A. J.; Kallio, R. E. Bacterial oxidation of 2-tridecanone to 1-undecanol. *J. Bacteriol.* 1967, 93 (2), 649-655.

37.   Forney, F.; Markovetz, A. An enzyme system for aliphatic methyl ketone oxidation. *Biochem. Biophys. Res. Commun.* 1969, 37 (1), 31-38.

38.   Norris, D.; Trudgill, P. The metabolism of cyclohexanol by *Nocardia globerula* CL1. *Biochem. J.* 1971, 121 (3), 363-370.

39.   Donoghue, N. A.; Norris, D. B.; Trudgill, P. W. The purification and properties of cyclohexanone oxygenase from *Nocardia globerula* CL1 and *Acinetobacter* NCIB 9871. *Eur. J. Biochem.* 1976, 63 (1), 175-192.

40.   Walsh, C. T.; Chen, Y. C. J. Enzymic Baeyer–Villiger oxidations by flavin-dependent monooxygenases. *Angew. Chem. Int. Ed. Engl.* 1988, 27, 333-343.

41.   Abril, O.; Ryerson, C. C.; Walsh, C.; Whitesides, G. M. Enzymatic Baeyer–Villiger type oxidations of ketones catalyzed by cyclohexanone oxygenase. *Bioorg. Chem.* 1989, 17 (1), 41-52.

42.   Zhang, Y.; Wu, Y.-Q.; Xu, N.; Zhao, Q.; Yu, H.-L.; Xu, J.-H. Engineering of cyclohexanone monooxygenase for the enantioselective synthesis of (S)-omeprazole. *ACS Sustain. Chem. Eng.* 2019, 7 (7), 7218-7226.

43.   Malito, E.; Alfieri, A.; Fraaije, M. W.; Mattevi, A. Crystal structure of a Baeyer–Villiger monooxygenase. *Proc. Natl. Acad. Sci. U. S. A.* 2004, 101, 13157–13162.

44.   Doukyu, N.; Ogino, H. Organic solvent-tolerant enzymes. *Biochem. Eng. J.* 2010, 48 (3), 270-282.

45.   Secundo, F.; Fiala, S.; Fraaije, M. W.; de Gonzalo, G.; Meli, M.; Zambianchi, F.; Ottolina, G. Effects of water miscible organic solvents on the activity and conformation of the Baeyer–Villiger monooxygenases from *Thermobifida fusca* and *Acinetobacter calcoaceticus*: a comparative study. *Biotechnol. Bioeng.* 2011, 108 (3), 491-499.

46.   Romero, E.; Castellanos, J. R.; Mattevi, A.; Fraaije, M. W. Characterization and Crystal Structure of a Robust Cyclohexanone Monooxygenase. *Angew. Chem. Int. Ed. Engl.* 2016, 55 (51), 15852-15855.

47.   Fürst, M. J. L. J.; Savino, S.; Dudek, H. M.; Gomez Castellanos, J. R.; Gutierrez de Souza, C.; Rovida, S.; Fraaije, M. W.; Mattevi, A. Polycyclic ketone monooxygenase from the thermophilic fungus *Thermothelomyces thermophila*: A structurally distinct biocatalyst for bulky substrates. *J. Am. Chem. Soc.* 2017, 139 (2), 627-630.

48. Kamerbeek, N. M.; van der Ploeg, R.; Olsthoorn, A. J.; Tahallah, N.; Heck, A. J.; Malito, E.; Fraaije, M. W.; Janssen, D. B. Exploring the role of the N-terminal domain of 4-hydroxyacetophenone monooxygenase. PhD Thesis, University of Groningen, Groningen, 2004.

49. van Berkel, W. J.; Kamerbeek, N. M.; Fraaije, M. W. Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts. *J. Biotechnol.* 2006, 124 (4), 670-689.

50. Fürst, M. J. L. J.; Fiorentini, F.; Fraaije, M. W. Beyond active site residues: overall structural dynamics control catalysis in flavin-containing and heme-containing monooxygenases. *Curr. Opin. Struct. Biol.* 2019, 59, 29-37.

51. Ferroni, F. M.; Tolmie, C.; Smit, M. S.; Opperman, D. J. Structural and catalytic characterization of a fungal Baeyer–Villiger monooxygenase. *PLoS One* 2016, 11 (7), e0160186.

52. Mirza, I. A.; Yachnin, B. J.; Wang, S.; Grosse, S.; Bergeron, H.; Imura, A.; Iwaki, H.; Hasegawa, Y.; Lau, P. C.; Berghuis, A. M. Crystal structures of cyclohexanone monooxygenase reveal complex domain movements and a sliding cofactor. *J. Am. Chem. Soc.* 2009, 131 (25), 8848-8854.

53. Yachnin, B. J.; McEvoy, M. B.; MacCuish, R. J.; Morley, K. L.; Lau, P. C.; Berghuis, A. M. Lactone-bound structures of cyclohexanone monooxygenase provide insight into the stereochemistry of catalysis. *ACS Chem. Biol.* 2014, 9 (12), 2843-2851.

54. Yachnin, B. J.; Sprules, T.; McEvoy, M. B.; Lau, P. C.; Berghuis, A. M. The substrate-bound crystal structure of a Baeyer–Villiger monooxygenase exhibits a Criegee-like conformation. *J. Am. Chem. Soc.* 2012, 134 (18), 7788-7795.

55. Messiha, H. L.; Ahmed, S. T.; Karuppiah, V.; Suardíaz, R.; Ascue Avalos, G. A.; Fey, N.; Yeates, S.; Toogood, H. S.; Mulholland, A. J.; Scrutton, N. S. Biocatalytic routes to lactone monomers for polymer production. *Biochemistry* 2018, 57 (13), 1997-2008.

56. Fürst, M. J. L. J.; Romero, E.; Gómez Castellanos, J. R.; Fraaije, M. W.; Mattevi, A. Side-chain pruning has limited impact on substrate preference in a promiscuous enzyme. *ACS Catal.* 2018, 8 (12), 11648-11656.

57. Leisch, H.; Shi, R.; Grosse, S.; Morley, K.; Bergeron, H.; Cygler, M.; Iwaki, H.; Hasegawa, Y.; Lau, P. C. Cloning, Baeyer–Villiger biooxidations, and structures of the camphor pathway 2-oxo-Delta$^3$-4,5,5-trimethylcyclopentenylacetyl-coenzyme A monooxygenase of *Pseudomonas putida* ATCC 17453. *Appl. Environ. Microbiol.* 2012, 78 (7), 2200-2212.

58. Martinoli, C.; Dudek, H. M.; Orru, R.; Edmondson, D. E.; Fraaije, M. W.; Mattevi, A. Beyond the protein matrix: probing cofactor variants in a Baeyer–Villiger oxygenation reaction. *ACS Catal.* 2013, 3 (12), 3058-3062.

59. Orru, R.; Dudek, H. M.; Martinoli, C.; Torres Pazmiño, D. E.; Royant, A.; Weik, M.; Fraaije, M. W.; Mattevi, A. Snapshots of enzymatic Baeyer–Villiger catalysis: oxygen activation and intermediate stabilization. *J. Biol. Chem.* 2011, 286 (33), 29284-29291.

60. Nguyen, T. D.; Choi, G.-E.; Gu, D.-H.; Seo, P.-W.; Kim, J.-W.; Park, J.-B.; Kim, J.-S. Structural basis for the selective addition of an oxygen atom to cyclic ketones by Baeyer–Villiger monooxygenase from *Parvibaculum lavamentivorans*. *Biochem. Biophys. Res. Commun.* 2019, 512 (3), 564-570.

61. Franceschini, S.; van Beek, H. L.; Pennetta, A.; Martinoli, C.; Fraaije, M. W.; Mattevi, A. Exploring the structural basis of substrate preferences in Baeyer–Villiger monooxygenases: insight from steroid monooxygenase. *J. Biol. Chem.* 2012, 287 (27), 22626-22634.

62. Ryerson, C. C.; Ballou, D. P.; Walsh, C. Mechanistic studies on cyclohexanone oxygenase. *Biochemistry* 1982, 21 (11), 2644-2655.

63. Ghisla, S.; Massey, V. Mechanisms of flavoprotein-catalyzed reactions. *Eur. J. Biochem.* 1989, 181 (1), 1-17.

64. Entsch, B.; Massey, V.; Ballou, D. P. Intermediates in flavoprotein catalyzed hydroxylations. *Biochem. Biophys. Res. Commun.* 1974, 57 (4), 1018-1025.

65. Ghisla, S.; Hastings, J. W.; Favaudon, V.; Lhoste, J.-M. Structure of the oxygen adduct intermediate in the bacterial luciferase reaction: $^{13}$C nuclear magnetic resonance determination. *Proc. Natl. Acad. Sci. U. S. A.* 1978, 75 (12), 5860-5863.

66. Sheng, D.; Ballou, D. P.; Massey, V. Mechanistic studies of cyclohexanone monooxygenase: chemical properties of intermediates involved in catalysis. *Biochemistry* 2001, 40 (37), 11156-11167.

67. Massey, V. Activation of molecular oxygen by flavins and flavoproteins. *J. Biol. Chem.* 1994, 269 (36), 22459-22462.

68. Jones, K. C.; Ballou, D. P. Reactions of the 4a-hydroperoxide of liver microsomal flavin-containing monooxygenase with nucleophilic and electrophilic substrates. *J. Biol. Chem.* 1986, 261 (6), 2553-2559.

69. Kadkhodayan, S.; Coulter, E. D.; Maryniak, D. M.; Bryson, T. A.; Dawson, J. H. Uncoupling Oxygen Transfer and Electron Transfer in the Oxygenation of Camphor Analogues by Cytochrome P450-CAM: direct observation of an intermolecular isotope effect for substrate c-h activation. *J. Biol. Chem.* 1995, 270 (47), 28042-28048.

70. Entsch, B.; Palfey, B. A.; Ballou, D. P.; Massey, V. Catalytic function of tyrosine residues in *para*-hydroxybenzoate hydroxylase as determined by the study of site-directed mutants. *J. Biol. Chem.* 1991, 266 (26), 17341-17349.

71. Holtmann, D.; Hollmann, F. The Oxygen Dilemma: A Severe Challenge for the Application of Monooxygenases? *ChemBioChem* 2016, 17 (15), 1391-1398.

72. Sucharitakul, J.; Prongjit, M.; Haltrich, D.; Chaiyen, P. Detection of a C4a-Hydroperoxyflavin Intermediate in the Reaction of a Flavoprotein Oxidase. *Biochemistry* 2008, 47 (33), 8485-8490.

73. Beaty, N. B.; Ballou, D. P. The oxidative half-reaction of liver microsomal FAD-containing monooxygenase. *J. Biol. Chem.* 1981, 256 (9), 4619-4625.

74. Mayfield, J. A.; Frederick, R. E.; Streit, B. R.; Wencewicz, T. A.; Ballou, D. P.; DuBois, J. L. Comprehensive Spectroscopic, Steady State, and Transient Kinetic Studies of a Representative Siderophore-associated Flavin Monooxygenase. *J. Biol. Chem.* 2010, 285 (40), 30375-30388.

75. Becvar, J. E.; Tu, S.-C.; Hastings, J. Activity and stability of the luciferase-flavin intermediate. *Biochemistry* 1978, 17 (9), 1807-1812.

76. Torres Pazmiño, D. E.; Baas, B. J.; Janssen, D. B.; Fraaije, M. W. Kinetic mechanism of phenylacetone monooxygenase from *Thermobifida fusca*. *Biochemistry* 2008, 47 (13), 4082-4093.

77. Polyak, I.; Reetz, M. T.; Thiel, W. Quantum mechanical/molecular mechanical study on the mechanism of the enzymatic Baeyer–Villiger reaction. *J. Am. Chem. Soc.* 2012, 134 (5), 2732-2741.

78. Sucharitakul, J.; Wongnate, T.; Chaiyen, P. Hydrogen peroxide elimination from C4a-hydroperoxy-flavin in a flavoprotein oxidase occurs through a single proton transfer from flavin N5 to a peroxide leaving group. *J. Biol. Chem.* 2011, 286 (19), 16900-16909.

79. Kamerbeek, N. M.; Fraaije, M. W.; Janssen, D. B. Identifying determinants of NADPH specificity in Baeyer–Villiger monooxygenases. *Eur. J. Biochem.* 2004, 271 (11), 2107-2116.

80. Doering, W. v. E.; Dorfman, E. Mechanism of the Peracid Ketone—Ester Conversion. Analysis of Organic Compounds for Oxygen-18. *J. Am. Chem. Soc.* 1953, 75 (22), 5595-5598.

81. Criegee, R. Die Umlagerung der Dekalin-peroxydester als Folge von kationischem Sauerstoff. *Liebigs Ann.* 1948, 560 (1), 127-135.

82. Grein, F.; Chen, A. C.; Edwards, D.; Crudden, C. M. Theoretical and Experimental Studies on the Baeyer–Villiger Oxidation of Ketones and the Effect of α-Halo Substituents. *J. Org. Chem.* 2006, 71 (3), 861-872.

83. Alvarez-Idaboy, J. R.; Reyes, L.; Cruz, J. A new specific mechanism for the acid catalysis of the addition step in the Baeyer–Villiger rearrangement. *Org. Lett.* 2006, 8 (9), 1763-1765.

84. Sever, R. R.; Root, T. W. Computational study of tin-catalyzed Baeyer–Villiger reaction pathways using hydrogen peroxide as oxidant. *J. Phys. Chem. B* 2003, 107 (39), 10848-10862.

85. Carlqvist, P.; Eklund, R.; Brinck, T. A Theoretical Study of the Uncatalyzed and $BF_3$-Assisted Baeyer–Villiger Reactions. *J. Org. Chem.* 2001, 66 (4), 1193-1199.

86. Vil', V. A.; dos Passos Gomes, G.; Bityukov, O. V.; Lyssenko, K. A.; Nikishin, G. I.; Alabugin, I. V.; Terent'ev, A. O. Interrupted Baeyer–Villiger rearrangement: building a stereoelectronic trap for the Criegee intermediate. *Angew. Chem. Int. Ed. Engl.* 2018, 130 (13), 3430-3434.

87. Polyak, I.; Reetz, M. T.; Thiel, W. Quantum mechanical/molecular mechanical study on the enantioselectivity of the enzymatic Baeyer–Villiger reaction of 4-hydroxycyclohexanone. *J. Phys. Chem. B* 2013, 117 (17), 4993-5001.

88. Rozzell Jr, J. D.; Benner, S. A. Combining enzymatic and chemical steps in the synthesis of biochemically valuable compounds: isotopically chiral methyl acetate. *J. Org. Chem.* 1983, 48 (8), 1190-1193.

89. Turner, R. B. Stereochemistry of the Peracid Oxidation of Ketones. *J. Am. Chem. Soc.* 1950, 72 (2), 878-882.

90. Alvarez-Idaboy, J. R.; Reyes, L.; Mora-Diez, N. The mechanism of the Baeyer–Villiger rearrangement: quantum chemistry and TST study supported by experimental kinetic data. *Org. Biomol. Chem.* 2007, 5 (22), 3682-3689.

91.  Doering, W. v. E.; Speers, L. The Peracetic Acid Cleavage of Unsymmetrical Ketones. *J. Am. Chem. Soc.* 1950, 72 (12), 5515-5518.

92.  Crudden, C. M.; Chen, A. C.; Calhoun, L. A. A Demonstration of the Primary Stereoelectronic Effect in the Baeyer–Villiger Oxidation of α-Fluorocyclohexanones. *Angew. Chem. Int. Ed. Engl.* 2000, 112 (16), 2973-2977.

93.  Chandrasekhar, S.; Deo Roy, C. Evidence for a stereoelectronic effect in the Baeyer–Villiger reaction: Introducing the intramolecular reaction. *Tetrahedron Lett.* 1987, 28 (50), 6371-6372.

94.  Noyori, R.; Kobayashi, H.; Sato, T. Remote substituent effects in the Baeyer–Villiger oxidation  II regioselection based on the hydroxyl group orientation in the tetrahedral intermediate. *Tetrahedron Lett.* 1980, 21 (26), 2573-2576.

95.  Itoh, Y.; Yamanaka, M.; Mikami, K. Theoretical study on the regioselectivity of Baeyer–Villiger reaction of α-Me-, -F-, -CF$_3$-cyclohexanones. *J. Org. Chem.* 2013, 78 (1), 146-153.

96.  Sauers, R.; Ahearn, G. The importance of steric effects in the Baeyer–Villiger oxidation. *J. Am. Chem. Soc.* 1961, 83 (12), 2759-2762.

97.  Bürgi, H.; Dunitz, J.; Shefter, E. Chemical reaction paths. IV. Aspects of O···C=O interactions in crystals. *Acta Crystallogr. B Struct. Cryst. Cryst. Chem.* 1974, 30 (6), 1517-1527.

98.  Branchaud, B. P.; Walsh, C. T. Functional group diversity in enzymic oxygenation reactions catalyzed by bacterial flavin-containing cyclohexanone oxygenase. *J. Am. Chem. Soc.* 1985, 107 (7), 2153-2161.

99.  Ferroni, F. M.; Tolmie, C.; Smit, M. S.; Opperman, D. J. Alkyl formate ester synthesis by a fungal Baeyer–Villiger monooxygenase. *ChemBioChem* 2017, 18 (6), 515-517.

100. Moonen, M. J.; Westphal, A. H.; Rietjens, I. M.; van Berkel, W. J. Enzymatic Baeyer–Villiger oxidation of benzaldehydes. *Adv. Synth. Catal.* 2005, 347 (7-8), 1027-1034.

101. Bisagni, S.; Summers, B.; Kara, S.; Hatti-Kaul, R.; Grogan, G.; Mamo, G.; Hollmann, F. Exploring the substrate specificity and enantioselectivity of a Baeyer–Villiger monooxygenase from *Dietzia* sp. D5: oxidation of sulfides and aldehydes. *Top. Catal.* 2014, 57, 366.

102. Torres Pazmiño, D. E.; Snajdrova, R.; Rial, D. V.; Mihovilovic, M. D.; Fraaije, M. W. Altering the substrate specificity and enantioselectivity of phenylacetone monooxygenase by structure-inspired enzyme redesign. *Adv. Synth. Catal.* 2007, 349 (8-9), 1361-1368.

103. Kamerbeek, N. M.; Olsthoorn, A. J.; Fraaije, M. W.; Janssen, D. B. Substrate specificity and enantioselectivity of 4-hydroxyacetophenone monooxygenase. *Appl. Environ. Microbiol.* 2003, 69 (1), 419-426.

104. Kumar, H.; Fraaije, M. W. Conversion of furans by Baeyer–Villiger monooxygenases. *Catalysts* 2017, 7 (6), 179.

105. Mendelovici, M.; Glotter, E. Epoxidation and Baeyer–Villiger oxidation of γ-hydroxy-αβ-unsaturated ketones on exposure to *m*-chloroperbenzoic acid. *Perkin Trans.* 1992,  (13), 1735-1740.

106. Reignier, T.; de Berardinis, V.; Petit, J. L.; Mariage, A.; Hamze, K.; Duquesne, K.; Alphand, V. Broadening the scope of Baeyer–Villiger monooxygenase activities toward α,β-unsaturated ketones: a promising route to chiral enol-lactones and ene-lactones. *Chem. Comm.* 2014, 50 (58), 7793-7796.

107. Bes, M. T.; Villa, R.; Roberts, S. M.; Wan, P. W.; Willetts, A. Oxidative biotransformations by microorganisms: production of chiral synthons by cyclopentanone monooxygenase from *Pseudomonas* sp. NCIMB 9872. *J. Mol. Catal. B Enzym.* 1996, 1 (3-6), 127-134.

108. van der Werf, M. J. Purification and characterization of a Baeyer–Villiger mono-oxygenase from *Rhodococcus erythropolis* DCL14 involved in three different monocyclic monoterpene degradation pathways. *Biochem. J.* 2000, 347 (3), 693-701.

109. Kadow, M.; Loschinski, K.; Saß, S.; Schmidt, M.; Bornscheuer, U. T. Completing the series of BVMOs involved in camphor metabolism of *Pseudomonas putida* NCIMB 10007 by identification of the two missing genes, their functional expression in *E. coli*, and biochemical characterization. *Appl. Microbiol. Biotechnol.* 2012, 96 (2), 419-429.

110. Ougham, H. J.; Taylor, D. G.; Trudgill, P. W. Camphor revisited: involvement of a unique monooxygenase in metabolism of 2-oxo-delta 3-4,5,5-trimethylcyclopentenylacetic acid by *Pseudomonas putida*. *J. Bacteriol.* 1983, 153 (1), 140-52.

111. Fiorentini, F.; Geier, M.; Binda, C.; Winkler, M.; Faber, K.; Hall, M.; Mattevi, A. Biocatalytic characterization of human FMO5: unearthing Baeyer–Villiger reactions in humans. *ACS Chem. Biol.* 2016, 11 (4), 1039-1048.

112. Hu, Y.; Dietrich, D.; Xu, W.; Patel, A.; Thuss, J. A.; Wang, J.; Yin, W. B.; Qiao, K.; Houk, K. N.; Vederas, J. C.; Tang, Y. A carbonate-forming Baeyer–Villiger monooxygenase. *Nat. Chem. Biol.* 2014, 10 (7), 552-554.

113. Drabowicz, J.; Mikołajczyk, M. Synthesis Of Sulfoxides. A Review. *Org. Prep. Proced. Int.* 1982, 14 (1-2), 45-89.

114. Light, D. R.; Waxman, D. J.; Walsh, C. Studies on the chirality of sulfoxidation catalyzed by bacterial flavoenzyme cyclohexanone monooxygenase and hog liver flavin adenine dinucleotide containing monooxygenase. *Biochemistry* 1982, 21 (10), 2490-2498.

115. Rioz-Martínez, A.; de Gonzalo, G.; Pazmiño, D. E. T.; Fraaije, M. W.; Gotor, V. Enzymatic Synthesis of Novel Chiral Sulfoxides Employing Baeyer–Villiger Monooxygenases. *Eur. J. Org. Chem.* 2010, (33), 6409-6416.

116. Gonzalo, G. d.; Pazmiño, D. E. T.; Ottolina, G.; Fraaije, M. W.; Carrea, G. Oxidations catalyzed by phenylacetone monooxygenase from *Thermobifida fusca*. *Tetrahedron: Asymmetry* 2005, 16 (18), 3077-3083.

117. de Gonzalo, G.; Fürst, M. J. L. J.; Fraaije, M. W. Polycyclic ketone monooxygenase (PockeMO): A robust biocatalyst for the synthesis of optically active sulfoxides. *Catalysts* 2017, 7 (10), 288.

118. de Gonzalo, G.; Torres Pazmiño, D. E.; Ottolina, G.; Fraaije, M. W.; Carrea, G. 4-Hydroxyacetophenone monooxygenase from *Pseudomonas fluorescens* ACB as an oxidative biocatalyst in the synthesis of optically active sulfoxides. *Tetrahedron: Asymmetry* 2006, 17 (1), 130-135.

119. Colonna, S.; Gaggero, N.; Carrea, G.; Pasta, P.; Alphand, V.; Furstoss, R. Enantioselective synthesis of *tert*-butyl *tert*-butanethiosulfinate catalyzed by cyclohexanone monooxygenase. *Chirality* 2001, 13 (1), 40-42.

120. Fraaije, M. W.; Kamerbeek, N. M.; Heidekamp, A. J.; Fortin, R.; Janssen, D. B. The prodrug activator EtaA from Mycobacterium tuberculosis is a Baeyer–Villiger monooxygenase. *J. Biol. Chem.* 2004, 279 (5), 3354-3360.

121. de Gonzalo, G.; Franconetti, A. Enantioselective sulfoxidations employing the thermostable cyclohexanone monooxygenase from *Thermocrispum municipale*. *Enzyme Microb. Technol.* 2018, 113, 24-28.

122. Carrea, G.; Redigolo, B.; Riva, S.; Colonna, S.; Gaggero, N.; Battistel, E.; Bianchi, D. Effects of substrate structure on the enantioselectivity and stereochemical course of sulfoxidation catalyzed by cyclohexanone monooxygenase. *Tetrahedron: Asymmetry* 1992, 3 (8), 1063-1068.

123. Zhang, Y.; Liu, F.; Xu, N.; Wu, Y.-Q.; Zheng, Y.-C.; Zhao, Q.; Lin, G.; Yu, H.-L.; Xu, J.-H. Discovery of two native Baeyer–Villiger monooxygenases for asymmetric synthesis of bulky chiral sulfoxides. *Appl. Environ. Microbiol.* 2018, 84 (14), e00638-18.

124. Bordewick, S.; Beier, A.; Balke, K.; Bornscheuer, U. T. Baeyer–Villiger monooxygenases from *Yarrowia lipolytica* catalyze preferentially sulfoxidations. *Enzyme Microb. Technol.* 2018, 109, 31-42.

125. Reetz, M. T.; Daligault, F.; Brunner, B.; Hinrichs, H.; Deege, A. Directed evolution of cyclohexanone monooxygenases: enantioselective biocatalysts for the oxidation of prochiral thioethers. *Angew. Chem. Int. Ed. Engl.* 2004, 116 (31), 4170-4173.

126. Ang, E. L.; Alvizo, O.; Behrouzian, B.; Clay, M.; Collier, S.; Eberhard, E.; Fu, F. J.; Song, S.; Smith, D.; Widegren, M. Biocatalysts and methods for the synthesis of armodafinil. US9267159B2, 2016.

127. Bong, Y. K.; Clay, M. D.; Collier, S. J.; Mijts, B.; Vogel, M.; Zhang, X.; Zhu, J.; Nazor, J.; Smith, D.; Song, S. Synthesis of prazole compounds. US9631181B2, 2011.

128. Hughes, D. L. Biocatalysis in drug development—highlights of the recent patent literature. *Org. Process Res. Dev.* 2018, 22 (9), 1063-1080.

129. Ottolina, G.; Bianchi, S.; Belloni, B.; Carrea, G.; Danieli, B. First asymmetric oxidation of tertiary amines by cyclohexanone monooxygenase. *Tetrahedron Lett.* 1999, 40 (48), 8483-8486.

130. Colonna, S.; Pironti, V.; Pasta, P.; Zambianchi, F. Oxidation of amines catalyzed by cyclohexanone monooxygenase. *Tetrahedron Lett.* 2003, 44 (4), 869-871.

131. Brondani, P. B.; de Gonzalo, G.; Fraaije, M. W.; Andrade, L. H. Selective oxidations of organoboron compounds catalyzed by Baeyer–Villiger monooxygenases. *Adv. Synth. Catal.* 2011, 353 (11-12), 2169-2173.

**1**

132. Brondani, P. B.; Dudek, H.; Reis, J. S.; Fraaije, M. W.; Andrade, L. H. Exploiting the enantioselectivity of Baeyer–Villiger monooxygenases via boron oxidation. *Tetrahedron: Asymmetry* 2012, 23 (9), 703-708.

133. Brondani, P. B.; Guilmoto, N. M. A. F.; Dudek, H. M.; Fraaije, M. W.; Andrade, L. H. Chemoenzymatic approaches to obtain chiral-centered selenium compounds. *Tetrahedron* 2012, 68 (51), 10431-10436.

134. Andrade, L. H.; Pedrozo, E. C.; Leite, H. G.; Brondani, P. B. Oxidation of organoselenium compounds. A study of chemoselectivity of phenylacetone monooxygenase. *J. Mol. Catal. B Enzym.* 2011, 73 (1), 63-66.

135. Colonna, S.; Gaggero, N.; Carrea, G.; Ottolina, G.; Pasta, P.; Zambianchi, F. First asymmetric epoxidation catalysed by cyclohexanone monooxygenase. *Tetrahedron Lett.* 2002, 43 (10), 1797-1799.

136. Rial, D. V.; Bianchi, D. A.; Kapitanova, P.; Lengar, A.; van Beilen, J. B.; Mihovilovic, M. D. Stereoselective desymmetrizations by recombinant whole cells expressing the Baeyer–Villiger monooxygenase from *Xanthobacter* sp. ZL5: A New Biocatalyst accepting structurally demanding substrates. *Eur. J. Org. Chem.* 2008, 2008 (7), 1203-1213.

137. Xu, J.; Peng, Y.; Wang, Z.; Hu, Y.; Fan, J.; Zheng, H.; Lin, X.; Wu, Q. Exploiting cofactor versatility to convert a fad-dependent Baeyer–Villiger monooxygenase into a ketoreductase. *Angewandte Chemie International Edition* 2019.

138. Entsch, B.; Cole, L. J.; Ballou, D. P. Protein dynamics and electrostatics in the function of *p*-hydroxybenzoate hydroxylase. *Arch. Biochem. Biophys.* 2005, 433 (1), 297-311.

139. Ballou David, P.; Entsch, B. In *Handbook of Flavoproteins. Complex Flavoproteins, Dehydrogenases and Physical Methods*, De Gruyter: 2013; Vol. 2, pp 1-28.

140. Zambianchi, F.; Fraaije, M. W.; Carrea, G.; de Gonzalo, G.; Rodríguez, C.; Gotor, V.; Ottolina, G. Titration and assignment of residues that regulate the enantioselectivity of phenylacetone monooxygenase. *Adv. Synth. Catal.* 2007, 349 (8-9), 1327-1331.

141. Catucci, G.; Zgrablic, I.; Lanciani, F.; Valetti, F.; Minerdi, D.; Ballou, D. P.; Gilardi, G.; Sadeghi, S. J. Characterization of a new Baeyer–Villiger monooxygenase and conversion to a solely N-or S-oxidizing enzyme by a single R292 mutation. *Biochim. Biophys. Acta, Proteins Proteomics* 2016, 1864 (9), 1177-1187.

142. Fordwour, O. B.; Wolthers, K. R. Active site arginine controls the stereochemistry of hydride transfer in cyclohexanone monooxygenase. *Arch. Biochem. Biophys.* 2018, 659, 47-56.

143. Li, G.; Garcia-Borràs, M.; Fürst, M. J. L. J.; Ilie, A.; Fraaije, M. W.; Houk, K. N.; Reetz, M. T. Overriding traditional electronic effects in biocatalytic Baeyer–Villiger reactions by directed evolution. *J. Am. Chem. Soc.* 2018, 140 (33), 10464-10472.

144. Robinson, R.; Badieyan, S.; Sobrado, P. C4a-hydroperoxyflavin formation in N-hydroxylating flavin monooxygenases is mediated by the 2 -OH of the nicotinamide ribose of NADP⁺. *Biochemistry* 2013, 52 (51), 9089-9091.

145. Mascotti, M. L.; Lapadula, W. J.; Juri Ayub, M. The origin and evolution of Baeyer–Villiger monooxygenases (BVMOs): an ancestral family of flavin monooxygenases. *PLoS One* 2015, 10 (7), e0132689.

146. Butinar, L.; Mohorčič, M.; Deyris, V.; Duquesne, K.; Iacazio, G.; Claeys-Bruno, M.; Friedrich, J.; Alphand, V. Prevalence and specificity of Baeyer–Villiger monooxygenases in fungi. *Phytochemistry* 2015, 117, 144-153.

147. Szolkowy, C.; Eltis, L. D.; Bruce, N. C.; Grogan, G. Insights into Sequence–Activity Relationships amongst Baeyer–Villiger Monooxygenases as Revealed by the Intragenomic Complement of Enzymes from *Rhodococcus jostii* RHA1. *ChemBioChem* 2009, 10 (7), 1208-1217.

148. Wen, Y.; Hatabayashi, H.; Arai, H.; Kitamoto, H. K.; Yabe, K. Function of the cypX and moxY genes in aflatoxin biosynthesis in *Aspergillus parasiticus*. *Appl. Environ. Microbiol.* 2005, 71 (6), 3192-3198.

149. McGuire, S. M.; Townsend, C. A. Demonstration of Baeyer–Villiger oxidation and the course of cyclization in bisfuran ring formation during aflatoxin B1 biosynthesis. *Bioorg. Med. Chem. Lett.* 1993, 3 (4), 653-656.

150. Zeng, H.; Yin, G.; Wei, Q.; Li, D.; Wang, Y.; Hu, Y.; Hu, C.; Zou, Y. Unprecedented [5.5.5.6] Dioxafenstrane ring construction in fungal insecticidal sesquiterpene biosynthesis. *Angew. Chem. Int. Ed. Engl.* 2019, 58 (20), 6569-6573.

151. Hu, J.; Li, H.; Chooi, Y.-H. A Fungal dirigent protein controls the stereoselectivity of multicopper oxidase-catalyzed phenol coupling in viriditoxin biosynthesis. *J. Am. Chem. Soc.* 2019, 141 (20), 8068-8072.

152. Urquhart, A. S.; Hu, J.; Chooi, Y.-H.; Idnurm, A. The fungal gene cluster for biosynthesis of the antibacterial agent viriditoxin. *Fungal Biol. Biotechnol.* 2019, 6 (1), 9.

153. Jiang, J.; Tetzlaff, C. N.; Takamatsu, S.; Iwatsuki, M.; Komatsu, M.; Ikeda, H.; Cane, D. E. Genome mining in *Streptomyces avermitilis*: A biochemical Baeyer–Villiger reaction and discovery of a new branch of the pentalenolactone family tree. *Biochemistry* 2009, 48 (27), 6431-6440.

154. Iwaki, H.; Hasegawa, Y.; Wang, S.; Kayser, M. M.; Lau, P. C. Cloning and characterization of a gene cluster involved in cyclopentanol metabolism in *Comamonas* sp. strain NCIMB 9872 and biotransformations effected by *Escherichia coli*-expressed cyclopentanone 1,2-monooxygenase. *Appl. Environ. Microbiol.* 2002, 68 (11), 5671-5684.

155. Iwaki, H.; Wang, S.; Grosse, S.; Bergeron, H.; Nagahashi, A.; Lertvorachon, J.; Yang, J.; Konishi, Y.; Hasegawa, Y.; Lau, P. C. Pseudomonad cyclopentadecanone monooxygenase displaying an uncommon spectrum of Baeyer–Villiger oxidations of cyclic ketones. *Appl. Environ. Microbiol.* 2006, 72 (4), 2707-2720.

156. Kostichka, K.; Thomas, S. M.; Gibson, K. J.; Nagarajan, V.; Cheng, Q. Cloning and characterization of a gene cluster for cyclododecanone oxidation in *Rhodococcus ruber* SC1. *J. Bacteriol.* 2001, 183 (21), 6478-6486.

157. Huijbers, M. M. E.; Montersino, S.; Westphal, A. H.; Tischler, D.; van Berkel, W. J. H. Flavin dependent monooxygenases. *Arch. Biochem. Biophys.* 2014, 544, 2-17.

158. Dai, X.; Mashiguchi, K.; Chen, Q.; Kasahara, H.; Kamiya, Y.; Ojha, S.; DuBois, J.; Ballou, D.; Zhao, Y. The Biochemical mechanism of auxin biosynthesis by an *Arabidopsis* YUCCA flavin-containing monooxygenase. *J. Biol. Chem.* 2013, 288 (3), 1448-1457.

159. Jensen, C. N.; Cartwright, J.; Ward, J.; Hart, S.; Turkenburg, J. P.; Ali, S. T.; Allen, M. J.; Grogan, G. A Flavoprotein monooxygenase that catalyses a Baeyer–Villiger reaction and thioether oxidation using NADH as the nicotinamide cofactor. *ChemBioChem* 2012, 13 (6), 872-878.

160. Riebel, A.; de Gonzalo, G.; Fraaije, M. W. Expanding the biocatalytic toolbox of flavoprotein monooxygenases from *Rhodococcus jostii* RHA1. *J. Mol. Catal. B Enzym.* 2013, 88, 20-25.

161. Riebel, A.; Fink, M. J.; Mihovilovic, M. D.; Fraaije, M. W. Type II flavin-containing monooxygenases: A new class of biocatalysts that harbors Baeyer–Villiger monooxygenases with a relaxed coenzyme specificity. *ChemCatChem* 2014, 6 (4), 1112-1117.

162. Huang, L.; Romero, E.; Ressmann, A. K.; Rudroff, F.; Hollmann, F.; Fraaije, M. W.; Kara, S. Nicotinamide adenine dinucleotide-dependent redox-neutral convergent cascade for lactonizations with type II flavin-containing monooxygenase. *Adv. Synth. Catal.* 2017, 359 (12), 2142-2148.

163. Gibson, M.; Nur-e-alam, M.; Lipata, F.; Oliveira, M. A.; Rohr, J. Characterization of kinetics and products of the Baeyer–Villiger oxygenase MtmOIV, the key enzyme of the biosynthetic pathway toward the natural product Anticancer Drug Mithramycin from *Streptomyces argillaceus. J. Am. Chem. Soc.* 2005, 127 (50), 17594-17595.

164. Iwaki, H.; Grosse, S.; Bergeron, H.; Leisch, H.; Morley, K.; Hasegawa, Y.; Lau, P. C. K. Camphor pathway redux: functional recombinant expression of 2,5- and 3,6-diketocamphane monooxygenases of *Pseudomonas putida* ATCC 17453 with their cognate flavin reductase catalyzing Baeyer–Villiger reactions. *Appl. Environ. Microbiol.* 2013, 79 (10), 3282-3293.

165. Kim, T.-W.; Hwang, J.-Y.; Kim, Y.-S.; Joo, S.-H.; Chang, S. C.; Lee, J. S.; Takatsuto, S.; Kim, S.-K. *Arabidopsis* CYP85A2, a cytochrome P450, mediates the Baeyer–Villiger oxidation of castasterone to brassinolide in brassinosteroid biosynthesis. *Plant Cell* 2005, 17 (8), 2397-2412.

166. Nomura, T.; Kushiro, T.; Yokota, T.; Kamiya, Y.; Bishop, G. J.; Yamaguchi, S. The last reaction producing brassinolide is catalyzed by cytochrome P450s, CYP85A3 in tomato and CYP85A2 in *Arabidopsis. J. Biol. Chem.* 2005, 280 (18), 17873-17879.

167. Isupov, M. N.; Schroder, E.; Gibson, R. P.; Beecher, J.; Donadio, G.; Saneei, V.; Dcunha, S. A.; McGhie, E. J.; Sayer, C.; Davenport, C. F.; Lau, P. C.; Hasegawa, Y.; Iwaki, H.; Kadow, M.; Balke, K.; Bornscheuer, U. T.; Bourenkov, G.; Littlechild, J. A. The oxygenating constituent of 3,6-diketocamphane monooxygenase from the CAM plasmid of Pseudomonas putida: the first crystal structure of a type II Baeyer-Villiger monooxygenase. *Acta Cryst. D* 2015, 71 (11), 2344-2353.

168. Beam, M. P.; Bosserman, M. A.; Noinaj, N.; Wehenkel, M.; Rohr, J. r. Crystal Structure of Baeyer–Villiger monooxygenase MtmOIV, the key enzyme of the mithramycin biosynthetic pathway. *Biochemistry* 2009, 48 (21), 4476-4487.

169. Fraaije, M. W.; Wu, J.; Heuts, D. P.; van Hellemond, E. W.; Spelberg, J. H.; Janssen, D. B. Discovery of a thermostable Baeyer–Villiger monooxygenase by genome mining. *Appl. Environ. Microbiol.* 2005, 66 (4), 393-400.

170. Seo, M. J.; Zhu, D.; Endo, S.; Ikeda, H.; Cane, D. E. Genome mining in *Streptomyces.* Elucidation of the role of Baeyer–Villiger monooxygenases and non-heme iron-dependent dehydrogenase/oxygenases in the final steps of the biosynthesis of pentalenolactone and neopentalenolactone. *Biochemistry* 2011, 50 (10), 1739-1754.

171. Park, J.; Kim, D.; Kim, S.; Kim, J.; Bae, K.; Lee, C. The analysis and application of a recombinant monooxygenase library as a biocatalyst for the Baeyer-Villiger reaction. *J. Microbiol. Biotechnol.* 2007, 17 (7), 1083-1089.

172. Riebel, A.; Dudek, H. M.; de Gonzalo, G.; Stepniak, P.; Rychlewski, L.; Fraaije, M. W. Expanding the set of rhodococcal Baeyer-Villiger monooxygenases by high-throughput cloning, expression and substrate screening. *Appl. Microbiol. Biotechnol.* 2012, 95 (6), 1479-1489.

173. Fraaije, M. W.; Kamerbeek, N. M.; van Berkel, W. J.; Janssen, D. B. Identification of a Baeyer–Villiger monooxygenase sequence motif. *FEBS Lett.* 2002, 518 (1-3), 43-47.

174. Rebehmed, J.; Alphand, V.; de Berardinis, V.; de Brevern, A. G. Evolution study of the Baeyer–Villiger monooxygenases enzyme family: Functional importance of the highly conserved residues. *Biochimie* 2013, 95 (7), 1394-1402.

175. Tolmie, C.; Smit, M.; Opperman, D. Alternative splicing of the aflatoxin-associated baeyer–villiger monooxygenase from *Aspergillus flavus*: Characterisation of MoxY isoforms. *Toxins* 2018, 10 (12), 521.

176. Völker, A.; Kirschner, A.; Bornscheuer, U.; Altenbuchner, J. Functional expression, purification, and characterization of the recombinant Baeyer-Villiger monooxygenase MekA from *Pseudomonas veronii* MEK700. *Appl. Environ. Microbiol.* 2008, 77 (6), 1251-1260.

177. Milker, S.; Goncalves, L. C. P.; Fink, M. J.; Rudroff, F. *Escherichia coli* fails to efficiently maintain the activity of an important flavin monooxygenase in recombinant overexpression. *Frontiers in Microbiology* 2017, 8 (2201).

178. Goncalves, L. C. P.; Kracher, D.; Milker, S.; Fink, M. J.; Rudroff, F.; Ludwig, R.; Bommarius, A. S.; Mihovilovic, M. D. Mutagenesis-independent stabilization of class B flavin monooxygenases in operation. *Adv. Synth. Catal.* 2017, 359 (12), 2121-2131.

179. Milker, S.; Fink, M. J.; Oberleitner, N.; Ressmann, A. K.; Bornscheuer, U. T.; Mihovilovic, M. D.; Rudroff, F. Kinetic modeling of an enzymatic redox cascade in vivo reveals bottlenecks caused by cofactors. *ChemCatChem* 2017, 9 (17), 3420-3427.

180. Baldwin, C. V.; Woodley, J. M. On oxygen limitation in a whole cell biocatalytic Baeyer–Villiger oxidation process. *Biotechnol. Bioeng.* 2006, 95 (3), 362-369.

181. Baek, A. H.; Jeon, E. Y.; Lee, S. M.; Park, J. B. Expression levels of chaperones influence biotransformation activity of recombinant *Escherichia coli* expressing *Micrococcus luteus* alcohol dehydrogenase and *Pseudomonas putida* Baeyer–Villiger monooxygenase. *Biotechnol. Bioeng.* 2015, 112 (5), 889-895.

182. van Beek, H. L.; Wijma, H. J.; Fromont, L.; Janssen, D. B.; Fraaije, M. W. Stabilization of cyclohexanone monooxygenase by a computationally designed disulfide bond spanning only one residue. *FEBS Open Bio* 2014, 4, 168-174.

183. Schmidt, S.; Genz, M.; Balke, K.; Bornscheuer, U. T. The effect of disulfide bond introduction and related Cys/Ser mutations on the stability of a cyclohexanone monooxygenase. *J. Biotechnol.* 2015, 214, 199-211.

184. Fordwour, O. B.; Luka, G.; Hoorfar, M.; Wolthers, K. R. Kinetic characterization of acetone monooxygenase from Gordonia sp. strain TY-5. *AMB Express* 2018, 8 (1), 181.

185. Bisagni, S.; Summers, B.; Kara, S.; Hatti-Kaul, R.; Grogan, G.; Mamo, G.; Hollmann, F. Exploring the substrate specificity and enantioselectivity of a Baeyer–Villiger monooxygenase from *Dietzia* sp. D5: Oxidation of sulfides and aldehydes. *Top. Catal.* 2014, 57 (5), 366-375.

186. Leipold, F.; Wardenga, R.; Bornscheuer, U. T. Cloning, expression and characterization of a eukaryotic cycloalkanone monooxygenase from *Cylindrocarpon radicicola* ATCC 11011. *Appl. Environ. Microbiol.* 2012, 94 (3), 705-717.

187. Beneventi, E.; Niero, M.; Motterle, R.; Fraaije, M.; Bergantino, E. Discovery of Baeyer-Villiger monooxygenases from photosynthetic eukaryotes. *J. Mol. Catal. B Enzym.* 2013, 98, 145-154.

188. Kamerbeek, N. M.; Moonen, M. J.; Van Der Ven, J. G.; Van Berkel, W. J.; Fraaije, M. W.; Janssen, D. B. 4-Hydroxyacetophenone monooxygenase from *Pseudomonas fluorescens* ACB. A novel flavoprotein catalyzing Baeyer–Villiger oxidation of aromatic compounds. *Eur. J. Biochem.* 2001, 268 (9), 2547-2557.

189. Liu, Y. Y.; Li, C. X.; Xu, J. H.; Zheng, G. W. Efficient synthesis of methyl 3-acetoxypropionate by a newly identified Baeyer-Villiger monooxygenase. *Appl. Environ. Microbiol.* 2019, 85 (11), e00239-19.

190. Weiss, M.; Denger, K.; Huhn, T.; Schleheck, D. Two enzymes of a complete degradation pathway for linear alkylbenzenesulfonate (LAS) surfactants: 4-sulfoacetophenone Baeyer-Villiger monooxygenase and 4-sulfophenylacetateesterase in *Comamonas testosteroni* KF-1. *Appl. Environ. Microbiol.* 2012, 78 (23), 8254-8263.

191. Itagaki, E. Studies on steroid monooxygenase from *Cylindrocarpon radicicola* ATCC 11011. Oxygenative lactonization of androstenedione to testololactone. *J. Biochem.* 1986, 99 (3), 825-32.

192. Leipold, F.; Rudroff, F.; Mihovilovic, M. D.; Bornscheuer, U. T. The steroid monooxygenase from *Rhodococcus rhodochrous*; a versatile biocatalyst. *Tetrahedron: Asymmetry* 2013, 24 (24), 1620-1624.

193. Dudek, H. M.; Fink, M. J.; Shivange, A. V.; Dennig, A.; Mihovilovic, M. D.; Schwaneberg, U.; Fraaije, M. W. Extending the substrate scope of a Baeyer–Villiger monooxygenase by multiple-site mutagenesis. *Appl. Microbiol. Biotechnol.* 2014, 98 (9), 4009-4020.

194. Fiorentini, F.; Romero, E.; Fraaije, M. W.; Faber, K.; Hall, M.; Mattevi, A. Baeyer–Villiger monooxygenase FMO5 as entry point in drug metabolism. *ACS Chem. Biol.* 2017, 12 (9), 2379-2387.

195. Bosserman, M. A.; Downey, T.; Noinaj, N.; Buchanan, S. K.; Rohr, J. Molecular insight into substrate recognition and catalysis of Baeyer–Villiger monooxygenase MtmOIV, the key frame-modifying enzyme in the biosynthesis of anticancer agent mithramycin. *ACS Chem. Biol.* 2013, 8 (11), 2466-2477.

196. Kołek, T.; Szpineter, A.; Świzdor, A. Baeyer-Villiger oxidation of DHEA, pregnenolone, and androstenedione by Penicillium lilacinum AM111. *Steroids* 2008, 73 (14), 1441-1445.

197. Świzdor, A.; Panek, A.; Milecka-Tronina, N. Microbial Baeyer–Villiger oxidation of 5α-steroids using *Beauveria bassiana*. A stereochemical requirement for the 11α-hydroxylation and the lactonization pathway. *Steroids* 2014, 82, 44-52.

198. Javid, M.; Nickavar, B.; Vahidi, H.; Faramarzi, M. A. Baeyer-Villiger oxidation of progesterone by *Aspergillus sojae* PTCC 5196. *Steroids* 2018, 140, 52-57.

199. Mascotti, M. L.; Palazzolo, M. A.; Bisogno, F. R.; Kurina-Sanz, M. Biotransformation of dehydro-epi-androsterone by Aspergillus parasiticus: Metabolic evidences of BVMO activity. *Steroids* 2016, 109, 44-49.

200. Zhang, H.; Ren, J.; Wang, Y.; Sheng, C.; Wu, Q.; Diao, A.; Zhu, D. Effective multi-step functional biotransformations of steroids by a newly isolated *Fusarium oxysporum* SC1301. *Tetrahedron* 2013, 69 (1), 184-189.

201. de Gonzalo, G.; Mihovilovic, M. D.; Fraaije, M. W. Recent developments in the application of Baeyer–Villiger monooxygenases as biocatalysts. *ChemBioChem* 2010, 11 (16), 2208-2231.

202. Itagaki, E. Studies on steroid monooxygenase from Cylindrocarpon radicicola ATCC 11011. Purification and characterization. *J. Biochem.* 1986, 99 (3), 815-24.

203. Secundo, F.; Zambianchi, F.; Crippa, G.; Carrea, G.; Tedeschi, G. Comparative study of the properties of wild type and recombinant cyclohexanone monooxygenase, an enzyme of synthetic interest. *J. Mol. Catal. B Enzym.* 2005, 34 (1-6), 1-6.

204. Mascotti, M. L.; Juri Ayub, M.; Dudek, H.; Sanz, M. K.; Fraaije, M. W. Cloning, overexpression and biocatalytic exploration of a novel Baeyer-Villiger monooxygenase from *Aspergillus fumigatus* Af293. *AMB Express* 2013, 3 (1), 33.

205. Ferroni, F. M.; Smit, M. S.; Opperman, D. J. Functional divergence between closely related Baeyer-Villiger monooxygenases from Aspergillus flavus. *J. Mol. Catal. B Enzym.* 2014, 107, 47-54.

206. Mthethwa, K. S.; Kassier, K.; Engel, J.; Kara, S.; Smit, M. S.; Opperman, D. J. Fungal BVMOs as alternatives to cyclohexanone monooxygenase. *Enzyme Microb. Technol.* 2017, 106, 11-17.

207. Scherkus, C.; Schmidt, S.; Bornscheuer, U. T.; Groger, H.; Kara, S.; Liese, A. Kinetic insights into ε-caprolactone synthesis: Improvement of an enzymatic cascade reaction. *Biotechnol. Bioeng.* 2017, 114 (6), 1215-1221.

208. Staudt, S.; Bornscheuer, U. T.; Menyes, U.; Hummel, W.; Groger, H. Direct biocatalytic one-pot-transformation of cyclohexanol with molecular oxygen into varepsilon-caprolactone. *Enzyme Microb. Technol.* 2013, 53 (4), 288-92.

209. Engel, J.; Mthethwa, K. S.; Opperman, D. J.; Kara, S. Characterization of new Baeyer-Villiger monooxygenases for lactonizations in redox-neutral cascades. *Mol. Catal.* 2019, 468, 44-51.

210. Beneventi, E.; Ottolina, G.; Carrea, G.; Panzeri, W.; Fronza, G.; Lau, P. C. K. Enzymatic Baeyer–Villiger oxidation of steroids with cyclopentadecanone monooxygenase. *J. Mol. Catal. B Enzym.* 2009, 58 (1-4), 164-168.

211. Alphand, V.; Carrea, G.; Wohlgemuth, R.; Furstoss, R.; Woodley, J. M. Towards large-scale synthetic applications of Baeyer–Villiger monooxygenases. *Trends Biotechnol.* 2003, 21 (7), 318-323.

212. Bisagni, S.; Hatti-Kaul, R.; Mamo, G. Cloning, expression and characterization of a versatile Baeyer-Villiger monooxygenase from Dietzia sp. D5. *AMB Express* 2014, 4 (23), 1-10.

213. Bisagni, S.; Abolhalaj, M.; de Brevern, A. G.; Rebehmed, J.; Hatti-Kaul, R.; Mamo, G. Enhancing the activity of a *Dietzia* sp. D5 Baeyer-Villiger monooxygenase towards cyclohexanone by saturation mutagenesis. *ChemistrySelect* 2017, 2 (24), 7169-7177.

214. Ceccoli, R. D.; Bianchi, D. A.; Fink, M. J.; Mihovilovic, M. D.; Rial, D. V. Cloning and characterization of the type I Baeyer-Villiger monooxygenase from *Leptospira biflexa*. *AMB Express* 2017, 7 (87), 1-13.

215. Rodríguez, C.; de Gonzalo, G.; Fraaije, M. W.; Gotor, V. Enzymatic kinetic resolution of racemic ketones catalyzed by Baeyer–Villiger monooxygenases. *Tetrahedron: Asymmetry* 2007, 18 (11), 1338-1344.

216. Rudroff, F.; Fink, M. J.; Pydi, R.; Bornscheuer, U. T.; Mihovilovic, M. D. First chemo-enzymatic synthesis of the (R)-Taniguchi lactone and substrate profiles of CAMO and OTEMO, two new Baeyer-Villiger monooxygenases. *Monatsh. Chem.* 2017, 148 (1), 157-165.

217. Rehdorf, J.; Zimmer, C. L.; Bornscheuer, U. T. Cloning, expression, characterization, and biocatalytic investigation of the 4-hydroxyacetophenone monooxygenase from *Pseudomonas putida* JD1. *Appl. Environ. Microbiol.* 2009, 75.

218. Woo, J.-M.; Jeon, E.-Y.; Seo, E.-J.; Seo, J.-H.; Lee, D.-Y.; Yeon, Y. J.; Park, J.-B. Improving catalytic activity of the Baeyer–Villiger monooxygenase-based *Escherichia coli* biocatalysts for the overproduction of (Z)-11-(heptanoyloxy) undec-9-enoic acid from ricinoleic acid. *Sci. Rep.* 2018, 8 (1), 10280.

219. Seo, J.-H.; Kim, H.-H.; Jeon, E.-Y.; Song, Y.-H.; Shin, C.-S.; Park, J.-B. Engineering of Baeyer-Villiger monooxygenase-based *Escherichia coli* biocatalyst for large scale biotransformation of ricinoleic acid into (Z)-11-(heptanoyloxy) undec-9-enoic acid. *Sci. Rep.* 2016, 6 (28223), 1-9.

220. Della Pina, C.; Falletta, E.; Rossi, M. A green approach to chemical building blocks. The case of 3-hydroxypropanoic acid. *Green Chem.* 2011, 13 (7), 1624-1632.

221. Willetts, A. Characterised flavin-dependent two-component monooxygenases from the CAM plasmid of *Pseudomonas putida* ATCC 17453 (NCIMB 10007): Ketolactonases by another name. *Microorganisms* 2019, 7 (1), 1-34.

222. Willetts, A.; Kelly, D. Flavin-Dependent redox transfers by the two-component diketocamphane monooxygenases of camphor-grown *Pseudomonas putida* NCIMB 10007. *Microorganisms* 2016, 4 (4), 38.

223. Trudgill, P.; DuBus, R.; Gunsalus, I. Mixed function oxidation VI. Purification of a tightly coupled electron transport complex in camphor lactonization. *J. Biol. Chem.* 1966, 241 (18), 4288-4290.

224. Willetts, A.; Kelly, D. Reply to the Comment by Littlechild and Isupov. *Microorganisms* 2017, 5 (3), 55.

225. Villa, R.; Willetts, A. Oxidations by microbial NADH plus FMN-dependent luciferases from *Photobacterium phosphoreum* and *Vibrio fischeri*. *J. Mol. Catal. B Enzym.* 1997, 2 (4-5), 193-197.

226. Williams, D. R.; Trudgill, P. W.; Taylor, D. G. Metabolism of 1, 8-cineole by a *Rhodococcus* species: ring cleavage reactions. *Microbiology* 1989, 135 (7), 1957-1967.

227. Rodríguez, D.; Quirós, L. M.; Braña, A. F.; Salas, J. A. Purification and characterization of a monooxygenase involved in the biosynthetic pathway of the antitumor drug mithramycin. *J. Bacteriol.* 2003, 185 (13), 3962-3965.

228. Wang, C.; Gibson, M.; Rohr, J.; Oliveira, M. A. Crystallization and X-ray diffraction properties of Baeyer–Villiger monooxygenase MtmOIV from the mithramycin biosynthetic pathway in Streptomyces argillaceus. *Acta Crystallogr. F* 2005, 61 (11), 1023-1026.

229. Romero, E.; Gómez Castellanos, J. R.; Gadda, G.; Fraaije, M. W.; Mattevi, A. Same Substrate, many reactions: oxygen activation in flavoenzymes. *Chem. Rev.* 2018, 118 (4), 1742-1769.

230. Henderson, M. C.; Siddens, L. K.; Morré, J. T.; Krueger, S. K.; Williams, D. E. Metabolism of the anti-tuberculosis drug ethionamide by mouse and human FMO1, FMO2 and FMO3 and mouse and human lung microsomes. *Toxicol. Appl. Pharmacol.* 2008, 233 (3), 420-427.

231. Chen, G.-P.; Poulsen, L.; Ziegler, D. Oxidation of aldehydes catalyzed by pig liver flavin-containing monooxygenase. *Drug Metab. Disposition* 1995, 23 (12), 1390-1393.

232. Zhao, Y. Auxin biosynthesis and its role in plant development. *Annu. Rev. Plant Biol.* 2010, 61, 49-64.

233. Löwe, J.; Blifernez-Klassen, O.; Baier, T.; Wobbe, L.; Kruse, O.; Gröger, H. Type II flavoprotein monooxygenase PsFMO_A from the bacterium Pimelobacter sp. Bb-B catalyzes enantioselective Baeyer-Villiger oxidations with a relaxed cofactor specificity. *J. Biotechnol.* 2019, 294, 81-87.

234. Beier, A.; Bordewick, S.; Genz, M.; Schmidt, S.; van den Bergh, T.; Peters, C.; Joosten, H.-J.; Bornscheuer, U. T. Switch in cofactor specificity of a Baeyer–Villiger monooxygenase. *ChemBioChem* 2016, 17 (24), 2312-2315.

235. Dudek, H. M.; Torres Pazmiño, D. E.; Rodriguez, C.; de Gonzalo, G.; Gotor, V.; Fraaije, M. W. Investigating the coenzyme specificity of phenylacetone monooxygenase from *Thermobifida fusca*. *Appl. Environ. Microbiol.* 2010, 88 (5), 1135-1143.

236. Jensen, C. N.; Ali, S. T.; Allen, M. J.; Grogan, G. Mutations of an NAD(P)H-dependent flavoprotein monooxygenase that influence cofactor promiscuity and enantioselectivity. *FEBS Open Bio* 2013, 3, 473-8.

237. Jensen, C. N.; Ali, S. T.; Allen, M. J.; Grogan, G. Exploring nicotinamide cofactor promiscuity in NAD (P) H-dependent flavin containing monooxygenases (FMOs) using natural variation within the phosphate binding loop. Structure and activity of FMOs from *Cellvibrio* sp. BR and *Pseudomonas stutzeri* NF13. *J. Mol. Catal. B Enzym.* 2014, 109, 191-198.

238. Opperman, D. J.; Reetz, M. T. Towards practical Baeyer–Villiger-monooxygenases: Design of cyclohexanone monooxygenase mutants with enhanced oxidative stability. *ChemBioChem* 2010, 11 (18), 2589-2596.

239. Bornadel, A.; Hatti-Kaul, R.; Hollmann, F.; Kara, S. Enhancing the productivity of the bi-enzymatic convergent cascade for ε-caprolactone synthesis through design of experiments and a biphasic system. *Tetrahedron* 2016, 72 (46), 7222-7228.

240. Fürst, M. J. L. J.; Boonstra, M.; Bandstra, S.; Fraaije, M. W. Stabilization of cyclohexanone monooxygenase by computational and experimental library design. *Biotechnol. Bioeng.* 2019, 116 (9), 2167-2177.

241. Liang, Q.; Wu, S. Nonconserved hinge in Baeyer–Villiger monooxygenase affects catalytic activity and stereoselectivity]. *Sheng Wu Gong Cheng Xue Bao* 2015, 31 (3), 361-374.

242. Bong, Y. K.; Song, S.; Nazor, J.; Vogel, M.; Widegren, M.; Smith, D.; Collier, S. J.; Wilson, R.; Palanivel, S. M.; Narayanaswamy, K.; Mijts, B.; Clay, M. D.; Fong, R.; Colbeck, J.; Appaswami, A.; Muley, S.; Zhu, J.; Zhang, X.; Liang, J.; Entwistle, D. Baeyer–Villiger monooxygenase-mediated synthesis of esomeprazole as an alternative for kagan sulfoxidation. *J. Org. Chem.* 2018, 83 (14), 7453-7458.

243. Hummel, W.; Gröger, H. Strategies for regeneration of nicotinamide coenzymes emphasizing self-sufficient closed-loop recycling systems. *J. Biotechnol.* 2014, 191, 22-31.

244. Wang, X.; Saba, T.; Yiu, H. H.; Howe, R. F.; Anderson, J. A.; Shi, J. Cofactor NAD(P)H regeneration inspired by heterogeneous pathways. *Chem* 2017, 2 (5), 621-654.

245. Doig, S. D.; Avenell, P. J.; Bird, P. A.; Gallati, P.; Lander, K. S.; Lye, G. J.; Wohlgemuth, R.; Woodley, J. M. Reactor operation and scale-up of whole cell Baeyer-Villiger catalyzed lactone synthesis. *Biotechnol. Prog.* 2002, 18 (5), 1039-1046.

246. Sieben, M.; Steinhorn, G.; Müller, C.; Fuchs, S.; Chin, L. A.; Regestein, L.; Büchs, J. Testing plasmid stability of Escherichia coli using the continuously operated shaken BIOreactor system. *Biotechnol. Prog.* 2016, 32 (6), 1418-1425.

247. Kadisch, M.; Willrodt, C.; Hillen, M.; Bühler, B.; Schmid, A. Maximizing the stability of metabolic engineering-derived whole-cell biocatalysts. *Biotechnology journal* 2017, 12 (8), 1600170.

248. Doig, S. D.; Simpson, H.; Alphand, V.; Furstoss, R.; Woodley, J. M. Characterization of a recombinant *Escherichia coli* TOP10 [pQR239] whole-cell biocatalyst for stereoselective Baeyer–Villiger oxidations. *Enzyme Microb. Technol.* 2003, 32 (3-4), 347-355.

249. Woodley, J. M. Accelerating the implementation of biocatalysis in industry. *Appl. Microbiol. Biotechnol.* 2019, 103 (12), 4733-4739.

250. Faber, K., *Biotransformations in organic chemistry*. 6th ed.; Springer: Berlin, Germany, 2011.

251. Xue, R.; Woodley, J. M. Process technology for multi-enzymatic reaction systems. *Bioresour. Technol.* 2012, 115, 183-195.

252. Wachtmeister, J.; Rother, D. Recent advances in whole cell biocatalysis techniques bridging from investigative to industrial scale. *Curr. Opin. Biotechnol.* 2016, 42, 169-177.

253. Dascier, D.; Kambourakis, S.; Hua, L.; Rozzell, J. D.; Stewart, J. D. Influence of cofactor regeneration strategies on preparative-scale, asymmetric carbonyl reductions by engineered *Escherichia coli*. *Org. Process Res. Dev.* 2014, 18 (6), 793-800.

254. Hollmann, F.; Arends, I. W.; Buehler, K. Biocatalytic redox reactions for organic synthesis: nonconventional regeneration methods. *ChemCatChem* 2010, 2 (7), 762-782.

255. Muschiol, J.; Peters, C.; Oberleitner, N.; Mihovilovic, M. D.; Bornscheuer, U. T.; Rudroff, F. Cascade catalysis–strategies and challenges en route to preparative synthetic biology. *Chem. Comm.* 2015, 51 (27), 5798-5811.

256. Sattler, J. H.; Fuchs, M.; Mutti, F. G.; Grischek, B.; Engel, P.; Pfeffer, J.; Woodley, J. M.; Kroutil, W. Introducing an in situ capping strategy in systems biocatalysis to access 6-aminohexanoic acid. *Angew. Chem. Int. Ed. Engl.* 2014, 53 (51), 14153-14157.

257. Milker, S.; Fink, M. J.; Rudroff, F.; Mihovilovic, M. D. Non-hazardous biocatalytic oxidation in Nylon-9 monomer synthesis on a 40 g scale with efficient downstream processing. *Biotechnol. Bioeng.* 2017, 114 (8), 1670-1678.

258. Sudheer, P. D.; Yun, J.; Chauhan, S.; Kang, T. J.; Choi, K.-Y. Screening, expression, and characterization of Baeyer-Villiger monooxygenases for the production of 9-(nonanoyloxy) nonanoic acid from oleic acid. *Biotechnology and Bioprocess Engineering* 2017, 22 (6), 717-724.

259. Jeon, E. Y.; Baek, A. H.; Bornscheuer, U. T.; Park, J. B. Enzyme fusion for whole-cell biotransformation of long-chain sec-alcohols into esters. *Appl. Environ. Microbiol.* 2015, 99 (15), 6267-6275.

260. Sudheer, P. D.; Seo, D.; Kim, E.-J.; Chauhan, S.; Chunawala, J.; Choi, K.-Y. Production of (Z)-11-(heptanoyloxy) undec-9-enoic acid from ricinoleic acid by utilizing crude glycerol as sole carbon source in engineered Escherichia coli expressing BVMO-ADH-FadL. *Enzyme Microb. Technol.* 2018, 119, 45-51.

261. Delgove, M. A.; Valencia, D.; Solé, J.; Bernaerts, K. V.; De Wildeman, S. M.; Guillén, M.; Álvaro, G. High performing immobilized Baeyer-Villiger monooxygenase and glucose dehydrogenase for the synthesis of ε-caprolactone derivative. *Appl. Catal., A.* 2019, 572, 134-141.

262. Solé, J.; Brummund, J.; Caminal, G.; Alvaro, G.; Schürmann, M.; Guillén, M. Enzymatic synthesis of trimethyl-ε-caprolactone: process intensification and demonstration at 100 liter scale. *Org. Process Res. Dev.* 2019.

263. Schulz, F.; Leca, F.; Hollmann, F.; Reetz, M. T. Towards practical biocatalytic Baeyer-Villiger reactions: applying a thermostable enzyme in the gram-scale synthesis of optically-active lactones in a two-liquid-phase system. *Beilstein J. Org. Chem.* 2005, 1 (10), 1-9.

264. Valencia, D.; Guillén, M.; Fürst, M. J. L. J.; López-Santín, J.; Álvaro, G. An immobilized and highly stabilized self-sufficient monooxygenase as biocatalyst for oxidative biotransformations. *J. Chem. Technol. Biotechnol.* 2017, 93 (4), 985-993.

265. Krajčovič, T.; Bučko, M.; Vikartovská, A.; Lacík, I.; Uhelská, L.; Chorvát, D.; Neděla, V.; Tihlaříková, E.; Gericke, M.; Heinze, T. Polyelectrolyte complex beads by novel two-step process for improved performance of viable whole-cell Baeyer-Villiger monoxygenase by immobilization. *Catalysts* 2017, 7 (11), 353.

266. Bucko, M.; Gemeiner, P.; Schenkmayerova, A.; Krajcovic, T.; Rudroff, F.; Mihovilovic, M. D. Baeyer-Villiger oxidations: biotechnological approach. *Appl. Environ. Microbiol.* 2016, 100 (15), 6585-6599.

267. Catucci, G.; Gao, C.; Sadeghi, S. J.; Gilardi, G. Chemical applications of Class B flavoprotein monooxygenases. *Rendiconti Lincei* 2017, 28 (1), 195-206.

268. Torres Pazmiño, D. E.; Dudek, H. M.; Fraaije, M. W. Baeyer-Villiger monooxygenases: recent advances and future challenges. *Curr. Opin. Chem. Biol.* 2010, 14 (2), 138-144.

269. Mallin, H.; Wulf, H.; Bornscheuer, U. T. A self-sufficient Baeyer-Villiger biocatalysis system for the synthesis of varepsilon-caprolactone from cyclohexanol. *Enzyme Microb. Technol.* 2013, 53 (4), 283-287.

270. Schmidt, S.; Scherkus, C.; Muschiol, J.; Menyes, U.; Winkler, T.; Hummel, W.; Groger, H.; Liese, A.; Herz, H. G.; Bornscheuer, U. T. An Enzyme cascade synthesis of epsilon-caprolactone and its oligomers. *Angew. Chem. Int. Ed. Engl.* 2015, 54 (9), 2784-2787.

271. Srinivasamurthy, V. S.; Böttcher, D.; Bornscheuer, U. T. A multi-enzyme cascade reaction for the production of 6-hydroxyhexanoic acid. *Z. Naturforsch. C Bio. Sci.* 2019, 74 (3-4), 71-76.

272. Pennec, A.; Hollmann, F.; Smit, M. S.; Opperman, D. J. One-pot conversion of cycloalkanes to lactones. *ChemCatChem* 2015, 7 (2), 236-239.

273. Oberleitner, N.; Peters, C.; Rudroff, F.; Bornscheuer, U. T.; Mihovilovic, M. D. In vitro characterization of an enzymatic redox cascade composed of an alcohol dehydrogenase, an enoate reductases and a Baeyer–Villiger monooxygenase. *J. Biotechnol.* 2014, 192, 393-399.

274. Peters, C.; Rudroff, F.; Mihovilovic, M. D.; Bornscheuer, U. T. Fusion proteins of an enoate reductase and a Baeyer-Villiger monooxygenase facilitate the synthesis of chiral lactones. *Biol. Chem.* 2017, 398 (1), 31-37.

275. Wu, J. T.; Wu, L. H.; Knight, J. A. Stability of NADPH: Effect of various factors on the kinetics of degradation. *Clin. Chem.* 1986, 32 (2), 314-319.

276. Aalbers, F. S.; Fraaije, M. W. Enzyme fusions in biocatalysis: coupling reactions by pairing enzymes. *ChemBioChem* 2018, 20 (1), 20-28.

277. Mourelle-Insua, Á.; Aalbers, F. S.; Lavandera, I.; Gotor-Fernández, V.; Fraaije, M. W. What to sacrifice? Fusions of cofactor regenerating enzymes with Baeyer-Villiger monooxygenases and alcohol dehydrogenases for self-sufficient redox biocatalysis. *Tetrahedron* 2019, 75 (13), 1832-1839.

278. Huang, L.; Aalbers, F. S.; Tang, W.; Röllig, R.; Fraaije, M. W.; Kara, S. Convergent cascade catalyzed by monooxygenase–alcohol dehydrogenase fusion applied in organic media. *ChemBioChem* 2019, 20 (13), 1653-1658.

279. Klibanov, A. M. Improving enzymes by using them in organic solvents. *Nature* 2001, 409 (6817), 241.

280. Castellana, M.; Wilson, M. Z.; Xu, Y.; Joshi, P.; Cristea, I. M.; Rabinowitz, J. D.; Gitai, Z.; Wingreen, N. S. Enzyme clustering accelerates processing of intermediates through metabolic channeling. *Nat. Biotechnol.* 2014, 32 (10), 1011-1018.

281. Kuzmak, A.; Carmali, S.; von Lieres, E.; Russell, A. J.; Kondrat, S. Can enzyme proximity accelerate cascade reactions? *Sci. Rep.* 2019, 9 (1), 455.

282. Wheeldon, I.; Minteer, S. D.; Banta, S.; Barton, S. C.; Atanassov, P.; Sigman, M. Substrate channelling as an approach to cascade reactions. *Nat. Chem.* 2016, 8 (4), 299-309.

283. Aalbers, F. S.; Fraaije, M. W. Coupled reactions by coupled enzymes: alcohol to lactone cascade with alcohol dehydrogenase–cyclohexanone monooxygenase fusions. *Appl. Microbiol. Biotechnol.* 2017, 101 (20), 7557-7565.

284. Unsworth, L. D.; van der Oost, J.; Koutsopoulos, S. Hyperthermophilic enzymes– stability, activity and implementation strategies for high temperature applications. *FEBS J.* 2007, 274 (16), 4044-4056.

# Optimizing the linker length for fusing an alcohol dehydrogenase with a cyclohexanone monooxygenase

**Alejandro Gran-Scheuch**[§], Friso S. Aalbers[§], Yannick Woudstra, Loreto Parra and Marco W. Fraaije

[§]These authors contributed equally

# ABSTRACT

The use of enzymes in organic synthesis is highly appealing due their remarkably high chemo-, regio- and enantioselectivity. Nevertheless, for biosynthetic routes to be industrially useful, the enzymes must fulfill several requirements. Particularly, in case of cofactor-dependent enzymes self-sufficient systems are highly valuable. This can be achieved by fusing enzymes with complementary cofactor dependency. Such bifunctional enzymes are also relatively easy to handle, may enhance stability, and promote product intermediate channeling. However, usually the characteristics of the linker, fusing the target enzymes, are not thoroughly evaluated. A poor linker design can lead to detrimental effects on expression levels, enzyme stability and/or enzyme performance. In this chapter, the effect of the length of a glycine-rich linker was explored for the case study of ε-caprolactone synthesis through an alcohol dehydrogenase-cyclohexanone monooxygenase fusion system. The procedure includes cloning of linker variants, expression analysis, determination of thermostability and effect on activity and conversion levels of fifteen variants of different linker sizes. The protocols can also be used for the creation of other protein-protein fusions.

**Keywords:** Baeyer-Villiger monooxygenase, alcohol dehydrogenase, biocatalytic cascade, cofactor regeneration, cyclohexanone

# INTRODUCTION

Biocatalysis refers to the use of enzymes to catalyze the synthesis of chemicals. For this, the field has exploited the accumulated knowledge of catalytic mechanisms, kinetic parameters and structure-function relationships of enzymes. This fundamental knowledge enables the development of improved biocatalysts for the chemo-, regio- and enantioselective synthesis of (fine) chemicals[1,2]. The use of biocatalysts in the synthesis of industrially relevant chemicals, such as lactones, is well described[1,3,4]. These compounds are highly attractive as platform chemicals for the pharmaceutical and chemical industries, because they are often used as precursors for valuable molecules, such as floral scents, or for polymer synthesis[3–6]. For example, ε-caprolactone, is a valued 6-carbon lactone that can be used as a precursor for the synthesis of caprolactam[7]. This product is further polymerized to form nylon-6, a popular polyamide polymer that is used in the manufacturing industries because of its high strength and elasticity. In addition to the production of polyamides, the most significant chemical feature of ε-caprolactone is its capacity to polymerize through a straightforward ring-opening reaction[8,9]. The obtained product is polycaprolactone, a biodegradable thermo polyester used as additive in resins, coating, colorant materials and polyurethanes[10–12]. Nowadays, ε-caprolactone has a high global demand and is produced on a scale of over 10,000 tons per year[13]. However, the classical synthesis of the lactone is not eco-friendly, as this compound is generally obtained through a Baeyer-Villiger oxidation using peracetic acid and cyclohexanone in an anhydrous solvent such as acetone[7,12]. Otherwise, it can be obtained by letting an excess of cyclohexanone react with air at 25-50 °C in the presence of acetaldehyde and metal-oxidant catalysts. As a byproduct, acetic acid is obtained, which must be removed by distillation[7]. A biocatalytic process would offer an attractive greener alternative for the synthesis of ε-caprolactone.

Broadly, enzymes have evolved for millions of years to catalyze specific reactions with a high efficiency. The use of enzymes may reduce energy costs by performing catalysis at milder conditions than classical chemical methods, such as lower temperature and pressure[14–16]. Enzymes can be employed as i) isolated enzymes, ii) immobilized enzymes, iii) cell free extracts or iv) within whole cells[17,18]. The latter is the most widely used method for large scale reactions, since (1) it is the cheapest way to formulate enzymes, (2) it provides greater enzyme stability, and (3) it satisfies biochemical requirements such as the presence of cofactors and cosubstrates. However, the use of whole cells also implies more complex mass balance calculations and transfer concerns, such as inadequate supply of oxygen or substrate/product transfer in/out the cells[19–23]. Evidently, each of these approaches has its own pros and cons, and the application will depend on the specific reaction to be achieved.

Baeyer-Villiger monooxygenases (BVMO, E.C. 1.14.13) have been described to produce ε-caprolactone using cyclohexanone as a substrate[10,24–26]. In particular, type I BVMOs from

the class B flavin-dependent monooxygenases contain flavin adenine dinucleotide (FAD) as a prosthetic group, which is required for catalysis. In the catalytic cycle, the FAD is reduced by NADPH[27]. Subsequently, the reduced enzyme reacts with dioxygen to form a reactive peroxyflavin enzyme intermediate which reacts with a ketone to form an ester or lactone[28]. Release of the product and NADP[+] completes the catalytic cycle. The most studied BVMO for the synthesis of ε-caprolactone is a cyclohexanone monooxygenase (AcCHMO) isolated from the soil bacterium *Acinetobacter calcoaceticus* NCIMB 9871[29]. Although this enzyme shows a remarkably broad substrate scope, it exhibits very poor stability. Specifically, its apparent melting temperature ($T_M^{app}$) is around 36 °C and it has a low tolerance towards cosolvents[30]. Recently, a more robust CHMO homolog was described and crystallized[31]. This variant, TmCHMO, was discovered in the actinobacterium *Thermocrispum municipale* DSM 44069. Even though this BVMO displays high sequence identity of 57 % with AcCHMO, TmCHMO showed greater robustness in presence of cosolvents and a significantly higher $T_M^{app}$. A clear disadvantage for CHMO and BVMOs as biocatalysts is their NADPH dependence. This cosubstrate is relatively expensive and therefore problematic to be used for biosynthetic routes. To remedy this, CHMO has been explored as fusion enzyme with different NADPH regenerating enzymes, such as phosphite dehydrogenase, formate dehydrogenase, glucose dehydrogenase or alcohol dehydrogenase (ADH)[32,33]. Interestingly, some alcohol dehydrogenases (ADH, E.C.1.1.1.X) are able to oxidize cyclohexanol to cyclohexanone while reducing NAD(P)[+]. The combination of such an ADH with a BVMO can catalyze a cascade reaction, in which the ADH regenerates NADPH while producing the substrate for the BVMO. For the particular reaction of this study, ADH catalyzes the oxidation of cyclohexanol (forming NADPH) which is converted into ε-caprolactone by CHMO (consuming NADPH)[10,24,34,35].

For the biocatalytic cascade reaction above, it is attractive to create a fused bifunctional ADH-BVMO fusion. For such fusion, it is important to introduce a proper linker. The linker is expected to allow enough flexibility and proper folding of the fused enzymes. The close proximity of two active-sites may also increase the conversion rates by enhancing the channeling of the reaction intermediates[36]. Although this substrate channeling effect is questioned in some studies[37–39], and likely only applies when the concentration of the intermediate remains low, there are a number of studies that found higher conversions for a cascade reaction with fused enzymes compared to the equimolar combination of separate enzymes[40]. The ADH-BVMO cascade system has been previously studied for the optimization of the biocatalytic conversion of cyclohexanol into ε-caprolactone (**Figure 1**)[34,41]. However, the reaction design is not trivial and some critical parameters are worth mentioning. For example, a poor substrate/cosubstrate set up can adversely affect the conversion rates. Particularly, AcCHMO suffers from inhibition by cyclohexanol, cyclohexanone and caprolactone[42]. For TmCHMO, an activity reduction of 50 % was observed at 66 mM ε-caprolactone, and 75 % at 2 mM cyclohexanol[43]. Therefore, a good cascade strategy must be found that balances optimal kinetic conditions for both enzymes.
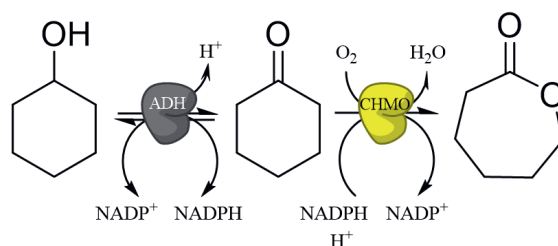
**Figure 1. TbADH-TmCHMO cascade reaction.** A schematic representation of the cascade reaction is shown. The cascade starts with the reversible oxidation of cyclohexanol to cyclohexanone and NADP$^+$ reduction by TbADH. Subsequently, the TmCHMO partner accepts the cyclohexanone and performs a Baeyer-Villiger oxidation. For this reaction, the hydride donor is consumed and lactone and water are produced. Then, the oxidized NADP$^+$ can be re-occupied for the first reaction.

In a previously published report, three different ADHs were explored as TmCHMO fusion partners: i) TbADH from *Thermoanaerobacter brockii*, ii) PfADH from *Pyrococcus furiosus* and iii) MiADH from *Mesotoga infera*[43]. To avoid product inhibition, a lipase was employed as additional biocatalyst, to decrease product concentration by catalyzing the ring opening of ε-caprolactone. The best performing fusion (TbADH-TmCHMO) achieved full conversion of 200 mM cyclohexanol with a turnover number (TON) of 13,000, while using separate enzymes only 41 % conversion was obtained (TON=5,600). Even though the TbADH-TmCHMO fusion gave promising results for the cascade reaction, the study did not try to optimize the linker. In that sense, both length and physicochemical characteristics of the linker can be tuned. Key aspects to consider are flexibility/rigidity and hydrophilicity/hydrophobicity of the linker sequence. Jeon et al. investigated the use of two different linkers in a similar two-step whole-cell biocatalytic process —conversion of long-chain *sec*-alcohols into esters—[44]. Fusions were evaluated consisting of an ADH from *Micrococcus luteus* NCTC2665 and BVMOs of *Pseudomonas putida* KT2440 or *Rhodococcus jostii* RHA1. The use of a glycine-rich linker had a better performance on the conversions than a rigid α-helix linker, and the fusion also outperformed the same combination of separate enzymes. Moreover, various studies have described that linkers with non-polar residues, such as the glycine-rich linker, are beneficial by enhancing the flexibility between both partners[45,46]. This would provide freedom for correct folding and conformational changes. In a critical study on flexible linkers, that also included computational modeling work, linker length and composition turned out to be highly important[47]. Shorter linkers and sequences containing more glycines increased FRET efficiencies between the ECFP and EYFP domains of the fusion. Other studies have described that the size of the linker can critically perturb the biocatalytic properties of fusions, such as enzyme activity, stability or coupling efficiency[48–50]. Therefore, in addition to the physicochemical composition of the linker, the length itself may have a key role on the biocatalytic features of the fusion. Thus, it is attractive to evaluate whether this is also the case for other fusions, such as the ADH-CHMO fusion, and to study the effects of linker length on its performance of its cascade reaction.

In this chapter, we describe an experimental procedure for exploring the effect of the length of a glycine-rich linker on the biocatalytic properties of a ADH-BVMO fusion. Specifically, the TbADH-TmCHMO fusion for the production of ε-caprolactone from cyclohexanol was studied. The glycine-rich linker SSGGSGGSGGSAGTA was selected based on previous work. Fusion variants were designed with different linker size using the P-LinK methodology[51]. The fusion enzymes were prepared with a linker size of 1 to 15 amino acids. The complete procedure is as follows: i) verifying the expression levels of the fusion, ii) spectrophotometrically analyze the oxidation state of TmCHMO, iii) evaluate the effect of the linker length on the thermostability, iv) activity and v) conversion levels.

## GENERAL METHOD AND STATISTICAL ANALYSIS

Common security measures for a biosafety level-1 laboratory must be followed. Personal protective elements as glasses, lab coat and chemically resistant gloves are required for the researcher when performing the experiments. For the statistical analysis, the determination of thermostability, activity and conversions level were performed in triplicates or duplicates. The results were analyzed using GraphPad Prism v6.05 for Windows (GraphPad Software, La Jolla, CA, United States). Statistical difference was evaluated using ordinary one-way ANOVA, using significance $p < 0.05$. Multiple comparisons against the control was performed using Dunnett's multiple comparison test. Nevertheless, the user can use other suitable software according to his preferences.

## MOLECULAR DESIGN OF LINKER VARIANTS

The most straightforward way to obtain a fusion enzyme is by ordering a synthetic DNA construct for expression of the respective protein. On the other hand, there are various ways to create a construct to express a fusion enzyme using the two respective genes. A couple of well-described and convenient cloning approaches are: 1) traditional restriction/ligation by adding a unique restriction site at the reverse primer of the first gene (for the first enzyme) and the forward primer of the second gene (for the second enzyme of the fusion), such that they can be ligated. It is worth mentioning that the stop codon of the first gene must be removed. 2) Golden Gate cloning, similar to restriction/ligation, except that it employs a restriction site (BsaI) that is not recognized after ligation, which greatly improves efficiency for ligation, and reduces steps compared to restriction/ligation. A disadvantage is that a Golden gate vector is required. Lastly, using 3) Gibson assembly/In-Fusion® cloning by designing primers with 15-30 base pairs (bp) overlapping regions, the amplified fragments can be assembled through an isothermal reaction with 3 enzymes: DNA polymerase, 5' exonuclease, and DNA ligase. In each of these methods there is some

room to build a linker region, by adding codons that code for a linker to the reverse primer of the first gene and to the forward primer of the second gene. Alternatively, with each method an existing linker can also be amplified as fragment, and ligated together with the two genes.

Once a construct coding for a fusion enzyme is obtained, different linker sequences can be designed and introduced. In this section, we describe the use of the P-LinK method that has been described to change the linker length between two domains of a cytochrome P450 monooxygenase[51]. This method is analogous to QuikChange with partially overlapping primers, with as difference that the primers are partly phosphorothioated, which enables efficient chemical ligation of the vectors after amplification (referred to as PLIC reaction from Phosphorothioated Ligation Independent Cloning). This is an advantage in terms of efficiency and number of steps needed to complete the construct compared to, for example, Gibson cloning. There are other methods to introduce different linkers, of which one recent example is detailed in Chapter 10 "Combinatorial Linker Engineering with iFLinkC" of this edition of Methods in Enzymology, as well as in a recent publication[52].

## Equipment

- PCR Thermocycler machine (T100 thermal Cycler, Bio-Rad).
- Shaker (Innova 44, New Brunswick Scientific).
- Thin-walled PCR tubes (DNAse free and/or sterile).
- Electrophoresis chamber (Bio-Rad).
- Thermomixer or water bath.
- Speedvac equipment.
- UV transilluminator (M20, Appligene).

## Reagents

- PCR oligonucleotides containing 12 phosphorothioates (PTO) (see note 1) at the 5'-end (see Procedure: Step 1 and Step 2).
- Pfu polymerase (as master mix, or with separate components: dNTPs mix, polymerase buffer).
- PCR cleanup kit (Qiagen).
- Plasmid MiniPrep kit (Qiagen).
- Sterile dH2O.
- Agarose gel (1 % w/v) (from agarose powder).
- DpnI enzyme (NEB).
- Sterile 80 % glycerol.
- Iodine solution (100 mM I2 in ethanol).
- Sterile LB medium (for 1 L, 10 g tryptone, 10 g NaCl and 5 g yeast extract). For LB agar, add agar-agar at 2 % w/v final concentration (20 g for 1 L).

- Escherichia coli NEB 10-β chemo competent cells, genotype: Δ(ara-leu) 7697 araD139 fhuA ΔlacX74 galK16 galE15 e14-φ80dlacZΔM15 recA1 relA1 endA1 nupG rpsL (StrR) rph spoT1 Δ(mrr-hsdRMS-mcrBC).

## Procedure

1. Design oligonucleotides as follows: the protocol has been optimized for having (at least) 12 phosphorothioates (PTOs) at the 5' end of the reverse primer and forward primers. For every 5 linker lengths (1-5 amino acids), the same reverse (RV) primer is used and 5 different forward (FW) primers. See the example in **table S1**.

2. Solubilize PTO oligonucleotide stocks to 100 μM in 10 mM TrisHCl pH 7.0 or autoclaved MilliQ water. Then, dilute to 20 μM for working stocks. Store the 100 μM stocks at -20 °C to preserve.

3. Set up PCR reactions with either PCR master mix, following the manufacturer's instruction, or using separate components, as follows in **table S2**. Include a reaction without any primers added, to serve as control for the DpnI digestion.

4. Transfer the tubes to a thermocycler and start the following PCR program (**Table S3**). The lid must be at 105 °C to prevent condensation.

5. Analyze the PCR products by loading them in a 1 % agarose gel (20-50 mL) with 1.5 μL of RotiSafe as gel stainer. Mix 3 μL of PCR sample with 0.5 μL of loading dye (6x), and add it to the well, then examine the gel under UV light in a UV transilluminator (**Figure 2a**) (see note 2).

6. Add 10 U of DpnI and CutSmart buffer (1x final concentration) to each reaction that gave product, and incubate them at 37 °C for at least 1 hour or overnight (16 hours).

7. Purify the PCR products with a PCR cleanup kit, and in the final step elute with 22 μL dH$_2$O, to ensure a high concentration of DNA.

8. Determine DNA concentration with a NanoDrop spectrophotometer. A concentration of approximately 0.02 – 0.04 pmol μL$^{-1}$ (=μM) is desired. To translate that value to ng μL$^{-1}$, the following online tool can be used: https://nebiocalculator.neb.com/#!/dsdnaamt (see note 3).

9. To increase the DNA concentration, one can apply the samples to a speedvac, to reach the 0.02 – 0.04 pmol μL$^{-1}$ concentration range. Optionally, an additional PCR reaction can be performed using the corresponding purified PCR product as template.

10. Place the purified PCR products and thin-wall PCR tubes on ice. Transfer 4 μL of each PCR product (at 0.02 – 0.04 pmol μL$^{-1}$) to a thin-wall PCR tube, as well as one for the control, and keep on ice.

11. To anneal the PCR product and form the circularized vector, a PLIC reaction is carried out. This reaction consists of two components: (1) the PCR product, and (2) iodine cleavage solution. This iodine solution can be prepared as follows **table S4.** Add 2 μL of iodine cleavage mixture to each thin-wall PCR tube that contains 4 μL of PCR reaction, for a total volume of 6 μL. The solution should be yellowish. Keep the solutions on ice (see note 4).

12. Transfer the tubes into PCR thermocycler with a preheated lid (80 °C), and start the program with an incubation of 10 minutes at 70 °C (PTO cleaving step). Then, for the annealing/hybridization step, incubate the tubes at 25 °C for 5 minutes After this step is done, the tubes can be transferred into ice, or stored at 4 °C. The solutions are ready to be used to transform chemically competent *E. coli* cells.

13. Set a thermomixer or water bath to 42 °C. Pre-warm 1 mL SOC or LB media to 37 °C in the thermo block, then transform the samples (PLIC reaction with each PCR product) and include the PLIC reaction control (without iodine) and a DpnI control.

14. Thaw 18 tubes of 50-100 µL chemically competent *E. coli* NEB 10-β cells on ice, and label them accordingly (1-18). Add the 2.5 µL to the respectively labeled tube and mix gently by flicking the tube, and directly place back on ice.

15. Incubate on ice for 20-30 minutes.

16. Perform a heat shock by incubating the tubes for 30 seconds at 42 °C in the thermomixer or water bath. Place them back into the ice after the 30 seconds.

17. Add 900 µL pre-warmed SOC or LB media to each tube, and incubate them at 37 °C for 60 minutes while shaking. This can be done in a large shaking incubator with a tube rack, or in a thermomixer at 700 r.p.m. This is necessary to ensure proper aeration.

18. Centrifuge the tubes at 5,000 g for 1 minute, remove 700-800 µL of the supernatant, and resuspend the remaining 100-200 µL with a pipette using sterile tips, and transfer to an LB agar plate with the appropriate antibiotic (50 µg mL$^{-1}$ ampicillin). Spread the cell suspension with sterile glass beads, an inoculation loop, a bend glass pipette, or a glass/metal spreader. If one of these tools was heated with a flame to sterilize, ensure that it is cooled down prior to touching the cell suspension.

19. Incubate the agar plates at 37 °C overnight.

20. If everything went well, the control plates should be empty or nearly empty, while the plates from the PLIC reaction should have >100 colonies. Typically, the more colonies, the higher the chance the correct construct will be obtained.

21. Pick a colony from each plate, except the controls, to inoculate 5 mL LB with 50 µg mL$^{-1}$ ampicillin, to grow overnight at 30 °C (>18 hours) or 37 °C (12-18 hours).

22. Transfer ~2 mL from the overnight culture to a 2 mL eppendorf tube, and spin down the cells at 5,000 g for 1 minute. Use a MiniPrep kit to purify the plasmid from this pellet, according to the kit's instructions. Store the remaining 3 mL of the culture at 4 °C.

23. Send the plasmid for sequencing, to ensure the right construct is obtained. After sequence confirmation, use the corresponding stored culture to make a -70/-80 °C glycerol stock by mixing 400 µL culture with 400 µL of sterile 80 % glycerol, and store at -70/-80 °C (see note 5). Alternatively, the corresponding plasmid can be used to transform chemically competent cells, and these can be used to make a -70/-80 °C stock (see note 6).
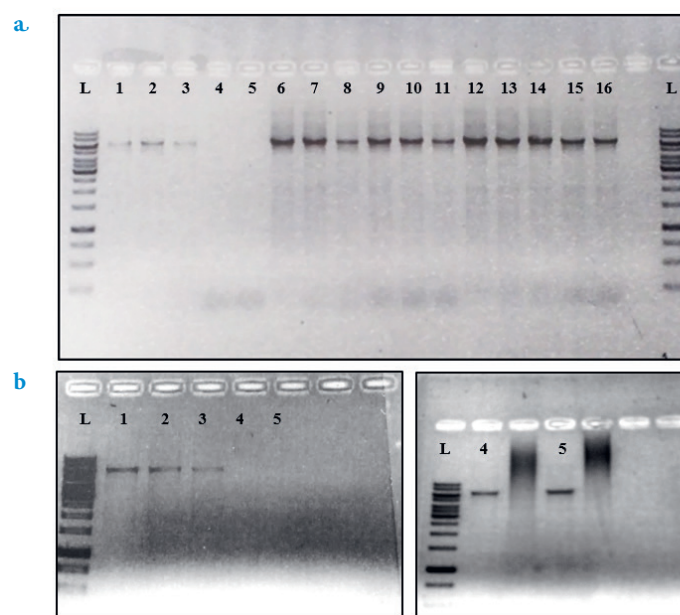
**Figure 2. Agarose gel analysis of the PCR products.** The numbers indicate the linker length of the PCR product. Lanes marked with L are loaded with the ladder (Page Ruler).

## Notes

- A notable disadvantage of this method is the need for phosphorothioated primers, which are quite costly or require in-house primer synthesis. There are other methods to introduce different linkers, of which one recent example is detailed in this edition of Methods in Enzymology, as well as in a recent publication 52. Moreover, the method detailed above could also be carried out with Gibson assembly instead of PLIC reaction, using regular primers, with some adaptations.
- Since not all PCR reactions yielded detectable products, the missing reactions were repeated (using the same program as above). For step 3, the annealing temperature parameter was modified to 62 °C (-1 °C/cycle), while for step 6, it was changed to 50 °C. This yielded product for linker lengths 1-3 (Figure 2b). To obtain PCR product for 4 and 5, the program from table S5 was used. In addition, one can add DMSO (3-10 %) to a PCR reaction. In our case, we added two PCR reactions that included 5 μL DMSO (10 % v/v).
- For this chapter, the PCR products were around 6,800 bp, since the vector is 6,789 bp. Hence, 0.02 – 0.04 pmol μL-1 corresponds to a DNA concentration of 84-168 ng μL-1.
- Optionally, an additional control can be added where 4 μL of PCR reaction is mixed with 2 μL of dH2O, to have an indication of whether the PLIC reaction worked.
- For this scenario case, variants with the 1-3 and 6-15 linker length were obtained, while the constructs with 4 and 5 amino acid were not.

- E. coli NEB 10-β is a derivate strain from DH10B, commonly preferred for high quality plasmid preparation, mostly due the deficiency in endonuclease I (endA1) and recombinase (recA1).

# PURIFICATION OF HIS$_{X6}$ TAGGED PROTEINS

This section describes the expression and purification of the obtained fused variants and the single components TmCHMO and TbADH after the cloning and sequence verification of the plasmid of each construct. The expression plasmids have a pBAD backbone that included a N-terminal hexa-histidine tag to facilitate the further purification. The pBAD expression system is tightly controlled using L-arabinose and within the plasmids contains the gene for ampicillin resistance. For an optimal yield, some aspects of protein expression should be considered, such as the evaluation of the optical density of the culture before induction, concentration of the inducer, time and temperature of induction, oxygenation levels, different types of plasmid backbones and host strains[53,54]. Then, after cell lysis the enzymes are obtained through immobilized metal affinity chromatography (IMAC) using nickel-sepharose HP resin. To increase the speed of the purification process, the resin is bed in a gravity-flow column. For a more precise purification, other systems can be employed, such as fast protein liquid chromatography (FPLC). Using FPLC, the purity of the sample can be even higher, but it would be more time consuming. Finally, the level of expression and purity of the proteins can be evaluated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). The fusions were well expressed in the host, and yields in the range of 300 mg L$^{-1}$ were obtained. Additionally, electrophoresis showed the expected size for the variants according to the molecular size estimation.

## Equipment
- Shaker (Innova 44, New Brunswick Scientific).
- Cooling centrifuge (Centrifuge 5804 R, Eppendorf).
- Sonicator (Sonics, Vibra Cell).
- Orbital rotator (Nutating Mixer, VWR International).
- Electrophoresis system (Power Pac HC and Mini-Protean Tetra system, Bio Rad).
- Nano drop spectrophotometer (NanoDrop 1000, Thermo Scientific).

## Buffers, strain and reagents
- Sterile LB medium (for 1 L, add 10 g tryptone, 10 g NaCl and 5 g yeast extract). For LB agar, add agar-agar at 2 % w/v final concentration.
- Sterile TB medium (for 1 L, add 12 g tryptone, 24 g yeast extract, 5 g glycerol and after autoclaving 100 mL sterile 10x TB salts). For 1 L of 10x TB salts dissolve 23.1 g KH2PO4 and 125.4 g K2HPO4.

- Lysis buffer (50 mM Tris HCl pH 8.0, 500 mM NaCl, 30 µM FAD, 1 mM PMSF, 10 mM MgCl2, 0.5 mg mL-1 DNAse I and 1 mg mL-1 lysozyme).
- Buffer A (50 mM Tris HCl pH 8.0, 500 mM NaCl and 5 mM imidazole).
- Buffer B (50 mM Tris HCl pH 8.0, 500 mM NaCl and 400 mM imidazole).
- Buffer C (50 mM Tris HCl pH 8.0 and 150 mM NaCl).
- Nickel-Sepharose resin (GE Healthcare).
- Ampicillin stock at 50 mg mL-1.

## Procedure

1. *E. coli* NEB 10-β glycerol stocks containing the plasmids with the cloned genes were taken from -80 °C freezer and kept on a cooling tube rack. In parallel, prepare sterile culture tubes with 5 mL LB supplemented with 50 µg mL$^{-1}$ ampicillin (see note 1).
2. Inoculate and grow the pre-cultures in a shaker at 37 °C with constant agitation (135 r.p.m.) over night.
3. The next morning, in a 250 mL flask inoculate at 1:100 in 50 mL TB medium supplemented with 50 µg mL$^{-1}$ ampicillin.
4. Incubate the flasks at 37 °C until desired $OD_{600nm}$ and induce at optimal expressions conditions (see note 2).
5. Centrifuge for 15 min at 3,000 g at 4 °C and discard the supernatants.
6. Gently suspend the pellet in 10 mL lysis buffer. Optionally, to facilitate the resuspension of the pellet in the lysis buffer, can be used syringe with a sharp needle.
7. Transfer the suspension into a cooled sonication vessel or into a plastic or glass container that is placed inside a beaker with water and ice to effectively cool the suspension while it is subjected to sonication.
8. Sonicate for 10 min at 70 % amplitude, 5 sec on and sec off.
9. To remove the insoluble fraction (cell debris, insoluble proteins) centrifuge the obtained cell extracts (CE) for 60 min at 15,000 g at 4 °C.
10. Filter the cell free extracts (CFE) using 0.45 µm filter(s).
11. Wash 3 mL of the nickel sepharose resin in a gravity flow column with 3-5 column volumes of milli-Q grade water.
12. Equilibrate the resin with at least 5 column volumes of buffer A.
13. Transfer the filtered CFEs and the 3 mL pre-equilibrated resin in suitable tubes and incubate them for 60 min in an orbital rotator at 4 °C.
14. Transfer each sample into a gravity flow column and discard the flow-through (eluate).
15. Wash with 10-15 column volumes of buffer A (see note 3).
16. Elute with 5 mL of buffer B (see note 4).
17. To remove the imidazole of the protein mixture, use a commercial size exclusion chromatography column, such as Econo-Pac® 10DG desalting prepacked gravity flow columns. Another suitable desalting column can also be used (if you will use another desalting column, follow the manufacturer instructions and skip until step 20). First,

wash the column with 20 mL of buffer C.

18. Subsequently, after buffer C reaches the top frit, load up to 3.3 mL of the eluted sample obtained in step 16 and discard the first 3.3 mL.

19. Add 4 mL of buffer C and collect the eluted 4 mL.

20. Determine the amount of protein using a Nano drop spectrophotometer.

21. To determine the level of purity, analyze the desalted samples through SDS-PAGE. Load 1-2 µg of purified protein per well (**Figure 3**).

22. Finally, flash freeze the protein samples with liquid nitrogen and store them at -80 °C until use. Try to prepare small volumes aliquots (equal to or less than 500 µL).
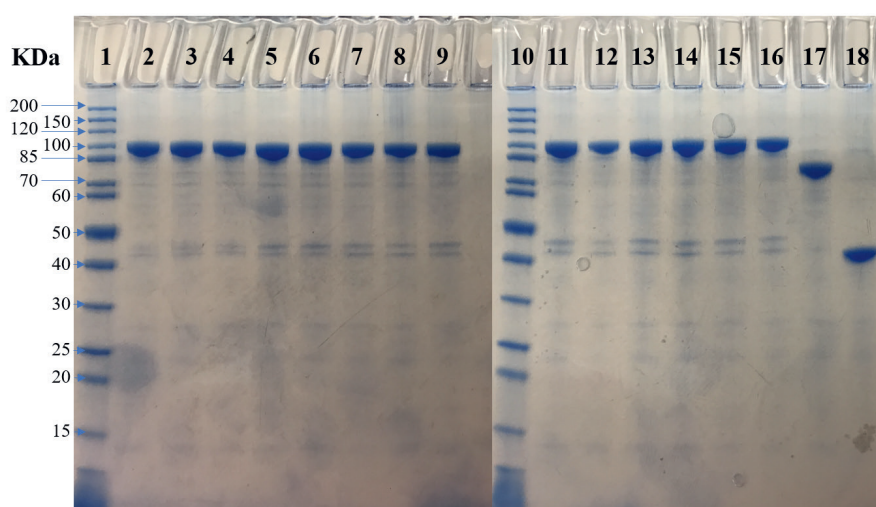


**Figure 3. SDS-PAGE of linker variants and single enzymes.** After purification, 1 µg of protein was loaded per well in an SDS-PAGE. Each well was named according to the size of the amino acid linker. (1) Protein ladder, (2) 1 amino acid (aa), (3) 2 aa, (4) 3 aa, (5) 6 aa, (6) 7 aa, (7) 8 aa, (8) 9 aa, (9) 10 aa, (10) protein ladder, (11) 11 aa, (12) 12 aa, (13) 13 aa, (14) 13 aa, (15) 14 aa, (16) 15 aa, (17) TmCHMO and (18) TbADH. The expected size of the fusions is ~102 KDa.

## Notes

1. During the inoculation and induction, the media (LB and TB), glasses, pipette tips and flasks must be autoclaved. During the protein purification process, keep the samples on ice the whole time.

2. For the pBAD expression system, first test small culture volumes in different expression conditions. We preferred to test 5 mL culture at 0.002-0.2 % w/v L-arabinose and incubate at 17, 24, 30 or 37 °C for 16, 24, 48 or 72 h. For this method, cells were induced at $OD_{600nm}$ 0.8 with 0.02 % L-arabinose for 48 h at 17 °C at 135 r.p.m. For more information about high level expression and purification of histidine tagged proteins see the handbook The QIAexpressionist [53].

3. Optionally, follow the absorbance at $280_{nm}$ and continue the washing with the binding buffer until the absorbance reach a steady baseline.
4. Flavin-dependent proteins, such as TmCHMO due to their bound prosthetic cofactor, have a characteristic UV-visible spectrum with high absorbance about 440nm. Consequently, these enzymes display a distinctive yellow color. Therefore, during purification, if there are relatively high amounts of protein, it is possible to easily see the protein mobility within the resin.

## UV-VISIBLE SPECTRAL ANALYSIS OF TMCHMO

As described above, TmCHMO contains FAD as prosthetic group. The bound flavin cofactor provides a useful probe to study and characterize the enzyme based on the UV-visible absorbance features of FAD[55]. FAD can exist in three different redox states: 1) the two-electrons reduced flavin hydroquinone, 2) the one-electron reduced flavin semiquinone and 3) the fully oxidized flavin[56]. Each redox state has a specific UV-visible absorbance spectrum, hence the redox state of the protein can be easily determined by spectrophotometric analysis[57]. Moreover, FAD has various protonation states, which alters the spectrophotometric properties. For example, the flavin semiquinone state absorbs at relatively long wavelengths ($>550_{nm}$), while the anionic flavin semiquinone displays a strong absorbance at $380_{nm}$[55,56]. Particularly, oxidized TmCHMO exhibits two absorbance maxima, at $376_{nm}$ and $440_{nm}$[58]. TmCHMO in the oxidized state has a molar extinction coefficient of 14.0 mM$^{-1}$ cm$^{-1}$ at $440_{nm}$, similar to other described flavoproteins. Another useful spectral feature for flavin-containing proteins is the $A_{280}$:$A_{440}$ absorbance ratio. The absorbance at $280_{nm}$ can be used to determine the amount of total protein, due to absorbance of aromatic amino acids. By measuring the $A_{280}$:$A_{440}$ it is possible to estimate the amount of holo protein (= FAD-containing enzyme) in a sample.

For this section, the absorbance spectra of the purified enzyme variants were recorded and the UV-visible spectra were analyzed. For all the fused enzymes, the typical spectral characteristics of oxidized FAD were observed (**Figure 4**). Furthermore, all tested linkers allowed binding of the FAD cofactor because the $A_{280}$:$A_{440nm}$ values ranged between 11.4-12.5 (**Table 1**).

**2**
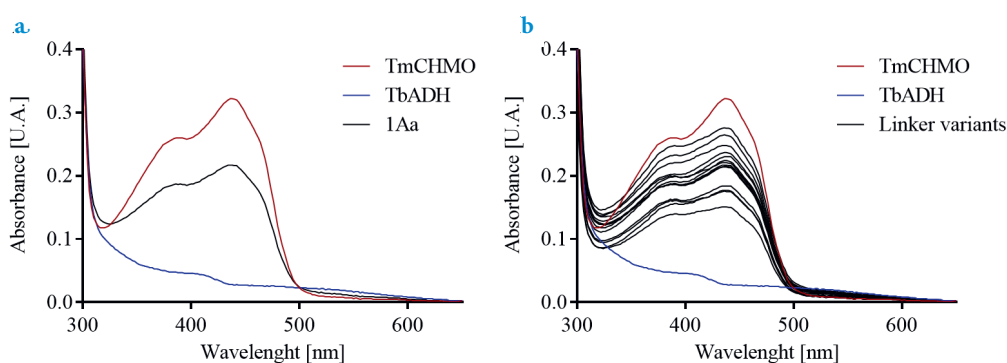


**Figure 4. Spectrophotometric determination of fused variants**. For each variant, the UV-Visible spectrum (300-600$_{nm}$) was recorded. For all purified FAD-containing proteins, the fully oxidized redox state and the distinctive peaks at 376 and 440$_{nm}$ were observed. The full spectra of (a) TmCHMO, TbADH and one linker variant is shown, and (b) for all the evaluated enzymes.

**Table 1. Spectrophotometric UV-Visible evaluation of fused variants.** The concentration of each variants was spectrophotometrically determined using the molar extinction coefficient of TmCHMO ($\varepsilon_{440nm}$ = 14.0 mM$^{-1}$ cm$^{-1}$). Additionally, the $A_{280}$:$A_{440}$ ratios were calculated for all the constructs.

| Variant | Yield [mg L$^{-1}$] | $A_{280}$:$A_{440}$ |
|---|---|---|
| TmCHMO | 500 | 11.7 |
| TbADH | 400 | - |
| 1 Aa | 300 | 12.5 |
| 2 Aa | 330 | 12.0 |
| 3 Aa | 360 | 12.4 |
| 6 Aa | 300 | 11.4 |
| 7 Aa | 330 | 12.1 |
| 8 Aa | 280 | 12.2 |
| 9 Aa | 260 | 12.6 |
| 10 Aa | 290 | 11.4 |
| 11 Aa | 350 | 10.9 |
| 12 Aa | 270 | 13.0 |
| 13 Aa | 350 | 11.7 |
| 13 Aa$^{*}$ | 350 | 11.9 |
| 14 Aa | 350 | 12.2 |
| 15 Aa | 250 | 12.3 |

## Equipment

- UV-visible spectrophotometer (see note 1) (Synergy H1 Microplate Reader, BioTek).

## Buffer

- Buffer A (50 mM Tris HCl pH 8.0 and 150 mM NaCl).

## Procedure

1. Take the protein samples from the -80 °C freezer to 0 °C (transfer them in a bucket or holder on ice). Keep the samples on cold until completely thawed, depending on the volume of the aliquot it should thaw in about 15 min.
2. Turn on the spectrophotometer. Leave it for at least 20 min which will allow the lamp to heat up and to measure more reliably. Depending on the spectrophotometer this time can vary. For optimal instrument performance follow the instructions of the manufacturer.
3. Measure buffer A as a blank (see note 2).
4. Prepare a first measurement and verify that the absorbance at $440_{nm}$ is equal to or less than 0.5 [A.U.]. Dilute with buffer A if necessary.
5. When the UV-visible spectrum has the requested absorbance, record the absorbance at $600_{nm}$, this value is later subtracted as protein background.

## Notes

1. For this section, a microplate reader with absorbance correction and 96-well plates were used. For each measurement, 300 µL protein samples were analyzed.
2. If a spectrophotometer and reusable cuvettes are used, cuvettes must be properly washed after each measurement. Optionally, wash the cuvette with a 20 % ethanol solution, then wash with milli-Q grade water and finally dry it.

## APPARENT MELTING TEMPERATURE DETERMINATION

Assessing the manner in which enzymes respond to temperature variation can give useful biophysical information. In general, thermostability analysis can be approached thermodynamically or kinetically[59-61]. Various methods have been extensively described to determine the temperature effect on purified protein samples, mainly by monitoring the perturbations on the secondary or tertiary structure by circular dichroism, differential scanning calorimetry, absorbance or fluorescence[62-67]. A faster, cheaper and easier alternative, is to evaluate the apparent melting temperature ($T_M^{app}$) by measuring fluorescence changes within a temperature gradient. The $T_M^{app}$ parameter is defined as the midpoint of the protein unfolding transition[68]. The ThermoFluor technique is a well-described method which requires a small amount of sample[69]. This technique permits to determine the protein $T_M^{app}$ by measuring the fluorescence intensity of a solvatochromic dye, such as ANS (1-anilino-8-naphthanlenesulfonate) or SYPRO orange[70,71]. The fundament of the method lies in the difference of the quantum yield of the dye when it is in water compared with its fluorescence in a nonpolar environment, such as when bound to hydrophobic moieties of unfolded proteins. During a temperature ramp, the sample is excited at a specific wavelength and the emission is recorded. At increasing temperatures,

the protein begins to unfold exposing its hydrophobic patches. The dye binds to the unfolded protein, increasing its fluorescence intensity. A similar technique, denominated ThermoFAD, is based on the same concept[72]. Yet, in ThermoFAD a solvatochromic dye is not required as it takes advantage of the intrinsic fluorescence of the flavin cofactor. The fluorescence of flavin cofactors is typically quenched when bound in a protein, resulting in a fluorescence increase when the protein unfolds. During the temperature ramp, the sample is excited at a wavelength range between 450 and $490_{nm}$ and the fluorescence is measured using an emission filter of $515$-$530_{nm}$. The data obtained is a sigmoidal curve of the fluorescence intensity over time. The apparent melting temperature is taken at the temperature at which the highest slope of the curve is observed. For this section, both techniques (ThermoFluor and ThermoFAD) were used and the results were summarized in **table 2**.

**Table 2. Apparent melting temperature determination.** Using a temperature ramp, thermostability for all constructs was measured using both ThermoFAD[a] and ThermoFluor[a,b] techniques. Duplicates results are shown. Each experiment was performed in duplicate.

| Variant | TmCHMO [°C][a,b] | TbADH [°C][b] |
|---------|------------------|---------------|
| TmCHMO | 53.0 | - |
| TbADH | - | 93.5 |
| 1 Aa | 51.5 | 90.5 |
| 2 Aa | 52.0 | 90.0 |
| 3 Aa | 52.0 | 93.0 |
| 6 Aa | 52.0 | 91.5 |
| 7 Aa | 52.0 | 91.0 |
| 8 Aa | 52.0 | 91.0 |
| 9 Aa | 52.0 | 91.0 |
| 10 Aa | 52.0 | 91.0 |
| 11 Aa | 52.0 | 90.5 |
| 12 Aa | 51.5 | 91.5 |
| 13 Aa | 52.0 | 91.5 |
| 13 Aa[*] | 52.0 | 91.0 |
| 14 Aa | 52.0 | 91.0 |
| 15 Aa | 52.0 | 91.0 |

## Equipment

- Real-time PCR detection system (CFX96 C1000 Touch Thermal Cycler, Bio-Rad).
- 96-Well qPCR Plates (Bio-Rad).
- Adhesive qPCR Plate Sealing Film (Bio-Rad).
- Bench centrifuge with plate adaptor (Centrifuge X-4001, Salmenkipp).

## Buffer

- Buffer A (50 mM Tris HCl pH 8.0 and 150 mM NaCl).
- Fluorescent dye as convenient, SYPRO orange was used for this section (see note 2).

## Procedure

1. Take the protein samples from the -80 °C freezer to 0 °C (transfer them in a bucket or holder on ice). Keep the samples on cold until completely thawed, depending on the volume of the aliquot it should thaw in about 15 min.
2. The concentration of the protein sample must be at least 1 mg mL$^{-1}$. Dilute with buffer A if is necessary. Perform the experiments in triplicate.
3. On the real-time PCR detection system (qPCR instrument), configure the temperature ramp program. The program has a gradient from 20 to 99 °C that measures the fluorescence intensity every 0.5 °C after a 10 sec delay for signal. Depending on the technique and dye, configure a suitable excitation and emission filter (see note 2).
4. Pipette 20 µL of protein sample into a well of the specified qPCR plate. Seal the plate with a sealing film (note 3 and 4).
5. Transfer the plate to the instrument and start the program.
6. Depending on the instrument in use, after the temperature gradient the slope of the sigmoidal curve can be plotted to determine the apparent melting temperature.

## Notes

1. Although there are some spectral changes among flavin-containing proteins, in general members of this class of enzymes show a fluorescence excitation maximum at 370-380$_{nm}$ and 440-450$_{nm}$. While for the fluorescence emission, the maximum intensity is at 535$_{nm}$.
2. For the ThermoFluor technique, SYPRO orange dye was used. The real-time PCR detection system was configured to excite between 450-490$_{nm}$ and detect fluorescence intensity using an emission filter in the 560-580$_{nm}$ range. If another solvatochromic dye is used, follow the instruction of the manufacturer.
3. Discard the bubbles of the plate by centrifugation. Use a delicate task wiper to clean fingerprints and impurities from the seal and to fix the plate on the instrument.
4. Both ThermoFAD and ThermoFluor techniques are extremely convenient: using 96-well qPCR plates it is possible to analyze up to 96 samples at the same time in less than 1 hour. Furthermore, depending on the objective of the study, the effect of the pH, miscible cosolvents, storage conditions, cosubstrates or presence of different additives can be easily analyzed.

# EVALUATION OF THE LINKER SIZE ON THE ENZYME ACTIVITY

For biocatalytic purposes, it is highly valuable to determine the effect of the linker length on enzyme activity. The activity of TmCHMO and TbADH can be readily measured by spectrophotometrically following the absorbance of NADPH at $340_{nm}$ ($\varepsilon_{340nm}$ = 6.22 mM$^{-1}$ cm$^{-1}$). For ADHs, the reaction is monitored by the measuring the formation of the reduced nicotinamide cosubstrate[73]. While for TmCHMO, the oxidation rate is followed[58]. However, determining the activity for fused enzymes that share their coenzyme is not trivial. In this case, one enzyme can disrupt the measurement for the other. Specifically, the reversible reaction of TbADH could affect the kinetic analysis for TmCHMO by NADPH-mediated reduction of the ketone. Nevertheless, for this scenario, the broad substrate acceptance of TmCHMO can be exploited and activity can be analyzed using an alternative substrate. TmCHMO can oxidize S-containing substrates, such as thioanisole (phenyl methyl sulfide). The use of this thioether as BVMO test substrate prevents interfering ADH activity. For this section, the oxidation of thioanisole (TmCHMO activity) or cyclohexanol (TbADH activity) was spectrophotometrically monitored.

## Equipment

- UV-visible spectrophotometer with orbital rotor (see note 1) (Synergy H1 Microplate Reader, BioTek).
- Multi pipette suitable for volumes of 300 μL.
- 96-well plates.

## Buffer and reagents

- Buffer A (50 mM Tris HCl pH 8.0 and 150 mM NaCl).
- NADP+ stock solution at 300 μM prepared in buffer A.
- NADPH stock solution at 300 μM prepared in buffer A.
- Cyclohexanol stock solution at 40 mM prepared in buffer A.
- Protein sample at 2.0 μM prepared in buffer A (note 2).
- Thioanisole (phenyl methyl sulfide) stock solution at 4.0 mM prepared in buffer A and 10 % v/v 1,4-dioxane as cosolvent.

## Procedure

1. Take the protein samples from the -80 °C freezer to 0 °C (transfer them in a bucket or holder on ice). Keep the samples on cold until completely thawed, depending on the volume of the aliquot it should thaw in about 15 min.
2. Turn on the spectrophotometer, this allows the lamps to heat up and measure in a more reliably way. Leave the spectrophotometer on for at least 20 min, depending on the equipment this time can vary, for a better instrument performance follow the instructions of the manufacturer.

3. In the spectrophotometer, configure a kinetic program measuring absorbance at $340_{nm}$ for 120 sec with constant orbital shaking, include a pathlength correction. Due the sensor delays between measurements, prepare two rows for each kinetic analysis.
4. For the kinetic analysis:
   A. For dehydrogenase activity, pipette 75 µL of cyclohexanol solution and 75 µL of protein sample into each well.
   B. For sulfide oxidation measurement, pipette 75 µL of thioanisole solution and 75 µL of protein sample into each well.
5. Gently suspend the solution by pipetting, avoid the formation of bubbles.
6. Using a multi pipette, transfer 150 µL of the NADP⁺ solution for the TbADH kinetic analysis. While for TmCHMO, transfer 150 µL of the NADPH solution (note 3). Quickly transfer the plate into the microplate reader and start the measurement.
7. After the kinetic measurement, continue with the next two rows.
8. To perform the analyze, plot the corrected absorbance values over time in seconds. First, determine the slope of the initial linear range on [A.U.] s⁻¹ units.
9. Then, translate the obtained slope to the rate of NADPH per seconds using the molar extinction coefficient of NADPH ($\varepsilon_{340nm} = 0.00622 \, \mu M^{-1} \, cm^{-1}$). The obtained units are in [$\mu M \, s^{-1}$].
10. Finally, correct the obtained data for the enzyme concentration in order to calculate the $k_{obs}$ [s⁻¹] (**Figure 5**).

## Notes

1. A microplate reader (with absorbance correction) and 96-well plates were used for this section. For each reaction, the final volume was 300 µL. If a spectrophotometer and reusable cuvettes are used, the cuvettes must be properly washed at each measurement. Optionally, wash the cuvette with 20 % ethanol solution, then wash with milli-Q grade water and finally dry it.
2. For the 2.0 µM protein solution in buffer A, keep all the tubes on ice until the reaction starts. Perform the experiments in triplicate.
3. For kinetic analysis, measurements must be rapid to avoid loss of the absorbance change of the first few seconds. Thereby, prepare all the reactions near to the instrument. After transferring NADPH or NADP⁺ solution to the wells, quickly mix using the multi pipette avoiding the formation of bubbles. Subsequently, transfer the plate in the spectrophotometer and start the measurement.
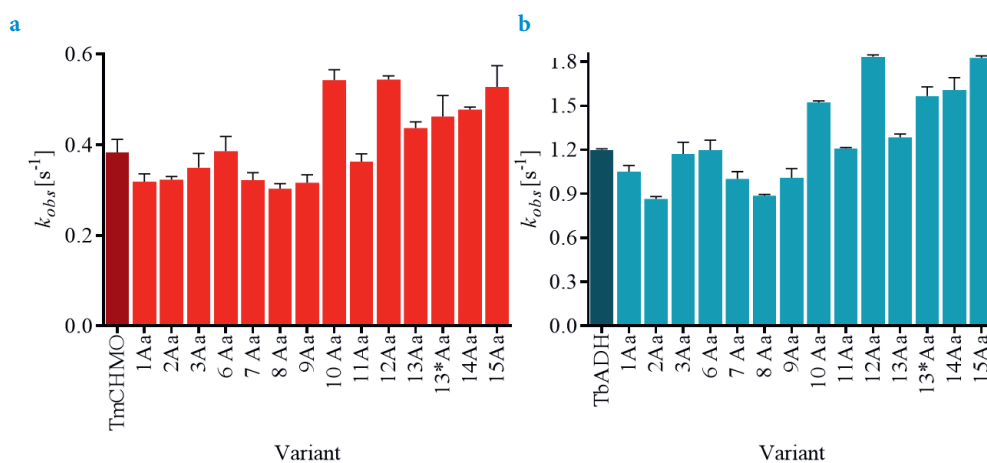
**Figure 5. Effect of linker length on enzyme activity.** For each variant, alcohol dehydrogenase and sulfoxidation activity was evaluated. (A) For sulfoxidation activity, NADPH oxidation rates were studied using 1.0 mM thioanisole and 150 µM NADPH. (B) For alcohol dehydrogenase activity, the NADP+ reduction rates were measured at 10 mM cyclohexanol and 150 µM NADP+. . As a control, the single non-fused enzymes were also evaluated. Each experiment was performed in triplicate.

## EVALUATION OF FUSED VARIANTS IN SMALL-SCALE BIOCONVERSIONS

As a final biocatalytic evaluation, the linker-variants can be tested in small-scale conversions and the dimensionless turnover number (TON) can be determined. The TON is defined as the ratio of the obtained moles of product divided by the moles of used enzyme. For this calculation, the cyclohexanol conversion and the produced ketone and lactone are monitored. By varying the experimental conditions, such as the temperature or pH, optimal parameters for biocatalytic performance of the fusion enzymes can be determined (see note 1). Due the chemical nature of the substrate and products, the conversions were monitored using gas chromatography (GC). This technique resolves analytes based on their volatility and polarity[74]. For this section, compounds were analyzed and quantified using a GC coupled to a mass spectrometer (MS) with electron ionization and quadruple separation. GC is nowadays often coupled to MS as it results in a powerful analytical tool[75,76]. The general principle of coupled GC-MS is the same for all other coupled MS techniques. Broadly, the GC instrument separates the components into fractions which are transferred to the MS module. Then, the ions are detected, resulting in a MS spectrum for each analyte. Finally, a total ion current of the continuous mass scanning is obtained[77]. GC-MS is a mature and sensitive technique, which is often equipped with powerful data analysis software for the identification of molecular ions or distinctive fragmented ion patterns. Finally, the identification is automatically performed

by comparing with spectral MS libraries, such as databases from the National Institute of Standards and Technology (NIST), Golm Metabolome Database (GMD), or the Human Metabolome Database (HMDB)[74,77].

## Equipment

- Shaker (Innova 44, New Brunswick Scientific).
- Vortex.
- Bench centrifuge (Heraeus Fresco 17 Centrifuge, Thermo Scientific).
- GC-MS (GC-MS QP2010 ultra with electron ionization and quadrupole separation, Shimadzu) (note 2).
- HP-5MS column (Agilent, 30 m x 0.25 mm x 0.25 μm).

## Buffer and reagent

- Buffer A (50 mM Tris HCl pH 8.0 and 150 mM NaCl).
- Buffer B (50 mM Tris HCl pH 8.0, 40 μM FAD and 150 mM NaCl).
- Protein mixtures at 2.0 μM in buffer B.
- NADP+ stock solution at 300 μM prepared in buffer A.
- NADPH stock solution at 300 μM prepared in buffer A.
- Cyclohexanol stock solution at 60 mM prepared in buffer A.
- Cyclohexanone stock solution at 60 mM prepared in buffer A.
- ε-caprolactone standard solutions (1.0-15 mM).
- FAD stock solution at 1.0 mM in buffer A.
- Ethyl acetate with 0.025 % v/v mesitylene (external standard).
- MgSO4 anhydrous.

## Procedure

1. Take the protein samples from the -80 °C freezer to 0 °C (transfer them in a bucket or holder on ice). Keep the samples on cold until completely thawed, depending on the volume of the aliquot it should thaw in about 15 min.
2. Transfer 125 μL of cyclohexanol solution in a 10 mL vial (see note 3) (for a TmCHMO control reaction, transfer 125 μL of cyclohexanone solution). Subsequently, transfer 125 μL of enzyme mixture.
3. Subsequently, transfer 250 μL of NADP+ solution to the 10 mL vial and incubate them in a shaker at 25 or 37 °C for 24 h (for the TmCHMO control reaction, add 250 μL of NADPH solution).
4. After incubation, transfer the solution in a 2 mL tube and add 500 μL of ethyl acetate (1 volume reaction).
5. Vortex the tubes for 30 sec, subsequently centrifuge at 16,000 g and transfer the organic layer to another tube.
6. Repeat step 4 and 5 two more times collecting the organic layer in a same tube.

7. Add a spatula of $MgSO_4$ (roughly 5 mg) in the organic solution and vortex it for 30 sec to remove the residual water for further GC analysis.

8. Centrifuge the tubes at 16,000 g and transfer the organic supernatants in GC-vials (use suitable inserts if necessary).

9. Perform the same extraction procedure for ε-caprolactone standard solutions to prepare a calibration curve.

10. For the GC analysis, configure the temperature program as:
    o   Hold 5 min at 50 °C.
    o   Increase 5 °C min$^{-1}$ for 2 min.
    o   Increase 2 °C min$^{-1}$ for 10 min.
    o   Hold for 10 min.

11. To calculate the substrate conversions, integrate the peaks of the substrate, mesitylene and product(s) from the chromatogram. Then, normalize the integrated values with the area of the external standard. Compare the results with a control reaction carried out without enzyme and quantify the percentage of conversion.

12. Finally, if cyclohexanone is observed in the chromatogram, calculate the production of ε-caprolactone using the calibration curve and determine the TON values (note 5) (**Table 3**).

## Notes

1. To find optimal conditions for conversion, two enzyme concentrations (0.5 and 5 μM), two substrate concentrations (10-20 mM) and two different temperatures (24 and 37 °C) were tested. The condition of 0.5 μM enzyme and 15 mM cyclohexanol at 37 °C for 24 h allowed to observe differences between the TON values when comparing the performance of the different fusion enzymes, avoiding full or very low substrate conversions.

2. If the mass spectrometer is not available for the identification, bioconversions can also be determined by regular GC analysis. In that case, the analytical standards of cyclohexanol, cyclohexanone and ε-caprolactone must be previously tested.

3. For bioconversions using monooxygenases as biocatalysts, the liquid: gas volume ratio must allow enough dioxygen for the reaction. For this methodology, 500 μL reaction volumes were prepared in 10 mL vial.

4. For the analysis and identification of the molecules, LabSolutions GC-MS software from Shimadzu with a MS-spectral library from NIST was used. Another suitable GC-MS and library may be used as preferred.

5. If the presence of cyclohexanone in the chromatogram is not observed, the conversion of the substrate can be used for the TON determination. Standards with cyclohexanol, cyclohexanone and ε-caprolactone must be evaluated in advance.

**Table 3. Bioconversions analysis of self-sufficient bifunctional fusions.** Conversion levels and TONs for linker variants and the single enzymes are shown. For each bioconversion 15 mM cyclohexanol and 0.5 µM of enzyme were used. Reactions were performed in buffer A at 37 °C and 135 r.p.m. for 24 h. Conversions were calculated based on the substrate depletion from the GC-MS analysis. The turnover number was calculated as the ratio of moles of product obtained divided by the moles of enzyme used in the reaction. Each experiment was performed in duplicate.

| Variant | Conversion [%] | TON±SEM |
|---|---|---|
| TbADH + TmCHMO | 39 | 11,700+600 |
| 1 Aa | 57 | 17,100±150 |
| 2 Aa | 65 | 19,500±600 |
| 3 Aa | 66 | 19,800±900 |
| 6 Aa | 62 | 18,600±900 |
| 7 Aa | 69 | 20,700±1800 |
| 8 Aa | 55 | 16,500±600 |
| 9 Aa | 49 | 14,700±1200 |
| 10 Aa | 58 | 17,400±900 |
| 11 Aa | 56 | 16,800±300 |
| 12 Aa | 46 | 13,800±300 |
| 13 Aa | 69 | 20,700±1200 |
| 13 Aa[*] | 70 | 21,000±300 |
| 14 Aa | 85 | 25,550±900 |
| 15 Aa | 48 | 14,400±900 |

## SUMMARY AND CONCLUSIONS

This chapter explores the effect of the linker size on the biocatalytic properties of a self-sufficient multifunctional TbADH-TmCHMO fusion biocatalyst. 14 different linker variants of the previously reported glycine-rich sequence were obtained and evaluated. A variation of the 13-amino acid-linker was also obtained (G13D). All the obtained variants exhibited a high expression level (250 and 360 mg L$^{-1}$) and were obtained as FAD-containing proteins. The latter is suggested by the obtained $A_{280}$:$A_{440}$ ratios. The protocol for enzyme purification includes an excess of FAD during cell lysis which may promote holo enzyme formation. It was previously reported that *E. coli* cells may not be able to support biosynthesis of sufficient amounts of the FAD cofactor[22,78]. Regarding the thermostability, for TmCHMO —using the ThermoFAD method— it was observed that the length of the linker did not show a noticeable effect on its unfolding profile. On the other hand, using the ThermoFluor method —where the $T_M^{app}$ of ADH is also monitored— the linker size had only a small effect compared with the not-fused TbADH. For all the constructs, the ADH fusion partner exhibited a somewhat lower $T_M^{app}$. For 3 variants the differences were up to 3 °C. Nevertheless, for TbADH this effect is not operationally significant due its

relatively high thermostability ($T_M^{app}$ = 93.5 °C). Regarding the effect on enzyme activity, the alcohol oxidation and sulfoxidation activities were similar for the 9 shortest variants. The fusions with the length of 10,12 and 15 amino acids, showed for both activities a slightly increased $k_{obs}$. Interestingly, the G13D variant exhibited a minor increase in the ADH activity, yet the same sulfoxidation activity when compared with the other 13 amino acids linker.

The small-scale bioconversions resulted in highest turnover numbers of the fusions with 2, 3, 6, 7, 13 and 14 amino acid linkers (TONs between 20,000-25,000). While for the reaction with the single enzymes, the obtained TON was around 12,000. Clearly, the use of optimal linkers results in superior performance. Where the fusion with the 14 amino acids linker displayed the highest TON and seems to be the best candidate for biocatalytic purposes. For the conversions, both 13 amino acids length variants showed the same TON. Intriguingly, the conversion level for the 15 amino acids linker was rather low (TON=14,400). This variant exhibited the highest activity for the alcohol oxidation and sulfoxidation. There is not a clear explanation for this observation. Perhaps it is related to a product inhibition issue. The drop in the conversion level is unlikely to stem from a lower protein stability. There were no significant differences in the $T_M^{app}$ values. Observed effects of different linkers may be correlated to a change in the orientation and/or interaction between the fusion partners, that may affect the TON. Measurements of biocatalyst lifetime robustness would be valuable to establish the best-performing linker.

This contribution provides an easy and efficient protocol to generate a collection of fusion enzymes with different linkers. This allowed to determine the effect of the linker size on the properties of the self-sufficient TbADH-TmCHMO fusions for the synthesis of ε-caprolactone. The employed procedure can be easily adapted for generating a similar library of fusion enzymes from other enzymes. Also, the other protocols, such as the protocol to determine apparent melting temperatures of a collection of (flavo)proteins, can be of value in other enzyme engineering or biocatalytic studies. Such easy-to-use protocols will help to improve procedures for generating optimized biocatalysts.

# SUPPORTING INFORMATION CHAPTER II

## Tables

**Table S1. List of primers for the linker variant design.**

| | PTO region (12)[1] | Linker region | Gene-overlapping region |
|---|---|---|---|
| RV1 | GGCCAGAATAACAAC | – | TGGTTTGATCAGATC |
| $T_m$[2] | 47 °C | – | 45 °C |
| FW1 | GTTGTTATTCTGGCC | TCG | ATGAGCACCACCCAGACCCCG |
| FW2 | GTTGTTATTCTGGCC | TCGAGT | ATGAGCACCACCCAGACCCCG |
| FW3 | GTTGTTATTCTGGCC | TCGAGTGGT | ATGAGCACCACCCAGACCCCG |
| FW4 | GTTGTTATTCTGGCC | TCGAGTGGTGGC | ATGAGCACCACCCAGACCCCG |
| FW5 | GTTGTTATTCTGGCC | TCGAGTGGTGGCTCT | ATGAGCACCACCCAGACCCCG |
| $T_m$[2] | 47 °C | 10-56 °C | 69 °C |
| RV2 | AGAGCCACCACT[3] | – | CGAGGCCAGAATAACAACTGGTTTG |
| $T_m$ | 53 °C | – | 68 °C |
| FW6 | AGTGGTGGCTCT | GGT | <u>ATGAGCACCACCCAGACCCCG</u>GACC |
| FW7 | AGTGGTGGCTCT | GGTGGG | <u>ATGAGCACCACCCAGACCCCG</u>GACC |
| FW8 | AGTGGTGGCTCT | GGTGGGAGC | <u>ATGAGCACCACCCAGACCCCG</u>GACC |
| FW9 | AGTGGTGGCTCT | GGTGGGAGCGGT | <u>ATGAGCACCACCCAGACCCCG</u>GACC |
| FW10 | AGTGGTGGCTCT | GGTGGGAGCGGTGGC | <u>ATGAGCACCACCCAGACCCCG</u>GACC |
| FW11 | AGTGGTGGCTCT | GGTGGGAGCGGTGGCTCA | <u>ATGAGCACCACCCAGACCCCG</u>GACC |
| $T_m$ | 53 °C | 10-74 °C | 80 °C |
| RV3 | TGAGCCACCGCT[1] | – | CCCACCAGAGCCACCACTCGAGGCC AGAATAACAACTGGTTTG |
| $T_m$ | 59 °C | – | 82 °C |
| FW12 | AGCGGTGGCTCA | GCT | <u>ATGAGCACCACCCAGACCCCG</u>GACC |
| FW13 | AGCGGTGGCTCA | GCTGGT | <u>ATGAGCACCACCCAGACCCCG</u>GACC |
| FW14 | AGCGGTGGCTCA | GCTGGTACC | <u>ATGAGCACCACCCAGACCCCG</u>GACC |
| FW15 | AGCGGTGGCTCA | GCTGGTACCGCG | <u>ATGAGCACCACCCAGACCCCG</u>GACC |
| FW16 | AGCGGTGGCTCA | GCTGGTACCGCGGGC | <u>ATGAGCACCACCCAGACCCCG</u>GACC |
| $T_m$ | 59 °C | 10-71 °C | 82 °C |

[1]The PTO region of the FW primers and RV primer are always the reverse complement of one another. Ensure that this hybridization area has a melting temperature higher than 42 °C to be stable during heat shock.
[2]Melting temperature of region of primer as calculated by NEB Tm calculator (https://tmcalculator.neb.com/). The linker region shows here codes for: SSGGS; the first part of our glycine-rich linker (SSGGSGGSGGSAGTAG).
[3]The PTO region for both RV6-11 and RV12-16 are actually in the linker region, hence the name in the middle column is now 'linker extension region'. Note that the Gene-overlapping region is slightly extended in this set of primers (underlined part is the same as in the 1-5 primers).

**Table S2. Polymerase chain reaction setup.**

| Volume | Components | Final concentration |
|--------|-----------|---------------------|
| 5 μL | 10x HF Polymerase buffer | 1x |
| 1 μL | 50x dNTP mixture | 0.2 mM[1] |
| 1 μL | FW primer | 0.4 μM |
| 1 μL | RV primer | 0.4 μM |
| 0.5 μL | Polymerase (e.g. Phusion) | 5.0 U |
| 1 μL | Template DNA | 0.4 ng μL$^{-1}$ |
| x μL | Sterile MilliQ / dH$_2$O up to 50 μL | |
| 50 μL | Final volume | |

[1]0.2 mM of each dATP, dCTP, dGTP, and dTTP

**Table S3. Thermocycling conditions.**

| Step | Time | Temperature |
|------|------|-------------|
| 1 | 2 min | 98 °C |
| 2 | 30 sec | 98 °C |
| 3 | 30 sec | 68 °C (-1 °C/cycle) |
| 4 | 7 min | 72 °C |
| | | Repeat step 2-4 for 12 cycles |
| 5 | 30 sec | 98 °C |
| 6 | 30 sec | 57 °C |
| 7 | 7 min | 72 °C |
| | | Repeat steps 5-7 for 15 cycles |
| 8 | 11 min | 72 °C |

**Table S4. Iodine solution reaction setup.**

| Volume | Component | Notes |
|--------|-----------|-------|
| 5 μl | TrisHCl buffer (500 mM; pH 9) | Since only a small volume is needed, prepare a low amount and store at 4 °C. |
| 3 μl | Iodine-EtOH solution (100 mM) | Keep the solution protected from light (in a non-transparent tube, or wrapped in aluminum foil), and keep on ice or at 4 °C. It is best to prepare this solution fresh. In our case, we mixed 0.254 g of iodine with 10 mL of 99 % ethanol. |
| 2 μl | dH$_2$O | |
| 10 μl | Total volume | This solution is kept on ice or at 4 °C. This volume can be used for 5 PLIC reactions. |

**Table S5.** Alternative thermoscycling conditions

| Step | Time | Temperature |
|---|---|---|
| 1 | 2 min | 98 °C |
| 2 | 30 sec | 98 °C |
| 3 | 30 sec | 70 °C (-0.5 °C/cycle) |
| 4 | 7 min | 72 °C |
| | | Repeat step 2-4 for 24 cycles |
| 5 | 30 sec | 98 °C |
| 6 | 30 sec | 58 °C |
| 7 | 7 min | 72 °C |
| | | Repeat steps 5-7 for 15 cycles |
| 8 | 11 min | 72 °C |

# REFERENCES

1. Woodley, J. M. New frontiers in biocatalysis for sustainable synthesis. *Current Opinion in Green and Sustainable Chemistry* vol. 21 22–26 (2020).

2. Adams, J. P., Brown, M. J. B., Diaz-Rodriguez, A., Lloyd, R. C. & Roiban, G. Biocatalysis: A Pharma Perspective. *Adv. Synth. Catal.* 361, adsc.201900424 (2019).

3. Delgove, M. A. F., Elford, M. T., Bernaerts, K. V & De Wildeman, S. M. A. Toward Upscaled Biocatalytic Preparation of Lactone Building Blocks for Polymer Applications. (2018) doi:10.1021/acs.oprd.8b00079.

4. Messiha, H. L. *et al.* Biocatalytic Routes to Lactone Monomers for Polymer Production. (2018).

5. Magali, P., Revel, de G. & Marchand, S. First identification of threep-menthane lactones and their potentialprecursor, menthofuran, in red wines. *Food Chem.* 217, 294–302 (2017).

6. Zhu, Y. *et al.* Enantiomeric distributions of volatile lactones and terpenoids in white teas stored for different durations. *Food Chem.* 320, 126632 (2020).

7. Ritz, J., Fuchs, H., Kieczka, H. & Moran, W. C. Caprolactam. *Ullmann's Encycl. Ind. Chem.* (2000).

8. Künkel, A. *et al.* Polymers, Biodegradable. in *Ullmann's Encyclopedia of Industrial Chemistry* 1–29 (Wiley-VCH Verlag GmbH & Co. KGaA, 2016).

9. Dahlhoff, G., Niederer, J. P. M. & Hoelderich, W. F. e-Caprolactam: new by-product free synthesis routes. *Catal. Rev.* 43, 381–441 (2001).

10. Mallin, H., Wulf, H. & Bornscheuer, U. T. A self-sufficient Baeyer--Villiger biocatalysis system for the synthesis of ε-caprolactone from cyclohexanol. *Enzyme Microb. Technol.* 53, 283–287 (2013).

11. Köpnick, H., Schmidt, M., Brügging, W., Rüter, J. & Kaminsky, W. Polyesters. *Ullmann's Encycl. Ind. Chem.* (2000).

12. Labet, M. & Thielemans, W. Synthesis of polycaprolactone: a revieww. (2009) doi:10.1039/b820162p.

13. Thaore, V., Chadwick, D. & Shah, N. Sustainable production of chemical intermediates for nylon manufacture: A techno-economic analysis for renewable production of caprolactone. *Chem. Eng. Res. Des.* 135, 140–152 (2018).

14. Fraaije, M. W. *et al.* Discovery of a thermostable Baeyer--Villiger monooxygenase by genome mining. *Appl. Microbiol. Biotechnol.* 66, 393–400 (2005).

15. Messiha, H. L. *et al.* Biocatalytic Routes to Lactone Monomers for Polymer Production. *Biochemistry* 57, 1997–2008 (2018).

16. Savoca, M. P., Tonoli, E., Atobatele, A. G. & Verderio, E. A. M. Biocatalysis by transglutaminases: a review of biotechnological applications. *Micromachines* 9, 562 (2018).

17. Woodley, J. M. New opportunities for biocatalysis: making pharmaceutical processes greener. *Trends Biotechnol.* 26, 321–327 (2008).

18. Wachtmeister, J. & Rother, D. Recent advances in whole cell biocatalysis techniques bridging from investigative to industrial scale. *Curr. Opin. Biotechnol.* 42, 169–177 (2016).

19. Kadisch, M., Willrodt, C., Hillen, M., Bühler, B. & Schmid, A. Maximizing the stability of metabolic engineering-derived whole-cell biocatalysts. *Biotechnol. J.* 12, 1600170 (2017).

20. Sieben, M. *et al.* Testing plasmid stability of Escherichia coli using the continuously operated shaken BIOreactor system. *Biotechnol. Prog.* 32, 1418–1425 (2016).

21. Doig, S. D. *et al.* Reactor operation and scale-up of whole cell Baeyer-Villiger catalyzed lactone synthesis. *Biotechnol. Prog.* 18, 1039–1046 (2002).

22. Milker, S. *et al.* Kinetic modeling of an enzymatic redox cascade in vivo reveals cofactor-caused bottlenecks. *ChemCatChem* 9, 3420–3427 (2017).

23. Baldwin, C. V. F. & Woodley, J. M. On oxygen limitation in a whole cell biocatalytic Baeyer–Villiger oxidation process. *Biotechnol. Bioeng.* 95, 362–369 (2006).

24. Staudt, S., Bornscheuer, U. T., Menyes, U., Hummel, W. & Gröger, H. Direct biocatalytic one-pot-transformation of cyclohexanol with molecular oxygen into ε-caprolactone. *Enzyme Microb. Technol.* 53, 288–292 (2013).

25. Parra, L. P., Acevedo, J. P. & Reetz, M. T. Directed evolution of phenylacetone monooxygenase as an active catalyst for the Baeyer-Villiger conversion of cyclohexanone to caprolactone. *Biotechnol. Bioeng.* 112, 1354–64 (2015).

26. Mthethwa, K. S. *et al.* Fungal BVMOs as alternatives to cyclohexanone monooxygenase. *Enzyme Microb. Technol.* 106, 11–17 (2017).

27. Van Berkel, W. J. H., Kamerbeek, N. M. & Fraaije, M. W. Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts. *J. Biotechnol.* 124, 670–689 (2006).

28. Romero, E., Gomez Castellanos, J. R., Gadda, G., Fraaije, M. W. & Mattevi, A. Same Substrate, Many Reactions: Oxygen Activation in Flavoenzymes. *Chem. Rev.* (2017) doi:10.1021/acs.chemrev.7b00650.

29. Donoghue, N. A., Norris, D. B. & Trudgill, P. W. The purification and properties of cyclohexanone oxygenase from Nocardia globerula CL1 and Acinetobacter NCIB 9871. *Eur. J. Biochem.* 63, 175–192 (1976).

30. Secundo, F. *et al.* Effects of Water Miscible Organic Solvents on the Activity and Conformation of the Baeyer–Villiger Monooxygenases From Thermobifida fusca and Acinetobacter calcoaceticus: A Comparative Study. *Biotechnol. Bioeng* 108, 491–499 (2011).

31. Romero, E., Gómez Castellanos, J. R., Mattevi, A. & Fraaije, M. W. Characterization and Crystal Structure of a Robust Cyclohexanone Monooxygenase. *Angew.Chem.Int.* 55, 15852–15855 (2016).

32. Mourelle-Insua, Á., Aalbers, F. S., Lavandera, I., Gotor-Fernández, V. & Fraaije, M. W. What to sacrifice? Fusions of cofactor regenerating enzymes with Baeyer-Villiger monooxygenases and alcohol dehydrogenases for self-sufficient redox biocatalysis. *Tetrahedron* 75, 1832–1839 (2019).

33. Torres Pazmiño, D. E. *et al.* Efficient biooxidations catalyzed by a new generation of self-sufficient Baeyer–Villiger monooxygenases. *ChemBioChem* 10, 2595–2598 (2009).

34. Schmidt, S. *et al.* An Enzyme Cascade Synthesis of ε-Caprolactone and its Oligomers. *Angew. Chemie Int. Ed.* 54, 2784–2787 (2015).

35. Sattler, J. H. *et al.* Introducing an in situ capping strategy in systems biocatalysis to access 6-aminohexanoic acid. *Angew. Chemie Int. Ed.* 53, 14153–14157 (2014).

36. Wheeldon, I. *et al.* Substrate channelling as an approach to cascade reactions. *Nat. Chem.* 8, 299 (2016).

37. Rabe, K. S., Müller, J., Skoupi, M. & Niemeyer, C. M. Cascades in compartments: en route to machine-assisted biotechnology. *Angew. Chemie Int. Ed.* 56, 13574–13589 (2017).

38. Kuzmak, A., Carmali, S., von Lieres, E., Russell, A. J. & Kondrat, S. Can enzyme proximity accelerate cascade reactions? *Sci. Rep.* 9, 1–7 (2019).

39. Ellis, G. A. *et al.* Artificial multienzyme scaffolds: Pursuing in vitro substrate channeling with an overview of current progress. *ACS Catal.* 9, 10812–10869 (2019).

40. Aalbers, F. S. & Fraaije, M. W. Enzyme fusions in biocatalysis: coupling reactions by pairing enzymes. *ChemBioChem* 20, 20–28 (2019).

41. Kohl, A., Srinivasamurthy, V., Böttcher, D., Kabisch, J. & Bornscheuer, U. T. Co-expression of an alcohol dehydrogenase and a cyclohexanone monooxygenase for cascade reactions facilitates the regeneration of the NADPH cofactor. *Enzyme Microb. Technol.* 108, 53–58 (2018).

42. Scherkus, C. *et al.* Kinetic insights into ε-caprolactone synthesis: Improvement of an enzymatic cascade reaction. *Biotechnol. Bioeng.* 114, 1215–1221 (2017).

43. Aalbers, F. S. & Fraaije, M. W. Coupled reactions by coupled enzymes: alcohol to lactone cascade with alcohol dehydrogenase–cyclohexanone monooxygenase fusions. *Appl. Microbiol. Biotechnol.* 101, 7557–7565 (2017).

44. Jeon, E. Y., Baek, A. H., Bornscheuer, U. T. & Park, J. B. Enzyme fusion for whole-cell biotransformation of long-chain sec-alcohols into esters. *Appl. Microbiol. Biotechnol.* 99, 6267–6275 (2015).

45. Argos, P. An investigation of oligopeptides linking domains in protein tertiary structures and possible candidates for general gene fusion. *J. Mol. Biol.* 211, 943–958 (1990).

46. Chen, X., Zaro, J. L. & Shen, W.-C. Fusion protein linkers: property, design and functionality. *Adv. Drug Deliv. Rev.* 65, 1357–1369 (2013).

47. van Rosmalen, M., Krom, M. & Merkx, M. Tuning the flexibility of glycine-serine linkers to allow rational design of multidomain proteins. *Biochemistry* 56, 6565–6574 (2017).

48. Tischler, D. *et al.* VpStyA1/VpStyA2B of Variovorax paradoxus EPS: An aryl alkyl sulfoxidase rather than a styrene epoxidizing monooxygenase. *Molecules* 23, 809 (2018).

49. Corrado, M. L., Knaus, T. & Mutti, F. G. A chimeric styrene monooxygenase with increased efficiency in asymmetric biocatalytic epoxidation. *ChemBioChem* 19, 679–686 (2018).

50. Heine, T. *et al.* Engineering styrene monooxygenase for biocatalysis: Reductase-epoxidase fusion proteins. *Appl. Biochem. Biotechnol.* 181, 1590–1610 (2017).

51. Belsare, K. D. *et al.* P-LinK: A method for generating multicomponent cytochrome P450 fusions with variable linker length. *Biotechniques* 57, 13–20 (2014).

52. Gräwe, A., Ranglack, J., Weyrich, A. & Stein, V. iFLinkC: an iterative functional linker cloning strategy for the combinatorial assembly and recombination of linker peptides with functional domains. *Nucleic Acids Res.* 48, e24–e24 (2020).

53. Qiagen. *The QIAexpressionist. A handbook for high-level expression and purification of 6xHis-tagged proteins", S* vol. 5 (2003).

54. Spriestersbach, A., Kubicek, J., Schaefer, F., Block, H. & Maertens, B. Purification of His-tagged proteins. in *Methods in enzymology* vol. 559 1–15 (Elsevier, 2015).

55. Chapman, S. K. & Reid, G. A. *Flavoprotein protocols*. vol. 131 (Springer Science & Business Media, 1999).

56. Massey, V. *The Chemical and Biological Versatility of Riboflavin*. (2000).

57. Macheroux, P. UV-visible spectroscopy as a tool to study flavoproteins. *Methods Mol. Biol.* 131, 1–7 (1999).

58. Romero, E., Castellanos, J. R. G., Mattevi, A. & Fraaije, M. W. Characterization and Crystal Structure of a Robust Cyclohexanone Monooxygenase. *Angew. Chemie - Int. Ed.* 55, 15852–15855 (2016).

59. Sanchez-Ruiz, J. M. Protein kinetic stability. *Biophys. Chem.* 148, 1–15 (2010).

60. Bommarius, A. S. & Paye, M. F. Stabilizing biocatalysts. *Chem. Soc. Rev.* 42, 6534–6565 (2013).

61. Polizzi, K. M., Bommarius, A. S., Broering, J. M. & Chaparro-Riggers, J. F. Stability of biocatalysts. *Curr. Opin. Chem. Biol.* 11, 220–225 (2007).

62. Hughes, A. V, Rees, P., Heathcote, P. & Jones, M. R. Kinetic analysis of the thermal stability of the photosynthetic reaction center from Rhodobacter sphaeroides. *Biophys. J.* 90, 4155–4166 (2006).

63. Faham, S. *et al.* Side-chain contributions to membrane protein structure and stability. *J. Mol. Biol.* 335, 297–305 (2004).

64. Lau, F. W. & Bowie, J. U. A method for assessing the stability of a membrane protein. *Biochemistry* 36, 5884–5892 (1997).

65. Galka, J. J., Baturin, S. J., Manley, D. M., Kehler, A. J. & O'neil, J. D. Stability of the glycerol facilitator in detergent solutions. *Biochemistry* 47, 3513–3524 (2008).

66. Grinberg, A. V, Gevondyan, N. M., Grinberg, N. V & Grinberg, V. Y. The thermal unfolding and domain structure of Na$^+$/K$^+$-exchanging ATPase. A scanning calorimetry study. *Eur. J. Biochem.* 268, 5027–5036 (2001).

67. Mancusso, R., Karpowich, N. K., Czyzewski, B. K. & Wang, D.-N. Simple screening method for improving membrane protein thermostability. *Methods* 55, 324–329 (2011).

68. Cummings, M. D., Farnum, M. A. & Nelen, M. I. Universal screening methods and applications of ThermoFluor®. *J. Biomol. Screen.* 11, 854–863 (2006).

69. Pantoliano, M. W. *et al.* High-density miniaturized thermal shift assays as a general strategy for drug discovery. *J. Biomol. Screen.* 6, 429–440 (2001).

70. Mezzasalma, T. M. *et al.* Enhancing recombinant protein quality and yield by protein stability profiling. *J. Biomol. Screen.* 12, 418–428 (2007).

71. Geerlof, A. *et al.* The impact of protein characterization in structural proteomics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 62, 1125–1136 (2006).

72. Forneris, F., Orru, R., Bonivento, D., Chiarelli, L. R. & Mattevi, A. ThermoFAD, a Thermofluor-adapted flavin ad hoc detection system for protein folding and ligand binding. *FEBS J.* 276, 2833–2840 (2009).

73. Bogin, O., Peretz, M. & Burstein, Y. Thermoanaerobacter brockii alcohol dehydrogenase: characterization of the active site metal and its ligand amino acids. *Protein Sci.* 6, 450–458 (1997).

74. Yip, L. Y., Chun, E. & Chan, Y. Proteomic and metabolomic approaches to biomarker discovery. in *Proteomic and Metabolomic Approaches to Biomarker Discovery* Chapter 8 (Academic Press, 2013).

75. Bothner, B. *et al.* Monitoring enzyme catalysis with mass spectrometry. *J. Biol. Chem.* 275, 13455–13459 (2000).

76. Mellon, F. A. Encyclopedia of food sciences and nutrition. in *Encyclopedia of food sciences and nutrition* 1294–1301 (Academic, 2003).

77. Lindon, J. C., Tranter, G. E. & Koppenaal, D. *Encyclopedia of spectroscopy and spectrometry*. (Academic Press, 2016).

78. Milker, S., Goncalves, L. C. P., Fink, M. J. & Rudroff, F. *Escherichia coli* fails to efficiently maintain the activity of an important flavin monooxygenase in recombinant overexpression. *Front. Microbiol.* 8, 1–9 (2017).
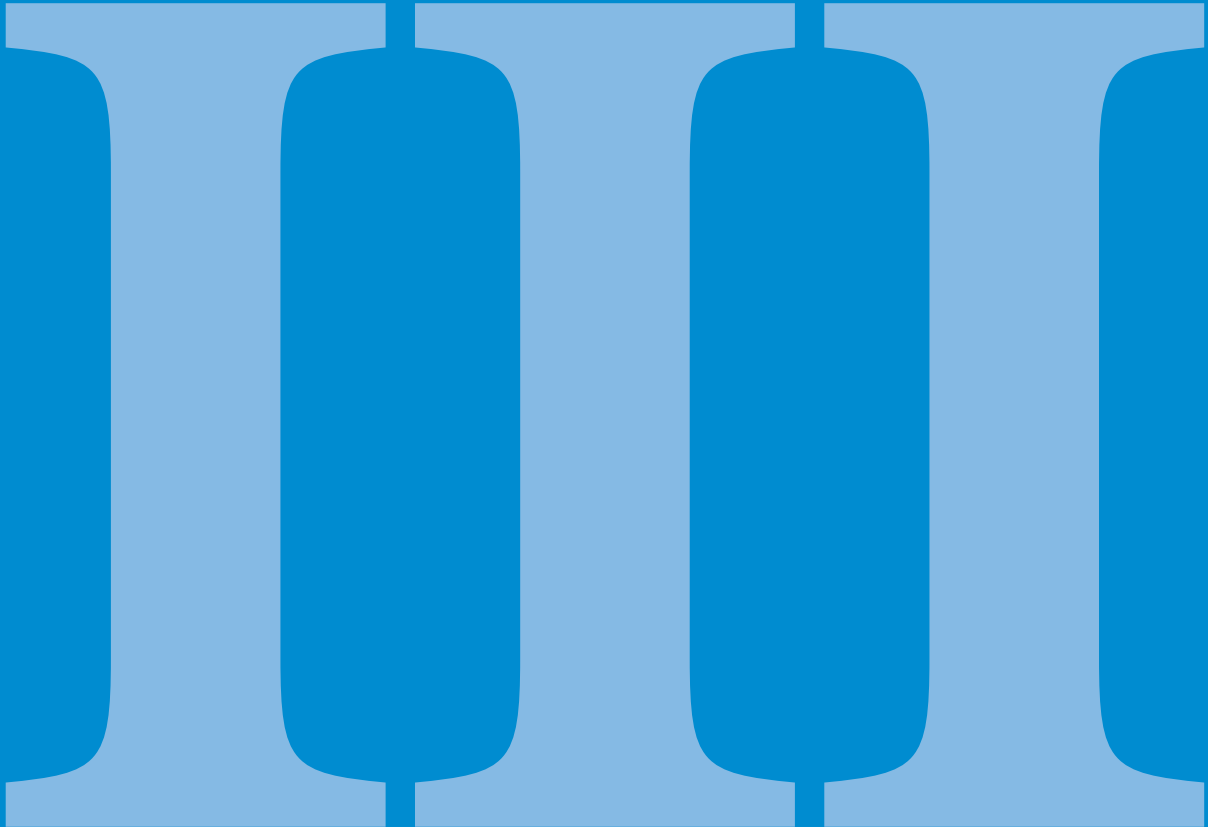
# Systematic assessment of uncoupling in flavoprotein oxidases and monooxygenases

**Alejandro Gran-Scheuch**, Loreto Parra and Marco W. Fraaije

# ABSTRACT

Flavin as a cofactor is an extremely versatile molecule that participates in a wide range of biochemical reactions. A special characteristic of the flavin cofactor, unique for a metal-free cofactor, is its ability to react in its reduced form with molecular oxygen. This is exploited in flavoprotein oxidases and monooxygenases. The flavin-mediated reduction of dioxygen can lead to various reactive oxygen species (ROS), primarily hydrogen peroxide and superoxide. No systematic analysis of the formation of such reduced oxygen species produced by flavoprotein oxidases and monooxygenases had been performed before. In this work, we investigated which ROS are formed by several prototype flavoprotein oxidases and monooxygenases: phenylacetone monooxygenase ($PAMO_{WT}$), an engineered PAMO variant that acts as an NADPH oxidase ($PAMO_{C65D}$), eugenol oxidase (EUGO), and 5-hydroxymethyl furfural oxidase (HMFO). The formed amounts of superoxide and hydrogen peroxide were measured under various conditions (different substrate concentrations, pH values and cosolvents). Interestingly, all flavoenzymes were found to produce, except for hydrogen peroxide, significant amounts of superoxide. Moreover, increased superoxide levels were measured at higher pH, which could be indicative for a pH-sensitive caged radical pair dissociation. To probe the effect of ROS formation on biocatalytic performance, conversions catalyzed by $PAMO_{WT}$ or EUGO, with or without catalase, were monitored. This revealed that catalase has a beneficial effect. No detrimental effect of the accumulation of superoxide on biocatalysis could be demonstrated. The results reveal that formation of ROS by flavoenzymes is highly dependent on the experimental conditions used. The results provide a better insight into the mechanism by which ROS is formed in flavoenzymes and may help studies or applications in which ROS formation should be promoted or minimized.

**Keywords:** Flavoprotein monooxygenase, flavoprotein oxidase, flavin, reactive oxygen species (ROS), superoxide, hydrogen peroxide

# INTRODUCTION

Flavin cofactors are tremendously versatile catalysts in nature. They enable a wide variety of enzyme-catalyzed (electro)chemical reactions[1–4]. This adaptability is partly explained by the ability of a flavin to accept or donate one or two electrons[1]. Flavoenzymes can exploit three different redox states: oxidized as flavoquinone ($Fl_{ox}$), one-electron-reduced as flavosemiquinone ($Fl_{sq}$), or the two-electron-reduced species as flavohydroquinone ($Fl_{red}$). On top of that, several other flavin-based cofactors have been encountered, such as prenylated flavins, flavin-N5-oxygen adducts, and deazaflavins[5–7]. When considering the canonical flavin cofactors FMN and FAD, it is known that they can exist in several protonation states, depending on the redox state and pH. The neutral and anionic states are the most relevant when considering flavoenzymes. Reduced flavoenzymes can utilize various oxidants. A special case is the high reactivity of a select set of flavoenzymes with molecular oxygen, a unique feature of redox enzymes employing an organic cofactor. In particular, it is known that free $Fl_{red}$ reacts slowly with dioxygen with a reaction rate of about 250 $M^{-1}s^{-1}$[8,9]. The resilience to react with dioxygen comes from the fact that in this reaction a spin-inversion has to occur[10]. Dioxygen is a diradical and paramagnetic molecule in the triplet ground state ($^3O_2$), with two unpaired electrons, which have parallel spins[10,11]. On the other hand, organic molecules —such as the N5-C(4α) locus of $Fl_{red}$— are usually in singlet state and exhibit paired electrons with opposite spins[10]. Consequently, for triplet molecular oxygen, the two-electron transfer from organic molecules is chemically spin-forbidden and requires a high activation energy[12]. At the same time, the redox potential of the $O_2/H_2O_2$ couple ($E_m$ = +270 mV) versus the free $Fl_{ox}/Fl_{red}$ couple ($E_m$ = -207 mV) makes the reaction basically irreversible[13]. The unfavorable thermodynamical reactivity prevents spontaneous dioxygen-mediated oxidation of organic molecules. It also allowed organisms to use dioxygen for respiration, helped by oxygen-reactive cofactors such as iron, copper, heme and flavin cofactors. For a select group of flavin-dependent enzymes, the monooxygenases and oxidases, the reoxidation of their reduced flavin occurs very rapidly with molecular oxygen with rates of $10^4$ - $10^6$ $M^{-1}s^{-1}$[14]. Clearly, the protein microenvironment around the flavin cofactor can accelerate (or slow down) flavin reactivity for $O_2$ by orders of magnitude[8,15,16].

It has been well established that variations of the residues in the flavin-binding pockets can perturb the physicochemical properties of the cofactor. The flavin microenvironment affects amongst others $pK_a$ values and redox potentials of the two-electron transfer reaction, while it can also have steric effects [13,16]. These effects explain the enormous differences of the oxygen reactivity among flavoenzymes. For example, the redox potentials for flavoenzymes can vary from -400 to +150 mV[4,16,17]. Furthermore, a redox potential of the $Fl_{sq}/Fl_{red}$ couple below than of $O_2/O_2^{\cdot-}$ would suggest a stabilization for the $Fl_{sq}$ causing favorable conditions for $O_2$ reactivity[16]. The redox potential for dioxygen is considerably

lower in hydrophobic solvents, hence desolvation of the active site may reduce the redox potential. This can decrease the reaction rate with $O_2$ but cannot be generalized for all flavoenzymes[16,18]. For example, the protein part of a flavoenzyme can also shield the flavin from dioxygen and thereby prohibit the reduced flavin to use dioxygen as electron acceptor[19]. Hence, the differences in the $O_2$ reactivity within flavoenzymes are due to a combination of factors which are modulated by the protein environment.

Broadly, flavoenzymes have been classified in three well-defined groups: flavin-dependent reductases/dehydrogenases, oxidases and monooxygenases. Only the two latter enzymes react efficiently with molecular oxygen. Flavin-dependent dehydrogenases react slowly or not at all with dioxygen and instead utilize other electron acceptors such as quinones or cytochromes[8,20]. For flavin-dependent oxidases and monooxygenases, it is generally assumed that the reaction between $Fl_{red}$ and dioxygen starts with a one-electron transfer from the singlet flavin at the N5-C4α locus to the triplet dioxygen (**Scheme 1**). Then, a short-lived caged radical pair between the neutral $Fl_{sq}$ and superoxide radical anion is formed. This first one-electron reduction overcomes the spin-barrier. The subsequent steps diverge dependent on the type of flavoenzyme. For oxidases, in most cases the reaction continues by a rapid second electron transfer from the $Fl_{sq}$ to the superoxide anion, generating the $Fl_{ox}$ species and hydrogen peroxide[20]. Consequently, molecular oxygen acts as an electron acceptor. For flavoprotein monooxygenases, the caged radical pair collapses forming a nucleophilic quasi-stable C(4α)-peroxyflavin adduct which can also be protonated to the electrophilic hydroperoxyflavin[21]. From there, $Fl_{ox}$ can be recovered by oxygen insertion into an organic substrate while reducing the other oxygen atom to water. In case a flavoprotein monooxygenase is not capable to react with an organic substrate, the peroxyflavin may decay to form hydrogen peroxide, essentially functioning as a NAD(P)H oxidase. The latter reaction is well known as uncoupling and it can be regarded as a futile use of the electrons. The third and last alternative, is the dissociation of the caged pair, forming flavin semiquinone radical and superoxide anion. This route is followed by a re-oxidation of the $Fl_{sq}$ to $Fl_{ox}$ by a second molecule of oxygen[15]. It is worth noting that, while the above-mentioned mechanisms of oxygen reactivity are valid for most studied flavoenzymes, it was recently discovered that a selected number of oxygen-consuming flavoenzymes operate via another mechanism. These oxidative flavoenzymes are capable to form a flavin-N5-peroxide and are not considered in this study[6].
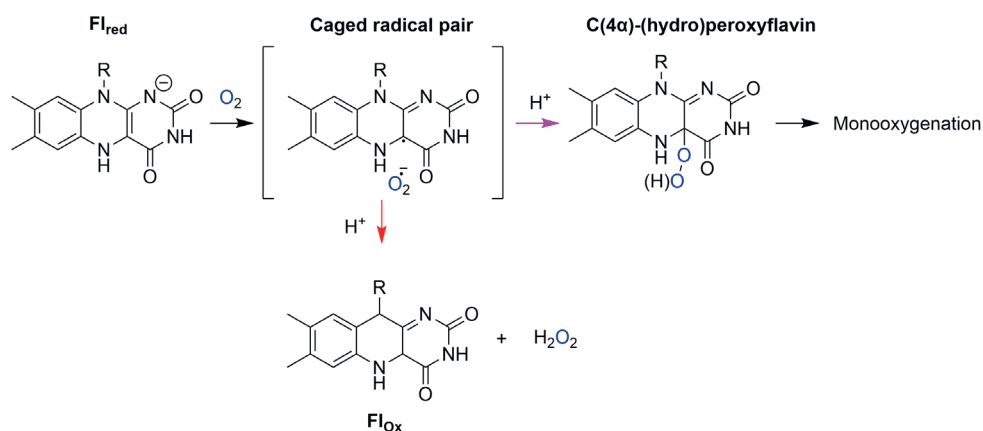
**Scheme 1. Reaction of reduced flavin with oxygen.** The reaction starts with transfer of one electron from the reduced flavin ($Fl_{red}$) to dioxygen, to form a caged radical pair between the flavin semiquinone and the superoxide anion. For oxidases, a second electron transfer generates $Fl_{Ox}$ and hydrogen peroxide (red arrow). For monooxygenases and certain oxidases, the radical pair collapses to form a C4α-(hydro)peroxyflavin species, which can further perform monooxygenation or hydrogen peroxide elimination (pink arrow). Scheme modified from Chaiyen et al. [20]

For flavoenzymes, a well-tuned dioxygen reactivity is crucial for efficient catalysis and preventing formation of harmful ROS side products. Therefore, understanding the features that govern the collapse or dissociation of the superoxide-flavin semiquinone caged radical pair is highly relevant. Additionally, for monooxygenases it is interesting to understand which conditions affect C4α-peroxyflavin intermediate stabilization and thus avoid a futile oxidation of hydride donors (NADH or NADPH) and/or ROS formation. Related to this, there are some previous studies that looked into the production of ROS formation by flavoproteins[22–24]. Yet, most studies on flavoprotein monooxygenase or oxidases neglect the potential issues that may be caused by uncoupling and ignore possible superoxide formation. For example, often activity of flavoprotein monooxygenases is measured (and reported) by merely monitoring NAD(P)H consumption. But such measurements cannot discriminate between oxygenation and uncoupling activity; it is the sum of both. Furthermore, typically, when uncoupling by monooxygenases is studied or oxidase activity is described, the only reduced oxygen species considered is hydrogen peroxide. As a comprehensive analysis of ROS formation by flavoprotein oxidases and monooxygenases was lacking, we investigated the formed reduced oxygen species generated by flavoprotein oxidases and monooxygenases. A better understanding of the conditions that prevent or promote ROS formation by flavoenzymes is highly relevant for many research areas, such as biotechnology (e.g. improve efficiency of biocatalysts) and biomedicine (e.g. trigger ROS formation).

## RESULTS AND DISCUSSIONS

In this work, we evaluated the production of hydrogen peroxide and superoxide anion by four well-characterized flavoenzymes. These are: i) the thermostable NADPH-dependent Baeyer-Villiger monooxygenase, phenylacetone monooxygenase ($PAMO_{WT}$) which acts on various substrates and has been extensively studied as biocatalyst[25], ii) the C65D PAMO mutant which acts as a NADPH oxidase as it shows an extremely high rate of uncoupling compared with the wild type enzyme[26], iii) eugenol oxidase (EUGO), a VAO-type oxidase containing a covalently bound FAD and which oxidizes a range of phenolic compounds[27], and iv) 5-hydroxymethylfurfural oxidase (HMFO), a GMC-type FAD-containing oxidase which can be used for the synthesis of the biobased polymer precursor, furan-2,5-dicarboxylic acid[28]. The production profile of hydrogen peroxide and superoxide of these enzymes was examined at different pH values, substrate concentrations and in the presence of cosolvents.

## PURIFICATION AND ASSAYS STANDARDIZATION

For assessing ROS generation by $PAMO_{WT}$, $PAMO_{C65D}$, EUGO and HMFO, the enzymes had to be obtained as pure and fully oxidized flavoenzymes. It is essential to use the enzymes in their oxidized form to eliminate background ROS formation through the contribution of semiquinone or fully reduced flavoprotein species. It is well known that the different redox states of the flavin cofactor can be distinguished based on the UV-Vis absorption features[29]. All enzymes were obtained as His-tagged proteins using the same host, *Escherichia coli* NEB 10β, and a pBAD expression vector. The flavoenzymes were purified by affinity chromatography resulting in yellow-colored samples. Only for the PAMO preparations, some absorbance features indicative for some flavin semiquinone species were observed. These samples were incubated with potassium ferricyanide —an oxidizing agent— and an excess of free FAD. After desalting, the UV-Vis spectra were recorded which revealed that fully oxidized samples of all four flavoproteins were obtained (**Figure S1a**).

While purity judged from SDS-PAGE was very high, significant interfering ROS scavenging activities were measured, resulting from minor catalase and superoxide dismutase (SOD) impurities. An additional affinity chromatography purification step was performed after which fractions were carefully tested for catalase and SOD activities. The fractions devoid of interfering enzyme activity were pooled, concentrated, desalted and used for further experiments. By catalase and SOD activity measurements, it was verified that all four purified enzyme samples did not contain any significant catalase or SOD activity (**Figure S2**). Also, the purity of the four flavoenzymes was again checked by SDS-PAGE (**Figure S1b**).

Lastly, the effect of the tetrazolium salt (WST-1) — the agent used for the detection of superoxide radical anion, see material and methods — was evaluated for its effect on both stability and activity of the studied enzymes. Concerning the thermostability of the studied flavoenzymes, there were no changes of the apparent melting temperatures ($T_M^{app}$) over the controls without WST-1 (**Fig. S3a**). Also, no significant effect of WST-1 on enzyme activity was found for all respective enzymes. Hence, WST-1 seems compatible to use in enzyme assays, by not disturbing enzyme functioning. With these results, we continued with these enzyme samples for further experiments.

## ASSESSMENT OF THE NADPH OXIDASE ACTIVITY FOR PAMO$_{WT}$ AND PAMO$_{C65D}$

Baeyer-Villiger monooxygenases and other flavoprotein monooxygenases are known to exhibit so-called uncoupling activity. This refers to the reaction in which the formed peroxyflavin intermediate does not react with the target substrate, but decays into oxidized flavin while forming hydrogen peroxide and/or superoxide. In a previous study, we have created a variant of PAMO that is crippled in stabilizing the peroxyflavin and therefore promotes the uncoupling reaction, essentially acting as an NADPH oxidase. The nature of the ROS formed by flavoprotein monooxygenases and the effects of conditions on the uncoupling reaction has hardly been studied before. In this study, the production of ROS (superoxide and hydrogen peroxide) by PAMO$_{WT}$ and its mutant PAMO$_{C65D}$ was carefully analyzed. First, we explored the uncoupling activity with NADPH as coenzyme in the absence of any substrate at pH 8.0. The levels of NADPH, hydrogen peroxide and superoxide were spectrophotometrically monitored for 30 min in the presence of 50, 100 or 200 µM of NADPH. As expected, the consumption of NADPH by PAMO$_{C65D}$ occurs faster than PAMO$_{WT}$, due to the fact that the uncoupling rates ($k_{unc}$) for the mutant is about 2 orders of magnitude higher than for PAMO$_{WT}$[26] (**Table S1**). Interestingly, for both enzymes, both the formation of superoxide and hydrogen peroxide was observed. The data reveal that, for both enzymes, the rate of formation of superoxide is 1-2 order of magnitude lower when compared with the formation of hydrogen peroxide. The amounts of hydrogen peroxide and superoxide formed after 30 minutes were determined. Interestingly, PAMO$_{WT}$ exhibited a significantly higher production of superoxide than the uncoupled mutant PAMO$_{C65D}$ (**Figure 1a**). PAMO$_{WT}$ produced about 5 times more superoxide than PAMO$_{C65D}$, generating up to 9 µM of superoxide. The mutant generated at most 2 µM superoxide (**Figure 1b**). This is in line with the observed kinetics: the initial rate of superoxide formation is relatively low when compared with the rate of hydrogen peroxide formation. These data indicate that the partitioning of superoxide / hydrogen peroxide production is favored towards hydrogen peroxide in both monooxygenases. While PAMO$_{WT}$ generates about 10-fold more hydrogen peroxide compared with superoxide, for

PAMO$_{C65D}$ the amount of superoxide is 100-fold lower compared with hydrogen peroxide. Intriguingly, the data show that both enzymes form significant amounts of superoxide. So far, the potential of flavoprotein monooxygenases to produce superoxide was unknown.
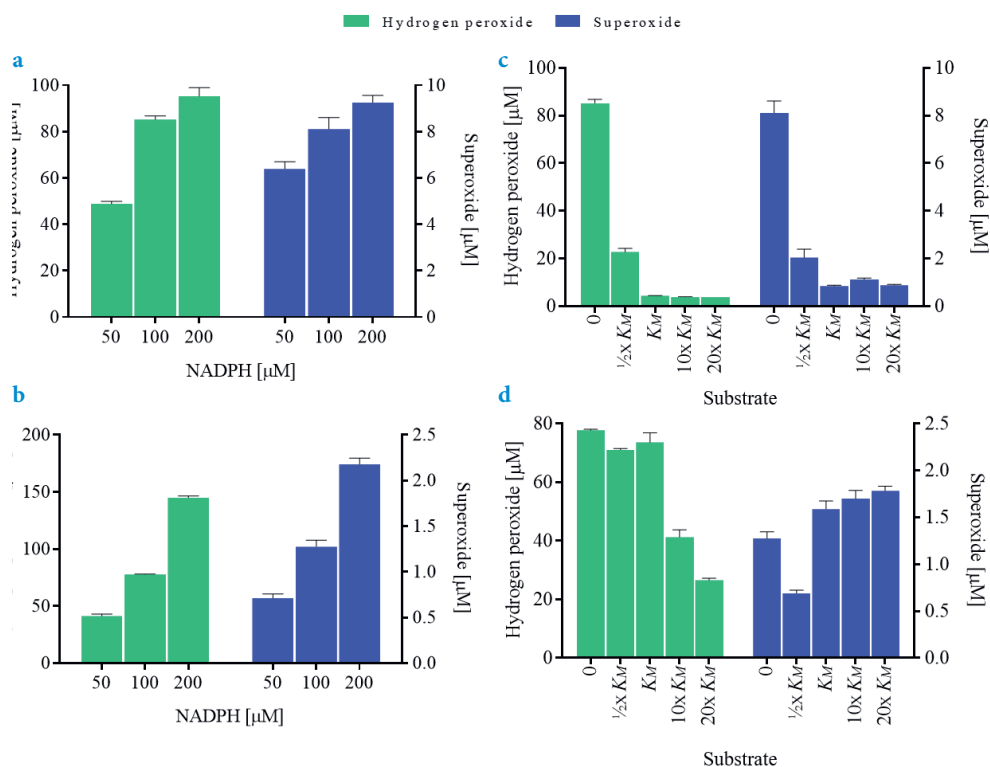


**Figure 1. Substrate and co-substrate effect over ROS production by PAMO$_{WT}$ and PAMO$_{C65D}$.** The total production of hydrogen peroxide (green) and superoxide anion (blue) was measured at different NADPH concentrations for (a) PAMO$_{WT}$ and (b) PAMO$_{C65D}$. Increasing amounts of phenylacetone were evaluated in the presence of 100 μM NADPH for both (c) PAMO$_{WT}$ and (d) PAMO$_{C65D}$. Reactions were carried out in 50 mM TRIS buffer pH 8.0, with 150 mM NaCl. Reactions were performed for 30 min at 25 °C in constant double-orbital shaking in triplicate.

Next, the analysis was done in the presence of different concentrations of phenylacetone to investigate the effect of the substrate. In the case of PAMO$_{WT}$, one would expect a drastic decrease in ROS formation due to the oxygenation of the substrate. We evaluated substrate concentrations equivalent to ½x $K_M$, $K_M$, 10x $K_M$ or 20x the $K_M$ value (equals to 25, 50, 500 and 1,000 μM) while using 100 μM NADPH[25]. Indeed, by increasing the substrate concentration, a drastic decrease in the production of hydrogen peroxide by PAMO$_{WT}$ was observed. At saturating amounts of phenylacetone, the production of hydrogen peroxide remained below 5 μM (**Figure 1c**). Also, superoxide formation decreased at high substrate concentrations to about 1 μM. This also translated in a relatively low uncoupling ratio of <5 % at saturating substrate concentrations (**Table S2**). This can be explained by

the fact that $PAMO_{WT}$ can form and stabilize a C(4α)-peroxyflavin intermediate which, after substrate binding, can perform a nucleophilic attack on the electron-poor carbonyl of phenylacetone. This cocked-gun mechanism is valid for all studied class B flavin-dependent monooxygenases and enables them to perform a rapid turnover when a proper substrate is present[30]. It ensures that NADPH oxidation is mainly coupled to substrate oxygenation, thereby preventing a waste of electrons. For $PAMO_{C65D}$, the decrease in hydrogen peroxide formation was less sharp when using phenylacetone. Also, the levels of formed superoxide were relatively constant (**Figure 1d**). In line with this, $PAMO_{C65D}$ exhibited high uncoupling (30 %-90 %), even at relatively high substrate concentrations (**Table S2**). This agrees with the fact that $PAMO_{C65D}$ is not effective in stabilizing the C(4α)-peroxyflavin intermediate[26]. This is due a reorientation of Asp66 toward the N5-C(4α) locus upon replacing Cys65 by an aspartate.

Subsequently, the effect of pH on ROS production was evaluated in the presence of 1.0 mM phenylacetone and 100 µM NADPH. Experiments were performed for 30 min in 50 mM KPi (pH 6.0), 50 mM TRIS (pH 7.0-9.0) or 50 mM CHES buffer (pH 9.0-10.0). For $PAMO_{WT}$, the uncoupling at different pHs ranged from 5.0 to 7.8 % (**Table S3**). Formation of hydrogen peroxide was rather constant (around 4 µM), while formation of superoxide was almost absent at low pH (pH 6.0), reaching 3 µM at high pH (**Figure 2a**). For $PAMO_{C65D}$, hydrogen peroxide and superoxide levels were highest at pH 9.0. Again, the amount of formed superoxide was dependent on pH, which was negligible at pH 6.0 (**Figure 2b**).
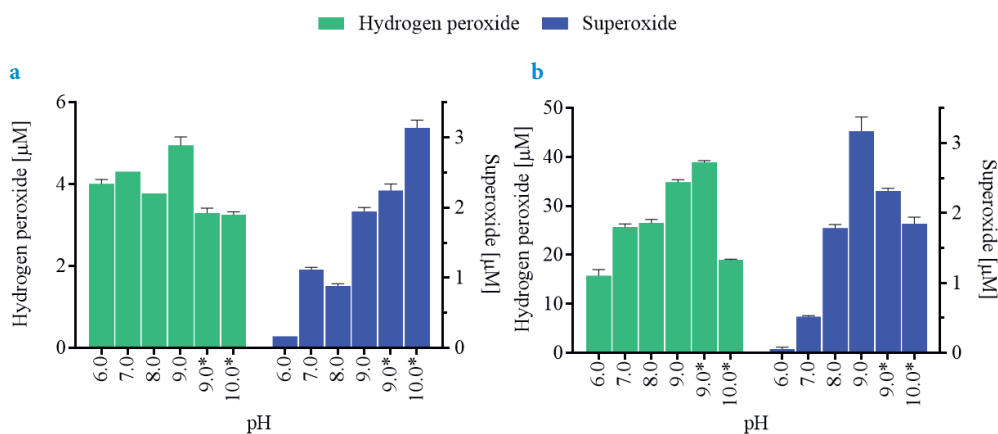


**Figure 2. Evaluation of generated ROS by $PAMO_{WT}$ at different pH values.** Measurements of total hydrogen peroxide (green) and superoxide anion (blue) were determined at increasing pH from 6.0 to 10.0 for (a) $PAMO_{WT}$ and (b) $PAMO_{C65D}$. Reactions were performed in in 50 mM KPi (pH 6.0), 50 mM TRIS (pH 7.0-9.0) or 50 mM CHES (9.0 and 10.0). All buffers contained 150 mM NaCl. Reactions were incubated for 30 min at 25 °C in constant double-orbital shaking. Reactions were performed in triplicate.

Interestingly, for PAMO$_{WT}$ and PAMO$_{C65D}$, the superoxide production was relatively high at high pH, generating up 2-3 μM superoxide at pH 10.0. In fact, for PAMO$_{WT}$, the amount of superoxide almost equaled the amount of hydrogen peroxide formed at pH 10.0 (**Figure S4c**). In line with the relatively high superoxide levels formed at high pH for both enzymes, the initial rates of superoxide formation of superoxide were also high at high pH (**Table S1**). At pH 6.0, these rates were extremely low. The observation that formation of superoxide is promoted at high pH suggests that dissociation of the caged radical pair is pH-sensitive.

Lastly, the total ROS production by PAMO in the presence of different concentrations of miscible cosolvents was evaluated. Reactions were performed in the presence of 1.0 mM phenylacetone at pH 8.0. Cosolvent concentrations up to 7.5 % v/v had little effect on NADPH depletion (**Figure S5a,c,e,g**). In fact, it has been reported that PAMO$_{WT}$ is quite tolerant towards solvents[31–33]. In general, the presence of cosolvent did not significantly affect ROS production. Similar to conversions in the absence of cosolvent, most uncoupling resulted in hydrogen peroxide while minor amounts of superoxide were detected. The uncoupling ratio of PAMO$_{WT}$ for all tested conditions was on average 4 % (**Table S3**). For PAMO$_{C65D}$, there was a trend to less hydrogen peroxide formation at higher amounts of cosolvent (**Figure S5b,d,f,h**). For methanol, ethanol, acetonitrile and isopropanol the uncoupling ratios were reduced from 36 % in the control to 17 %, 14 %, 24 % and 18 %, respectively (**Table S3**). Whilst, for both enzyme in all four cosolvents, the production of superoxide was low and relatively constant.

## ANALYSIS OF ROS PRODUCTION BY FLAVOPROTEIN OXIDASES

A similar analysis of ROS formation by two flavoprotein oxidases, EUGO and HMFO, was performed. Unlike monooxygenases, these flavoenzymes do not require an external hydride donor. In flavoprotein oxidases, the substrate reduces the flavin and then dioxygen acts as electron acceptor, typically generating hydrogen peroxide. In fact, activity of flavoprotein oxidases, such as glucose oxidase, is usually assayed using a peroxidase-coupled assay that relies on the formed hydrogen peroxide. Formation of superoxide by flavoprotein oxidases is normally not considered. First, we tested ROS formation using different substrate concentrations at pH 8.0. ROS production by EUGO and HMFO was analyzed in the presence of vanillyl alcohol and HMF at ½x $K_M$, $K_M$, 10x $K_M$ or 20x $K_M$ (corresponding to 25, 50, 500 and 1,000 μM vanillyl alcohol for EUGO and 0.75, 1.5, 15 and 30 mM HMF for HMFO) (**Figure S6**). This revealed that, similar to the monooxygenases, both oxidases generate hydrogen peroxide and superoxide as reduced dioxygen species. In both cases, the amounts of formed superoxide were relatively small. Even at high substrate concentrations, resulting in the most ROS, the levels of formed superoxide were rather low

(< 1.0 μM). The amounts of superoxide formed were detectable but in the range of 0.2-0.5 % of the amount of hydrogen peroxide formed.

Then, ROS levels were measured at different pH values, ranging from pH 6.0 to 10.0 (**Figure 3**). Reactions were carried out in the presence of substrate equivalent to 20x $K_M$. When comparing both oxidases, the formed hydrogen peroxide levels showed opposite trends. This is in line with their respective pH optima for activities. HMFO displays highest activity below pH 9.0, which is reflected in relatively little hydrogen peroxide formation at higher pH values (**Figure 3a**)[23]. While, for EUGO a significant rise of hydrogen peroxide production was detected by increasing the pH (**Figure 3b**). This is in accordance with the pH optimum for activity for EUGO which is around pH 9.0[27,28,34]. Interestingly, for the formation of superoxide, a clear and similar pH-dependent trend is seen for both oxidases. Similar to the PAMO variants described above, the highest superoxide anion generation was obtained at high pH values (pH 9.0 and 10.0). For HMFO, despite a drastic decrease in hydrogen peroxide production, up to 3.0 μM superoxide was formed at pH 10.0 (**Figure 3a**). Also for EUGO, significant higher levels of superoxide were observed at high pH. At pH 9.0, around 20 μM superoxide was formed (**Figure 3b**). Particularly, the production of superoxide at pH 9.0 and 10.0 was about 25 times higher than at pH 8.0. When using EUGO at pH 9.0, a significant amount (around 3 %) of the produced reduced oxygen species is superoxide. In line with this, the superoxide production rate for EUGO increased from below 0.001 s$^{-1}$ at pH 8.0 to 0.1 s$^{-1}$ at pH 9.0-10 (**Table S4**).
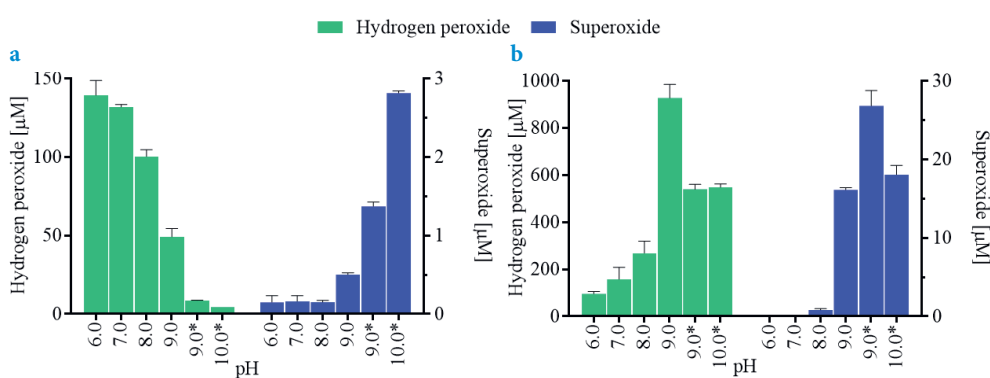


**Figure 3. Effect of pH in the production of hydrogen peroxide and superoxide anion by oxidases.** Determination of total hydrogen peroxide (green) and superoxide anion (blue) were measured at increasing pH from 6.0 to 10.0 for (a) HMFO and (b) EUGO. Reactions were performed in in 50 mM KPi (pH 6.0), 50 mM TRIS (pH 7.0-9.0) or 50 mM CHES* (9.0 and 10.0). All buffers contained 150 mM NaCl. Incubations were carried out 30 min at 25 °C in constant double-orbital shaking. Reactions were performed in triplicate.

Next, the effect of miscible cosolvents at pH 8.0 and 20x $K_M$ of substrate on ROS production was investigated. For this, methanol, ethanol, acetonitrile and isopropanol (at 2.5, 5.0 and 7.5 % v/v) were used. For HMFO the increase of cosolvents levels led to a slight decrease in oxidase activity (**Figure S7**). Particularly, a small decrease of hydrogen peroxide formation was observed while the superoxide levels were the same. Likewise, for EUGO, hydrogen peroxide production decreased upon adding cosolvents. As an exception, with isopropanol no significant effect on hydrogen peroxide production was found (**Figure S8**). On the other hand, the superoxide production at higher amounts of cosolvent tends to decrease, particularly in the presence of acetonitrile and isopropanol.

## EFFECT OF ROS SCAVENGERS ON PHENYLACETONE AND VANILLYL ALCOHOL CONVERSIONS

After evaluating the factors and stressors that could affect the production of hydrogen peroxide or superoxide, their potential detrimental effect on conversions was evaluated. For this, PAMO$_{WT}$ and EUGO were selected as biocatalysts for conversions at pH 9.0. This condition was found to result in relatively high levels of superoxide (vide supra). Conditions were chosen that resulted in partial conversion when standard conditions were used. In the case of PAMO$_{WT}$ (0.50 μM), the use of 50 mM phenylacetone resulted in 22 % conversion in 24 h. In the case of EUGO (0.15 μM), 20 mM of vanillyl alcohol was used as test concentration as it led to only 70 % conversion after 24 h. Then, to investigate the effect of the formed ROS, for both enzymes the reactions were performed with or without an excess of catalase (35 U mL$^{-1}$) and/or SOD (100 U mL$^{-1}$). For PAMO$_{WT}$, the presence of catalase increased the conversion to nearly 35 %, while the addition of SOD had a minor effect (25 % conversion) (**Figure 4a**). EUGO showed a similar response, where the presence of catalase increased the conversion from 70 % to above 80 % (**Figure 4b**). For EUGO, conversions in the presence of SOD did not show a beneficial effect. Furthermore, incubations with both catalase and SOD did not improve conversions compared to incubation with only catalase. Lastly, to ascertain a negligible effect of superoxide on the conversions, both reactions were evaluated with increasing amounts of SOD (0, 10, 100 and 500 U mL$^{-1}$), keeping the catalase concentration constant at 35 U mL$^{-1}$. For all four SOD concentrations, no statistical effect was found (**Figure 4a,b**).
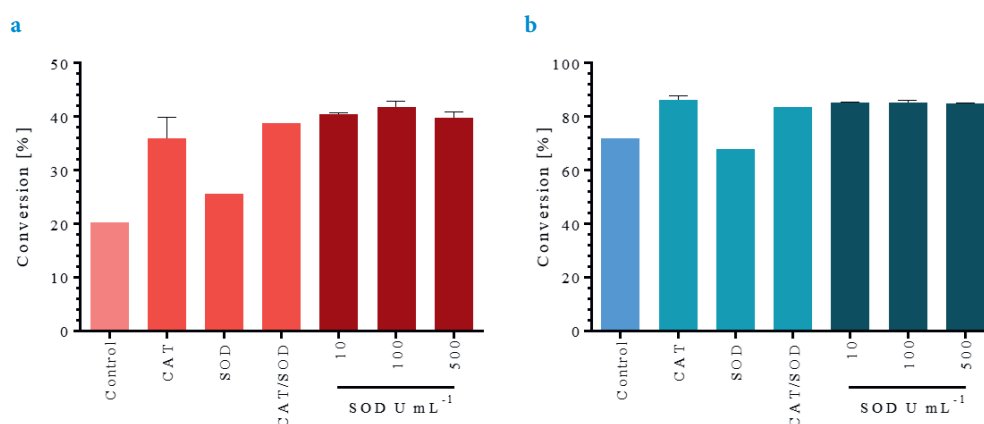
**Figure 4.** **Effect of catalase and superoxide dismutase in PAMO_WT and EUGO bioconversions.** Reactions using (a) 0.50 μM PAMO_WT or (b) 0.15 μM EUGO as biocatalysts were tested for the conversion of phenylacetone or vanillyl alcohol. For PAMO, reaction mixture included 200 mM sodium phosphite, 200 μM NADPH and 1.0 μM PTDH. For both, the presence of ROS scavengers was tested, 35 U mL$^{-1}$ catalase and (dark) increasing amounts of SOD (0-500 U mL$^{-1}$). Reactions were performed in 50 mM TRIS buffer pH 9.0, with 150 mM NaCl for 24 h at 24 °C with constant shaking. Reactions were performed in duplicate.

## DISCUSSION

### Effect of ROS in biomolecules

From cell biology studies, it has become clear that ROS are important signaling molecules for aerobic organisms. For example, in eukaryotes, ROS are involved in the regulation of growth, apoptosis, hypoxia, innate immunity and signaling processes[35–40]. The role of ROS is dynamic and can be regulated through the amount, duration, and localization of these species. An unbalance of the redox state can translate in altered ROS levels leading to cell damage or even death[38,41]. Specifically, targets and reactivity of ROS are highly regulated by 1) molecular proximity —this because of their short life-times— and 2) whether they perform one or two-electron oxidations[42]. On the other hand, for aerobic bacteria, dioxygen by-products, such as superoxide, hydrogen peroxide and hydroxyl radical are generated during cell growth. These are mainly produced by membrane–associated respiratory chain enzymes[43]. Therefore, ROS-mediated oxidative stress may be considered as an inevitable biological process, caused by univalent electron leaking. As defense against the damaging effect of ROS, aerobic organisms have developed a whole set of enzymes to deal with ROS, such as superoxide dismutase and catalase. Biochemically, ROS are relevant to study because they can perform redox modification to a wide diverse group of biomolecules, including lipids, DNA, RNA and proteins[11,44–49].

Superoxide, studied in this work, is an ionic specie with a single unpaired electron, and it is formed after the first sequential reduction of dioxygen. Despite its high reduction potential is not an aggressive oxidant because the anionic charge limits its reactivity with electron-rich centers[50]. However, this molecule can be protonated to hydroperoxyl radical, which is lipid soluble and a more reactive molecule[51]. This ROS can readily oxidize iron-sulfur clusters at a rate that is almost diffusion limited because of electrostatic attraction[42]. This leads to the inactivation of enzymes like aconitase and bacterial dehydratases, which are involved in amino acid synthesis[52–54]. Interestingly, neither PAMO$_{WT}$ nor EUGO small-scale conversions at high pH were significantly affected by superoxide. This may indicate that these flavoproteins are not so sensitive towards superoxide. In cells, superoxide is precursor of others more reactive oxygen species, such as peroxynitrite and hydrogen peroxide[10,11]. The latter, also evaluated in this work, is the results of the transfer of two one-electrons to dioxygen, and is a stronger and more stable non-radical oxidant. Hydrogen peroxide shows a high reactivity with amino acids with low-redox potential, such as methionine, cysteine, histidine and tryptophan[55–57]. This oxidation often leads to loss of enzyme activity, structural changes, aggregation and even immunogenicity[44,58,59]. Because the hydrogen peroxide dipole moment is similar to that of water, this molecule can diffuse into/through cell membranes. Consequently, even when present in the extracellular environment, hydrogen peroxide can cause redox stress within cells by crossing the membrane. Therefore, for both cell-based and enzyme-catalyzed conversions, hydrogen peroxide can be a significant oxidative stressor. The presence of an efficient ROS scavenger, such as catalase, would be required to avoid the inactivation and disruption of metabolic and catalytic processes[60]. In this sense, the high hydrogen peroxide levels achieved by oxidases under industrial concentrations of substrate, could hinder their use for applied regioselective and enantiospecific oxidations[61,62]. The use of ROS scavengers can help to mitigate this. For PAMO$_{WT}$ and EUGO, we could show that the addition of catalase significantly increased the catalytic performance (about 10 % higher conversion).

## Superoxide production

The oxygen dilemma has been well studied in the context of monooxygenase-based biocatalytic applications[63]. Monooxygenases need to generate a reactive oxygenating enzyme intermediate, which is prone to formation of futile ROS, lowering the biocatalytic performance. Some promising solutions have been described for this problem[64–66]. Nevertheless, the production of ROS, particularly superoxide generation, by flavoprotein monooxygenases and oxidases has been scantly studied. In the past, it was observed that flavoprotein oxidases and hydroxylases did not appear to produce significant amounts of superoxide anion[8,22]. Evidently, this was not the case for NADPH oxidase or xanthine oxidase, both of which are very well studied superoxide-producing enzymes[67,68]. Our study indicates that all tested flavoenzymes produce significant amounts of superoxide, and the superoxide levels were significantly higher at pH 9.0-10.0 (**Figure 3b**). This suggests that

superoxide production may be a more widespread phenomenon than previously thought. Although the dioxygen accessibility, desolvation and geometry contribute to increasing the reactivity of dioxygen with $Fl_{red}$, the spin incompatibility dictates initial formation of a transient radical pair upon the first one-electron transfer step. In any case, for some oxidases, it has been suggested that positive charges —such as His and Lys— close to C4α-locus could enhance the electrostatic stabilization of the transition state[69,70]. The latter would limit ROS leaking and lead to higher oxidation rates.

Interestingly, for pyranose-2 oxidase, a GMC-superfamily member, the initial one-electron transfer coupled to a proton transfer generates a hydroperoxyl radical[71,72]. Evidence is lacking to ascertain whether this proton-coupling mechanism is shared by most oxygen-utilizing flavoproteins. However, intriguingly, superoxide production was boosted by high pH for all four flavoenzyme tested in this study. This effect, could be related to the scenario of the pyranose-2 oxidase mechanism, where it could be that the initially formed superoxide would need to be protonated to accept the second one-electron to form hydrogen peroxide. The latter process would be disrupted at high pH. Nevertheless, there could be other explanations for this phenomenon. In the 1970s, Massey and coworkers studied extensively the redox chemistry of flavins and demonstrated for free flavins that the reactivity of $Fl_{sq}$ with dioxygen for the production of $Fl_{ox}$ and superoxide radical anion was pH-dependent[8,73]. Interestingly, this reaction exhibited a rate of $1 \times 10^8$ $M^{-1}s^{-1}$ for the anionic flavin radical and $1 \times 10^4$ $M^{-1}s^{-1}$ for the neutral $Fl_{sq}$. The high superoxide production could be involved in an even more complex reaction, due to the fact that free $Fl_{sq}$ exhibits a $pK_a$ of approximately 8.5 —for glucose oxidase, the $Fl_{sq}$ remains anionic down to pH 5.0[74]— and the thermodynamic balance of the neutral or anionic $Fl_{sq}$ could differ affecting the stabilization of the radical pair and the rates of ROS production.

## ROS as bottleneck in biocatalysis

To avoid detrimental oxidative stress, a better characterization and understanding of the parameters that govern the formation of ROS due to the uncoupling of flavin-dependent enzymes are highly relevant. For some proteins, the sensitivity to ROS generated by uncoupling might be approached by protein engineering. ROS-sensitive residues can be targeted for mutagenesis. Also, more polar residues can be designed to obstruct the access towards the active site or other sensitive parts of the protein structure [48,55]. ROS-sensitivity becomes even more complex if the modified amino acids are within the active site, or if the cofactor involved in the catalytic mechanism is oxidized[60,75]. This situation is critical for heme-dependent enzymes, where hydrogen peroxide can react with the iron from the heme group[60]. This reaction generates a ferryl species and the already named hydroxyl radical, which will oxidize the porphyrin. This modified heme-group can later interact with dioxygen to form verdo-heme and biliverdin. Subsequently, biliverdin can induce the release of the iron. This process is known as heme-bleaching, which irreversibly inactivates

the enzyme[75,76]. Another ROS-related situation is for the P450 monooxygenases, a class of enzymes with a heme-thiolate cofactor with a somewhat more complex mechanism than flavoprotein monooxygenases[60,63,77]. P450s perform highly selective oxyfunctionalisation using reducing equivalents. However, they often suffer of higher uncoupling levels than flavoprotein monooxygenases[78]. This is partly because, while flavoenzymes such as PAMO directly accept an hydride from NADPH, the P450 monooxygenases acquire electrons via two consecutive one-electron transfer processes with the help of a reductase (domain)[77]. These reductases have been associated to uncoupling. In some cases, the uncoupling can reach levels of up to 90 %[78,79]. As for flavoprotein monooxygenases, the uncoupling wastes highly valuables electrons. Despite their very attractive chemistry, this greatly limits the industrial applicability of P450 and flavoprotein monooxygenases for biocatalysis[63,80,81]. Although it has been studied extensively, the mechanism that leads to uncoupling in P450 monooxygenases is not fully understood yet, and it may be modulated by the substrate, pH, ionic force, oxygen/substrate concentration, or other factors[82–84]. Therefore, for monooxygenases, the profiles of ROS formation under different operational conditions would be valuable information (e.g. **Table S2** and **Table S3**). Specifically, accurate characterizations could help overcome the oxygen dilemma by distinguishing conditions in which the oxygenation is optimized and the ROS accumulation is minimal. The latter can facilitate downstream processing avoiding the loss of valuable hydrides.

The reactivity between flavoproteins and molecular oxygen is a captivating topic. It plays an important role in biology, is relevant for biotechnological applications, and is still not fully understood in terms of mechanisms. More insight is needed to fully understand how flavoenzymes react with dioxygen, what determines the fate of the reduced oxygen species being formed, and why flavoenzymes cannot completely prevent the formation of superoxide radical anion. With this study, we have revealed that superoxide and hydrogen peroxide are common ROS formed by flavoprotein oxidases and monooxygenases. Formation of both species can be tuned by varying reaction conditions. These insights may help to improve biocatalytic applications while they can also help in understanding biological processes.

## MATERIAL AND METHODS

### Expression and purification

All chemicals and reagents were purchased from Sigma-Aldrich or Difco, unless otherwise stated. Phosphite dehydrogenase (PTDH) from *Pseudomonas stutzeri* was obtained from Gecco-Biotech. The pBAD-based expression plasmids, coding sequences of $PAMO_{WT}$, $PAMO_{C65D}$, EUGO and HMFO that included a hexa-histidine tag were obtained as previously described[26–28]. The plasmids were transformed into chemocompetent *E. coli*

NEB 10β cells (New England Biolabs). For expression, single colonies were grown in Luria Bertani Broth supplemented with 50 μg mL$^{-1}$ ampicillin at 37 °C overnight with constant shaking. Aliquots of the precultures (1:100) were used to inoculate 400 mL of fresh Terrific Broth medium supplemented with ampicillin. Cultures were incubated at 37 °C with constant shaking until an OD$_{600nm}$ of 0.8 was reached. Then, expression was induced for 48 hours at 24 °C by adding L-arabinose at final 0.02 % w/v concentration. Cells were harvested and disrupted by sonication in lysis buffer (50 mM TRIS pH 8.0, 500 mM NaCl, 1.5 mg mL$^{-1}$ lysozyme, 10 μM FAD and 1.0 mM PMSF) for 10 min (5 s on and 5 s off) at 70 % amplitude (VCX30 Vibra-Cell, Sonic and materials Inc.). Subsequently, cell free extracts (CFE) were obtained by centrifugation at 10,000 r.p.m. for 60 min at 4 °C. For purification, CFE were filtered (at 0,45 μm), loaded into a buffer equilibrated Ni$^{+2}$-sepharose HP resin (GE Health Care) and incubated for 60 min at 4 °C in a rotating system. Later, the column was washed with 10 CV of 50 mM TRIS pH 8.0, 500 mM NaCl and 20 mM imidazole. Yellow fractions were collected after elution with 50 mM TRIS pH 8.0, 500 mM NaCl and 400 mM imidazole. Lastly, eluted fractions were desalted using an EconoPac 10-DG (Bio-Rad) pre-equilibrated with 50 mM TRIS pH 8.0 and 150 mM NaCl. Purity of each protein sample was analyzed and verified by SDS-PAGE and UV-visible spectrum (250-700$_{nm}$). When the proteins were shown to contain some Fl$_{sq}$, the samples were fully oxidized by incubating with 2.0 mM potassium ferricyanide (K$_3$[Fe(CN)$_6$] and 1.0 mM FAD at 4 °C overnight in a rotating system, then the samples were subsequently desalted as explained above.

A second purification procedure to eliminate catalase and SOD was performed through FPLC (AKTA Pure) using a Ni$^{+2}$-sepharose 6 Fast Flow column (HisTrap FF column, GE Health Care). An imidazole gradient was carried out from 5 mM to 400 mM with a flow of 1 mL min$^{-1}$ (12 CV). The fractions were tested for the presence of catalase activity (explained below) and the catalase-free fractions were pooled and desalted. For each protein, their own molar extinction coefficient was used for the quantification (εPAMO$_{441nm}$=12.4 mM$^{-1}$ cm$^{-1}$, εEUGO$_{441nm}$=14.2 mM$^{-1}$ cm$^{-1}$ or εHMFO$_{456nm}$=10.7 mM$^{-1}$ cm$^{-1}$). Finally, the obtained pure flavoproteins were flash frozen in liquid nitrogen and stored at -80 °C until use.

## SPECTROPHOTOMETRIC ANALYSIS

### Linear response determination

Determination of the linear response of the activity over different amount of proteins was evaluated to verify the linearity and feasibility of the experimental conditions (**Figure S9**). For PAMO, the NADPH depletion was followed at 340$_{nm}$ in the presence of 100 μM NADPH and in presence or absence of 1.0 mM phenylacetone. Meanwhile, for EUGO and HMFO, 1.0 mM vanillyl alcohol or 15 mM HMF were used to follow the oxidase activity. Then, the

activity was followed at $515_{nm}$ through an indirect HRP detection system[85]. Consequently, the amount of protein was selected for further experiments. For NADPH oxidase activity, 2.0 μM $PAMO_{WT}$ and 0.05 μM $PAMO_{C65D}$ were used. In presence of phenylacetone, both variants were used at 0.05 μM. While for EUGO and HMFO, reactions were carried out using 0.15 and 0.015 μM, respectively.

## Spectrophotometric detection and quantification

For the measurement of all three analytes (NADPH, hydrogen peroxide, and superoxide), a microplate reader with pathlength correction was used (Synergy H1 Hybrid multi-mode reader, Biotek). For the NADPH quantification, the absorbance at $340_{nm}$ was followed in time ($\varepsilon_{340nm}$ = 6,220 $M^{-1}$ $cm^{-1}$). For hydrogen peroxide detection, the xylenol orange method was used[86]. The reactions were incubated at different times, and subsequently 20 μL samples were taken and mixed with 180 μL of the xylenol solution (1:99, $(NH_4)_2Fe(SO_4)_2 \cdot 6H_2O$ and 2.5 M $H_2SO_4$: 100 mM sorbitol and 125 μM xylenol orange). The mixtures were incubated for 15 min and were measured by absorbance at $595_{nm}$. For each condition calibration curves were prepared, and the same protocol as explained above was followed. Possible interferences were analyzed and discarded during the method optimization. To evaluate presence of catalase in the samples, a calibration curve was performed in presence or absence of the flavoenzymes, and was compared the change by fold over the control. During AKTA purification, each fraction was incubated with 30 μM hydrogen peroxide and compared with a control without enzyme. For superoxide anion detection, a described coupling assay using a soluble tetrazolium salt was employed[86]. For this method, 200 μM of WST-1 was mixed at the different conditions (Dojindo Molecular Technologies, Inc). Each condition had its own prepared blank. For detection and quantification, the extinction coefficient of the produced WST-1 formazan was used ($\varepsilon_{438nm}$ = 37,000 $M^{-1}$ $cm^{-1}$). For SOD detection, a calibration curve with 0.02 U $mL^{-1}$ xanthine oxidase and increasing concentrations of xanthine were prepared. The calibration curves were made in presence or absence of the flavoenzymes, and SOD presence were evaluated by comparing the change by fold over the control. To evaluate the effect of WST-1, NADPH depletion and oxidase activities were tested in the presence or absence of 200 μM WST-1 (method explained above). Finally, the uncoupling ratios of $PAMO_{WT}$ and $PAMO_{C65D}$ were calculated as $\frac{[hydrogen\ peroxide]+[superoxide]}{[NADPH\ oxidized]} x\ 100$.

## Conditions for chemical effect over oxidases and monooxygenases

Each experiment was performed in triplicate in 300 μL final volume at 25 °C for 30 min, with double-orbital shaking. Reactions were prepared in 50 mM phosphate buffer (pH 6.0), 50 mM TRIS (pH 7.0-9.0) or 50 mM CHES (pH 9.0-10.0), all in the presence of 150 mM NaCl. For PAMO, increased amounts of NADPH were evaluated (50, 100 or 200 μM). For each experiment, NADPH was fresh prepared and spectrophotometrically verified.

For all the flavoenzymes, tested substrate concentrations were equivalent to 0, ½x $K_M$, $K_M$, 10x $K_M$ or 20x $K_M$ (0, 25, 50, 500 or 1000 µM phenylacetone for PAMO$_{WT}$ and PAMO$_{C65D}$, 25, 50, 500 and 1,000 µM vanillyl alcohol for EUGO or 0, 0.75, 1.5, 15 and 30 mM HMF for HMFO). For cosolvents effect analysis, methanol (MeOH), ethanol (EtOH), 1,2-dioxane, acetonitrile (ACN) or isopropanol at 0, 2.5, 5 or 7.5 % v/v were evaluated.

## Determination of melting temperature

The apparent melting temperatures ($T_M^{app}$) were determined by using the ThermoFAD method[87]. In triplicate, a sample with 1.0 mg mL$^{-1}$ of enzyme in absence or presence of 200 µM WST-1 was prepared in 50 mM TRIS buffer pH 8.0 and 150 mM NaCl. The mixtures were prepared in a 96-well PCR plate and exposed to a temperature ramp from 25 to 99 °C using an RT-PCR instrument (CFX96-Touch, Bio-Rad). The gradient increased the temperature well by 0.5 °C every 10 s. Fluorescence was measured using a 450–490 nm excitation filter and a 515–530 nm emission filter. The $T_M^{app}$ or unfolding temperature was determined as the maximum of the sigmoidal curve of the derivative of fluorescence against temperature.

## Bioconversions

Conversions of phenylacetone were carried out in duplicate using 0.5 µM PAMO$_{WT}$, 1.0 µM PTDH, 200 mM Na$_2$PO$_3$•6H$_2$0, 200 µM NADPH, 150 mM NaCl, 1.0-50 mM phenylacetone, 0-35 U mL$^{-1}$ catalase and 0-500 U mL$^{-1}$ SOD in 50 mM TRIS buffer pH 9.0, for 24 h at 24 °C with constant shaking. Then, the samples were extracted by mixing three times with one volume of ethyl acetate and 0.025 % v/v mesitylene as external standard. Subsequently, anhydrous sulfate magnesium was added to the organic layer to remove the residual water. The analytic analysis was performed using a HP-5MS agilent column in a GCMS-QP2010 Ultra (Shimadzu) with electron ionization and quadrupole separation. Chromatograms and MS spectra were analyzed using GCMSsolution Postrun Analysis 4.11 (Shimadzu) with a MS library designed by the manufacturer. The conversion of vanillyl alcohol was performed using 0.15 µM EUGO, 20 mM substrate, 0-35 U mL$^{-1}$ catalase and 0-500 U mL$^{-1}$ SOD in 50 mM TRIS buffer pH 9.0 for 24 h at 24 °C in constant shaking. For the analysis, a 200 µL sample was taken and mixed with 600 µL of acetonitrile. Later, after centrifugation the sample was analyzed by HPLC using an Altima C18-column and (A) 0.08% formic acid and (B) acetonitrile as eluents. The obtained peaks were compared with analytic standards.

## Statistical Analysis

All analyses were performed using GraphPad Prism v6.05 for Windows (GraphPad Software, La Jolla, CA, United States). To assess statistically significant differences One-way ANOVA test was used with a Dunnet's post-test to compare each value with the control, a $p$-value <0.05 was used for significance.

# SUPPORTING INFORMATION CHAPTER III

## Tables

**Table S1. PAMO rates of NADPH oxidation and ROS production.** Initial rates for NADPH depletion and ROS production of PAMO$_{WT}$ and PAMO$_{C65D}$ were measured at increasing concentrations of NADPH or phenylacetone. For rates determination at different pH, reactions were performed with 100 µM NADPH and 1.0 mM phenylacetone (20x $K_M$). Experiments were performed in triplicate, averaged results are shown.

| | | PAMO$_{WT}$ | | | PAMO$_{C65D}$ | | |
|---|---|---|---|---|---|---|---|
| | Condition | $k_{NADPH}$ [s$^{-1}$] | $k_{O_2^{\cdot-}}$ [s$^{-1}$] | $k_{H2O2}$ [s$^{-1}$] | $k_{NADPH}$ [s$^{-1}$] | $k_{O_2^{\cdot-}}$ [s$^{-1}$] | $k_{H2O2}$ [s$^{-1}$] |
| NADPH | 50 µM | 0.019 | <4.4x10$^{-3}$ | 0.020 | 1.75 | 0.024 | 1.14 |
| | 100 µM | 0.025 | <5.8x10$^{-3}$ | 0.023 | 1.99 | 0.028 | 1.23 |
| | 200 µM | 0.027 | <5.7x10$^{-3}$ | 0.025 | 2.27 | 0.032 | 2.22 |
| Substrate | $K_M$ | 1.01 | 0.019 | 0.012 | 1.30 | 0.030 | 1.28 |
| | 10x $K_M$ | 2.03 | 0.028 | 0.010 | 1.74 | 0.033 | 0.78 |
| | 20x $K_M$ | 2.00 | 0.024 | 0.030 | 1.79 | 0.036 | 0.53 |
| pH | 6.0 | 1.01 | <1.0x10$^{-4}$ | 0.017 | 0.70 | <1.2x10$^{-3}$ | 0.23 |
| | 7.0 | 2.02 | 0.020 | 0.012 | 2.08 | 0.012 | 0.31 |
| | 8.0 | 2.00 | 0.024 | 0.029 | 1.80 | 0.036 | 0.50 |
| | 9.0 | 1.20 | 0.046 | 0.019 | 1.40 | 0.050 | 0.44 |

**Table S2. Effect of the substrate concentration in the uncoupling ratio for PAMO$_{WT}$ and PAMO$_{C65D}$.** Ratios values were obtained from reactions performed in 50 mM TRIS buffer, pH 8.0 with 150 mM NaCl, 100 µM NADPH and increasing concentrations of phenylactenone. Reactions were carried out in triplicate at 25 °C for 30 min in constant shaking.

| | PAMO variants | |
|---|---|---|
| Phenylacetone | WT [%] | C65D [%] |
| 25 µM | 82.6 | 93.4 |
| 50 µM | 9.12 | 74.6 |
| 500 µM | 5.42 | 42.4 |
| 1000 µM | 4.96 | 29.3 |

**Table S3. Summary of total uncoupling ratios for PAMO$_{WT}$ and PAMO$_{C65D}$ at different pH and in presence of co-solvents.** All the reactions were tested in the presence of 1.0 mM phenylacetone, 100 μM NADPH and 150 mM NaCl. Evaluation of cosolvent effect were carried out in 50 mM TRIS buffer, pH 8.0. Reactions were performed in triplicate at 25 °C for 30 min with constant shaking.

| Feature | Condition | PAMO variants | |
|---|---|---|---|
| | | WT [%] | C65D [%] |
| pH | 50 mM KPi pH 6.0 | 6.6 | 29.6 |
| | 50 mM TrisHCl pH 7.0 | 6.3 | 31.0 |
| | 50 mM TrisHCl pH 8.0 | 5.0 | 29.3 |
| | 50 mM TrisHCl pH 9.0 | 7.8 | 42.3 |
| | 50 mM CHES pH 9.0 | 6.4 | 46.4 |
| | 50 mM CHES pH 10.0 | 7.2 | 34.0 |
| Co-solvent | Water | Control | 4.9 | 36.4 |
| | MeOH | 2.5 % v/v | 3.8 | 24.2 |
| | | 5 % v/v | 2.4 | 18.8 |
| | | 7.5 % v/v | 1.7 | 17.2 |
| | EtOH | 2.5 % v/v | 3.3 | 21.6 |
| | | 5 % v/v | 2.7 | 16.4 |
| | | 7.5 % v/v | 2.7 | 14.0 |
| | ACN | 2.5 % v/v | 4.4 | 33.7 |
| | | 5 % v/v | 3.8 | 27.2 |
| | | 7.5 % v/v | 3.4 | 24.0 |
| | Isop | 2.5 % v/v | 4.5 | 34.0 |
| | | 5 % v/v | 5.8 | 32.8 |
| | | 7.5 % v/v | 3.9 | 18.0 |

**Table S4. Superoxide radical anion production rates for EUGO.** Initial rates for superoxide radical were measured at different pHs. Experiments were performed in triplicate, averaged results are shown. Reaction were performed in the presence of 150 mM NaCl.

| $k_{O_2^{\cdot-}}$ [s$^{-1}$] | pH |
|---|---|
| $<1 \times 10^{-3}$ | 50 mM KPi 6.0 |
| $<1 \times 10^{-3}$ | 50 mM TRIS 7.0 |
| $<1 \times 10^{-3}$ | 50 mM TRIS 8.0 |
| 0.035 | 50 mM TRIS 9.0 |
| 0.11 | 50 mM CHES 9.0 |
| 0.067 | 50 mM CHES 10.0 |

# Figures
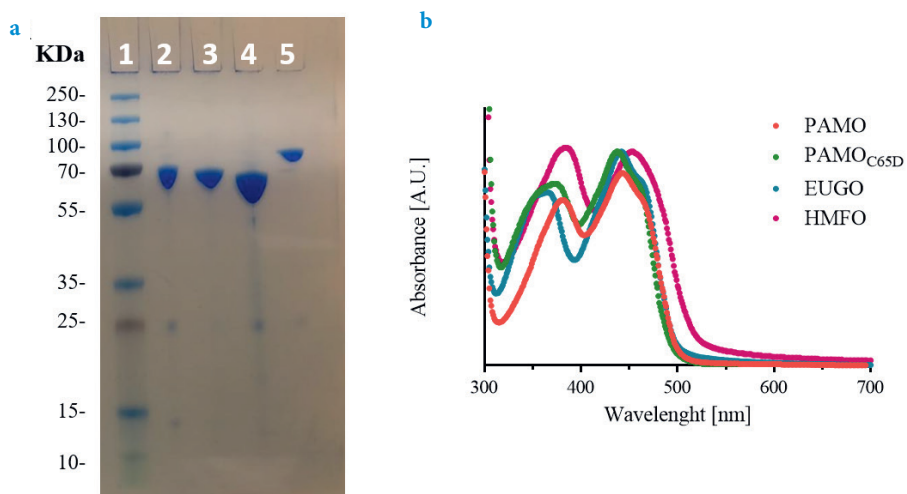


**Figure S1. Flavoprotein purifications.** (a) SDS-PAGE for purified oxidases and monooxygenases. 1) Ladder, 2) PAMO$_{WT}$, 3) PAMO$_{C65D}$, 4) EUGO and 5) HMFO. (b) UV-vis spectra for PAMO$_{WT}$ (red), PAMO$_{C65D}$ (green), EUGO (blue) and HMFO (pink). Ratio$_{280:440}$ was calculated to observe apo/ holoprotein relationship, for all four flavoenzymes the ratio was between 9-18.
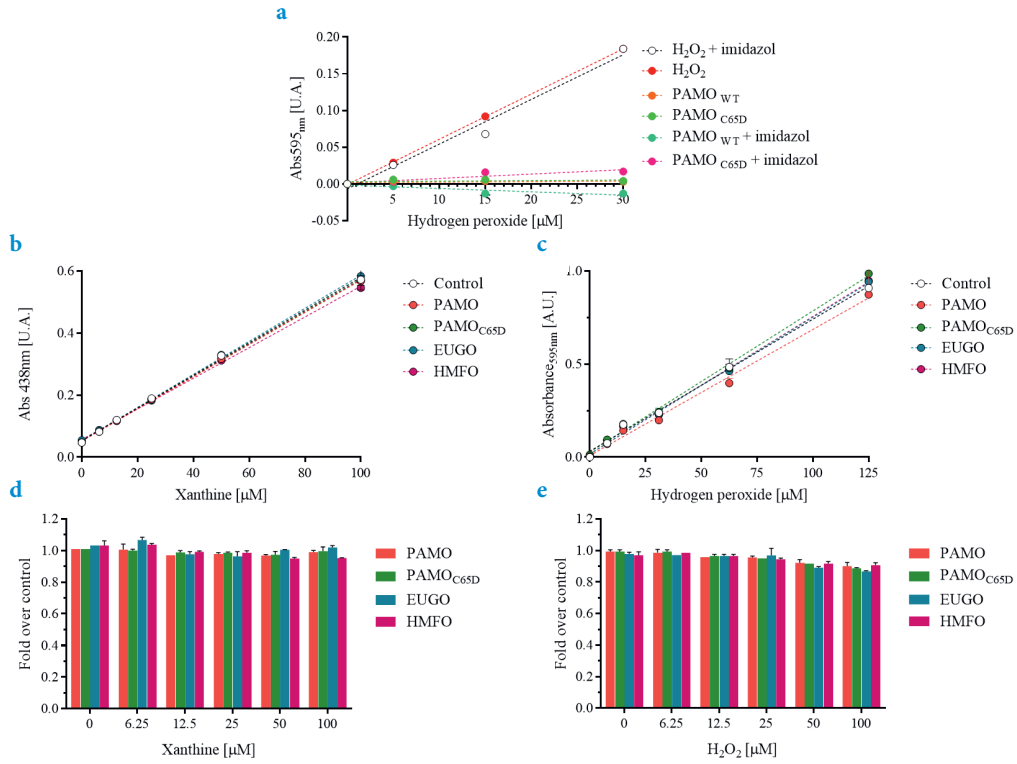
**Figure S2. Verification of the presence of ROS scavengers in purified proteins.** Calibration curves were prepared to confirm the presence of ROS scavengers. For the analysis a calibration curve of hydrogen peroxide was prepared in the presence of proteins samples from (a) first IMAC step and (c,e) after further IMAC purification. For superoxide anion, a curve of xanthine:xanthine oxidase was prepared and the (b,e) superoxide level was measured by WST-1 coupling method.
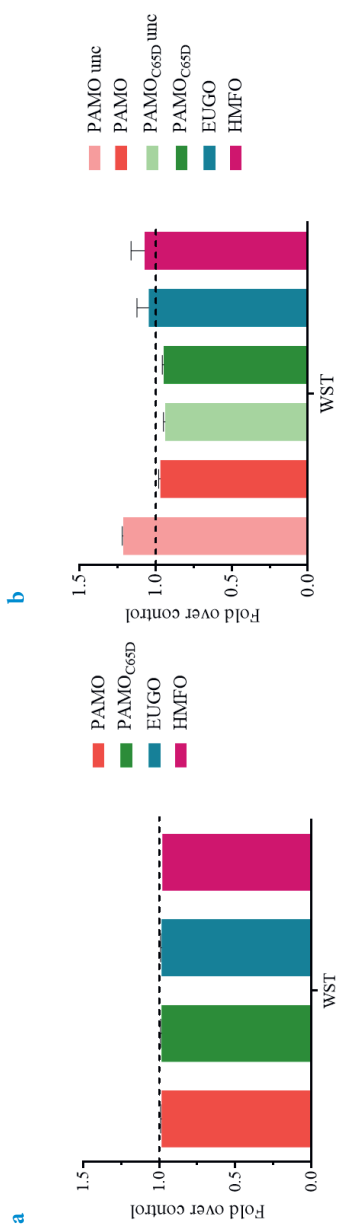
**Figure S3. Effect of tetrazolium salt WST-1 over thermostability and enzymatic activity.** The possible effect of WST-1 over $PAMO_{WT}$ (red), $PAMO_{C65D}$ (green), EUGO (blue) and HMFO (pink) was evaluated. Variation of fold over the control was analyzed for (a) thermostability via ThermoFAD method and (b) enzymatic activity by spectrophotometry. For monooxygenases, the activity was evaluated in the presence (dark) or absence (light) of phenylacetone following the absorbance at $340_{nm}$. For oxidases the activity was measured through the HRP indirect coupling system at $515_{nm}$.
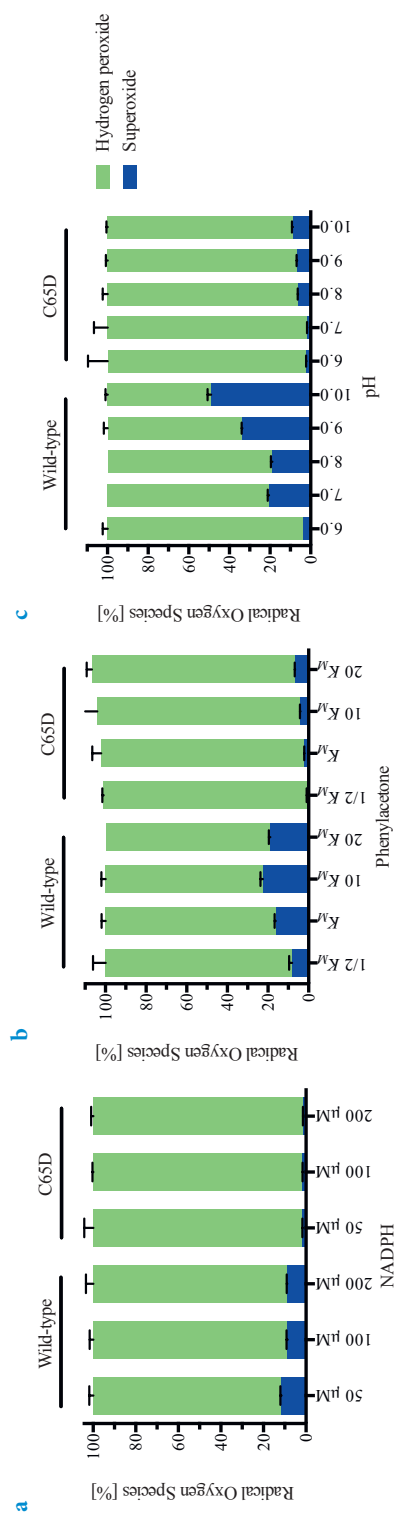


**Figure S4. Reactive Oxygen Species ratios for $PAMO_{WT}$ and $PAMO_{C65D}$.** The percentage composition of hydrogen peroxide (green) and superoxide (blue) levels at increasing concentration of (a) NADPH, (b) phenylacetone and (c) pH is shown.
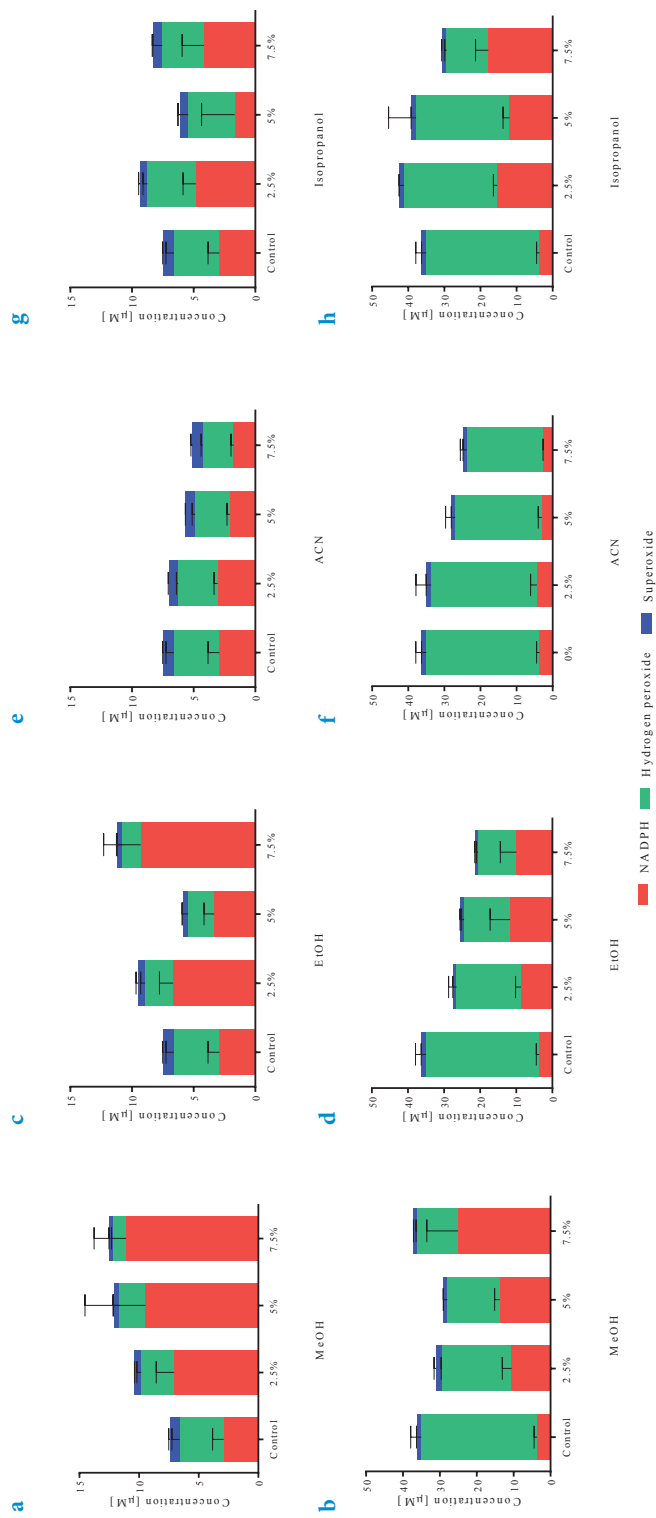
**Figure S5. Effect of co-solvent on ROS production by PAMO.** Measurements of hydrogen peroxide (green), NADPH (red) and superoxide anion (blue) were determined in presence of different co-solvent concentration for PAMO$_{WT}$ (upper panel) and PAMO$_{C65D}$ (bottom panel). Reactions were tested using 2.5, 5 and 7.5 % v/v of (a,b) methanol, (c,d) ethanol, (e,f) acetonitrile and (g,h) isopropanol. Reactions were performed in 50 mM TRIS pH 8.0 buffer with 100 µM NADH, 1 mM phenylacetone and 150 mM NaCl. All reactions were incubated for 30 min at 25 °C in constant double-orbital shaking. Each reaction was prepared in triplicate.

**3**

113
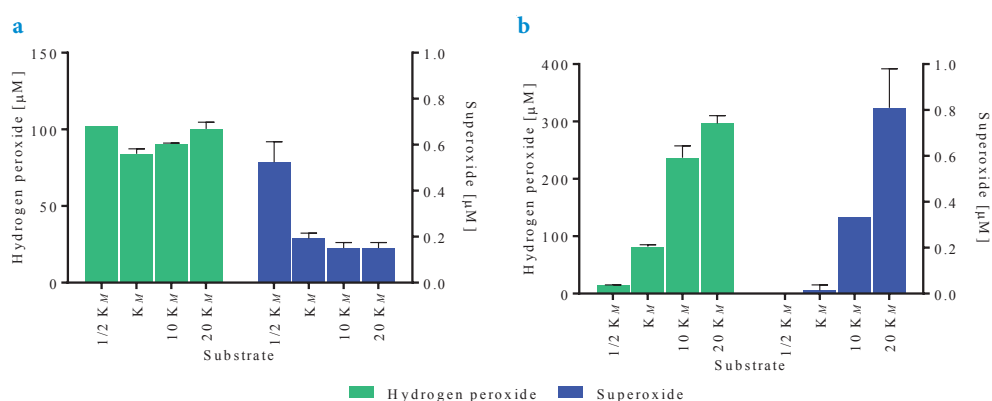
**Figure S6. Effect of substrate levels on the production of ROS by oxidases.** Reactions with equivalent to ½x $K_M$, $K_M$, 10x $K_M$ or 20x $K_M$ of (a) HMF for HMFO and (b) vanillyl alcohol for EUGO were evaluated for ROS production. Levels of hydrogen peroxide (green) and superoxide anion (blue) were measured in 50 mM TRIS buffer, pH 8.0 in presence of 150 mM NaCl for 30 min at 25 °C in constant double-orbital shaking. Reactions were performed in triplicate.
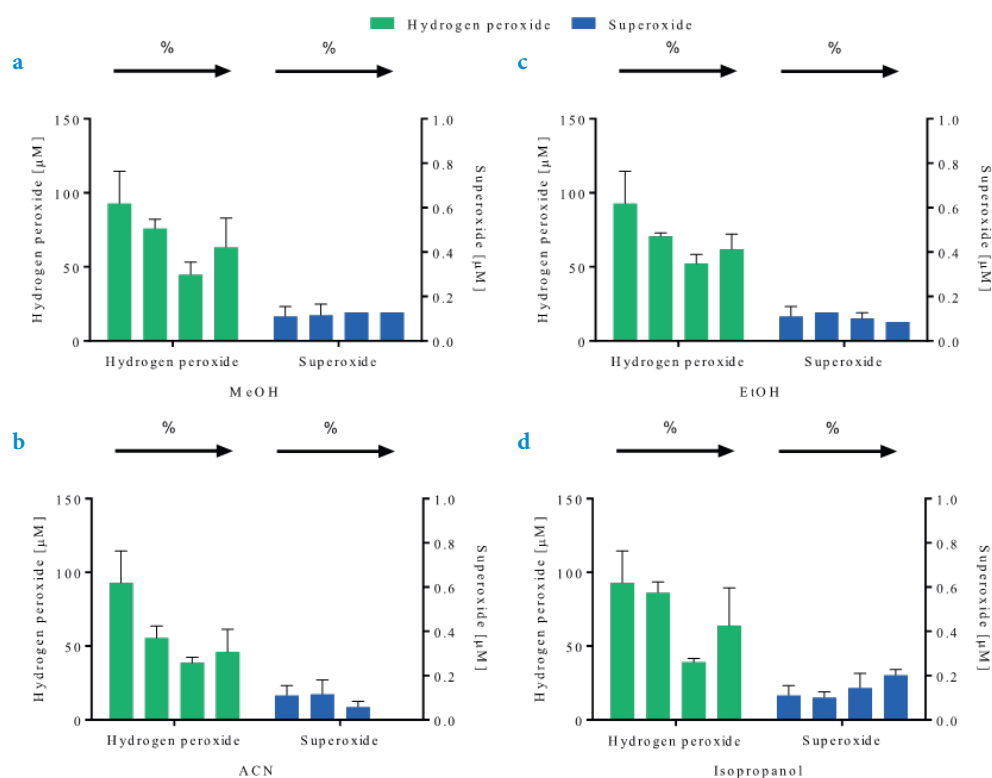


**Figure S7. Effect of the presence of co-solvent on HMFO.** Formation of hydrogen peroxide (green) and superoxide anion (blue) was analyzed in presence of different co-solvent concentration for HMFO. Reactions were tested at 2.5, 5 and 7.5 % v/v of (a) methanol, (b) ethanol, (c) acetonitrile and (d) isopropanol. Reactions were carried out in 50 mM TRIS buffer, pH 8.0 with 30 mM HMF and 150 mM NaCl. All reactions were performed for 30 min at 25 °C in constant double-orbital shaking. Reactions were performed in triplicate
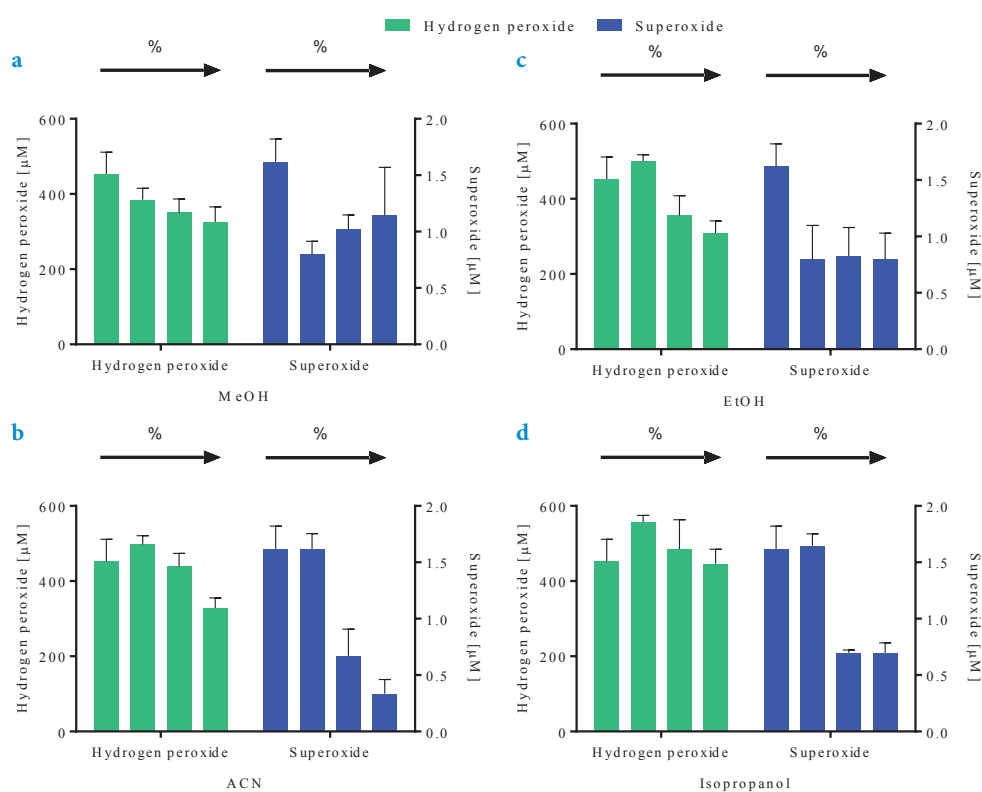
**Figure S8. Effect of the presence of co-solvent on EUGO.** Detection of hydrogen peroxide (green) and superoxide anion (blue) was analyzed in presence of different co-solvent concentration for EUGO. Reactions were tested at 2.5, 5 and 7.5 % v/v of (a) methanol, (b) ethanol, (c) acetonitrile and (d) isopropanol. Reactions were carried out in 50 mM TRIS buffer pH 8.0, with 1 mM vanillyl alcohol and 150 mM NaCl. All reactions were performed for 30 min at 25 °C in constant double-orbital shaking. Reactions were performed in triplicate.
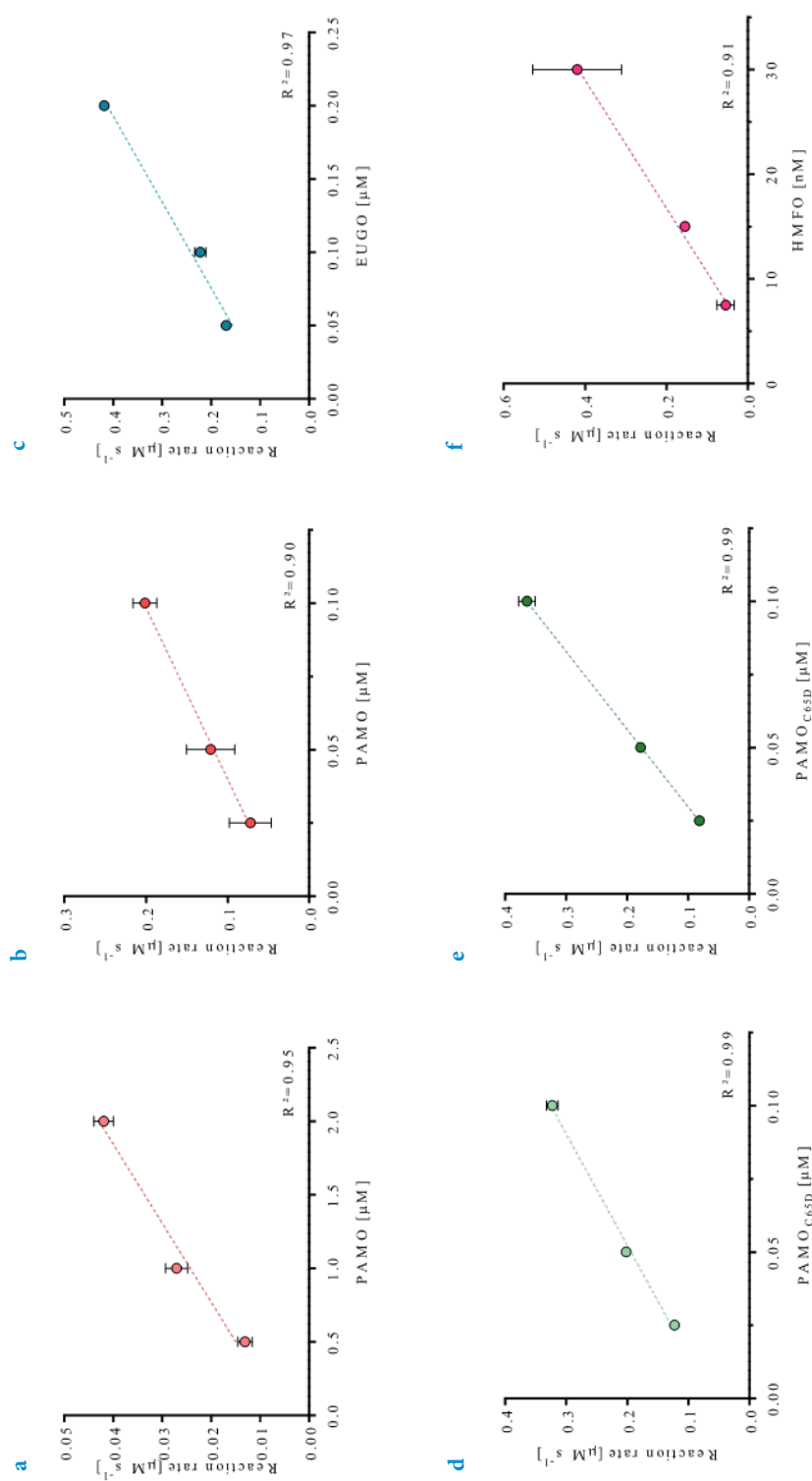
**Figure S9. Determination of linear response for enzymatic activity.** Activity at different protein concentration was tested to evaluate linearity and feasibility of experimental conditions. NADPH oxidase activity in presence of 100 μM NADPH was tested for (a) PAMO$_{WT}$ and (b) PAMO$_{C65D}$. Oxygenase activity in presence of 1.0 mM phenylacetone was evaluated for (c) PAMO$_{WT}$ and (d) PAMO$_{C65D}$. While, oxidase activity was tested for (d) EUGO and (e) HMFO using 1.0 mM vanillyl alcohol or 15 mM HMF, respectively.

# REFERENCES

1. Romero, E., Gómez Castellanos, J. R., Gadda, G., Fraaije, M. W. & Mattevi, A. Same substrate, many reactions: Oxygen activation in flavoenzymes. *Chem. Rev.* 118, 1742–1769 (2018).

2. Huijbers, M. M. E., Montersino, S., Westphal, A. H., Tischler, D. & van Berkel, W. J. H. Flavin dependent monooxygenases. *Arch. Biochem. Biophys.* 544, 2–17 (2014).

3. Fraaije, M. W. & Mattevi, A. Flavoenzymes: diverse catalysts with recurrent features. *Trends Biochem. Sci.* 25, 126–132 (2000).

4. Massey, V. Introduction: flavoprotein structure and mechanism. *FASEB J.* 9, 473–475 (1995).

5. Greening, C., Ahmed, F.H., Mohamed, A.E., Lee, B.M., Pandey, G., Warden, A.C., Scott, C., Oakeshott, J.G., Taylor, M.C. and Jackson, C.J. Physiology, biochemistry, and applications of F420-and Fo-dependent redox reactions. *Microbiol. Mol. Biol. Rev.* 80, 451–493 (2016).

6. Matthews, A., Saleem-Batcha, R., Sanders, J.N., Stull, F., Houk, K.N. & Teufel, R. Aminoperoxide adducts expand the catalytic repertoire of flavin monooxygenases. *Nat. Chem. Biol.* 16, 556–563 (2020).

7. Leys, D. Flavin metamorphosis: cofactor transformation through prenylation. *Curr. Opin. Chem. Biol.* 47, 117–125 (2018).

8. Massey, V. Activation of molecular oxygen by flavins and flavoproteins. *J. Biol. Chem.* 269, 22459–22462 (1994).

9. Gibson, Q. H., Massey, V. & Atherton, N. M. The nature of compounds present in mixtures of oxidized and reduced flavin mononucleotides. *Biochem. J.* 85, 369–383 (1962).

10. Cohen, G. Enzymatic/nonenzymatic sources of oxyradicals and regulation of antioxidant defensesa. *NYASA* 738, 8–14 (1994).

11. Hrycay, E. G. & Bandiera, S. M. Involvement of Cytochrome P450 in reactive oxygen species formation and cancer. in *Advances in Pharmacology* vol. 74 35–84 (Academic Press Inc., 2015).

12. Lewis, D. F. Oxidative stress: The role of cytochromes P450 in oxygen activation. *J. Chem. Technol. Biotechnol.* 77, 1095–1100 (2002).

13. Mayhew, S. G. The effects of pH and semiquinone formation on the oxidation--reduction potentials of flavin mononucleotide: a reappraisal. *Eur. J. Biochem.* 265, 698–702 (1999).

14. Mattevi, A. To be or not to be an oxidase: challenging the oxygen reactivity of flavoenzymes. *Trends Biochem. Sci.* 31, 276–283 (2006).

15. Massey, V. The chemical and biological versatility of riboflavin. (2000):283-296

16. Massey, V. The reactivity of oxygen with flavoproteins. *International Congress Series.* 1233, 3–11 (2002).

17. Heuts, D. P. H. M., Scrutton, N. S., McIntire, W. S. & Fraaije, M. W. What's in a covalent bond? *FEBS J.* 276, 3405–3427 (2009).

18. Wang, R. & Thorpe, C. Reactivity of medium-chain acyl-CoA dehydrogenase toward molecular oxygen. *Biochemistry* 30, 7895–7901 (1991).

19. Baron, R., Riley, C., Chenprakhon, P., Thotsaporn, K., Winter, R.T., Alfieri, A., Forneris, F., van Berkel, W.J., Chaiyen, P., Fraaije, M.W. & Mattevi, A. Multiple pathways guide oxygen diffusion into flavoenzyme active sites. *Proc. Natl. Acad. Sci.* 106, 10603–10608 (2009).

20. Chaiyen, P., Fraaije, M. W. & Mattevi, A. The enigmatic reaction of flavins with oxygen. *Trends Biochem. Sci.* 37, 373–380 (2012).

21. Van Berkel, W. J. H., Kamerbeek, N. M. & Fraaije, M. W. Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts. *J. Biotechnol.* 124, 670–689 (2006).

22. Massey, V., Strickland, S., Mayhew, S.G., Howell, L.G., Engel, P.C., Matthews, R.G., Schuman, M. & Sullivan, P.A. The production of superoxide anion radicals in the reaction of reduced flavins and flavoproteins with molecular oxygen. *Biochem. Biophys. Res. Commun.* 36, 891–897 (1969).

23. Daithankar, V. N., Wang, W., Trujillo, J. R. & Thorpe, C. Flavin-linked Erv-family sulfhydryl oxidases release superoxide anion during catalytic turnover. *Biochemistry* 51, 265–272 (2012).

24. Binda, C., Robinson, R.M., del Campo, J.S.M., Keul, N.D., Rodriguez, P.J., Robinson, H.H., Mattevi, A. & Sobrado, P. An unprecedented NADPH domain conformation in lysine monooxygenase NbtG provides insights into uncoupling of oxygen consumption from substrate hydroxylation. *J. Biol. Chem.* 290, 12676–12688 (2015).

25. Fraaije, M.W., Wu, J., Heuts, D.P., Van Hellemond, E.W., Spelberg, J.H.L. & Janssen, D.B. Discovery of a thermostable Baeyer-Villiger monooxygenase by genome mining. *Appl. Microbiol. Biotechnol.* 66, 393–400 (2005).

26. Brondani, P. B., Dudek, H. M., Martinoli, C., Mattevi, A. & Fraaije, M. W. Finding the switch: Turning a Baeyer-Villiger monooxygenase into a NADPH oxidase. *JACS* (2014).

27. Jin, J., Mazon, H., van den Heuvel, R. H. H., Janssen, D. B. & Fraaije, M. W. Discovery of a eugenol oxidase from *Rhodococcus* sp. strain RHA1. *FEBS J.* 274, 2311–2321 (2007).

28. Dijkman, W. P. & Fraaije, M. W. Discovery and characterization of a 5-hydroxymethylfurfural oxidase from *Methylovorus* sp. strain MP688. *Appl. Environ. Microbiol.* 80, 1082–1090 (2014).

29. Macheroux, P. UV-visible spectroscopy as a tool to study flavoproteins. *Methods Mol. Biol.* 131, 1–7 (1999).

30. Malito, E., Alfieri, A., Fraaije, M. W., Mattevi, A. & Matthews, R. G. Crystal structure of a Baeyer–Villiger monooxygenase. *Proceedings of the National Academy of Sciences*, 101, 13157-13162 (2004).

31. Andrade, L. H., Pedrozo, E. C., Leite, H. G. & Brondani, P. B. Oxidation of organoselenium compounds. A study of chemoselectivity of phenylacetone monooxygenase. *J. Mol. Catal. B Enzym.* 73, 63–66 (2011).

32. Rodríguez, C., De Gonzalo, G., Fraaije, M. W. & Gotor, V. Ionic liquids for enhancing the enantioselectivity of isolated BVMO-catalysed oxidations. *Green Chem.* 12, 2255–2260 (2010).

33. Secundo, F., Fiala, S., Fraaije, M.W., de Gonzalo, G., Meli, M., Zambianchi, F. & Ottolina, G. Effects of water miscible organic solvents on the activity and conformation of the Baeyer-Villiger monooxygenases from *Thermobifida fusca* and *Acinetobacter calcoaceticus*: A comparative study. *Biotechnol. Bioeng.* 108, 491–499 (2011).

34. Nguyen, Q.T., de Gonzalo, G., Binda, C., Rioz-Martínez, A., Mattevi, A. & Fraaije, M.W. Biocatalytic properties and structural analysis of eugenol oxidase from *Rhodococcus jostii* RHA1: A versatile oxidative biocatalyst. *ChemBioChem* 17, 1359 (2016).

35. Hamanaka, R. B. & Chandel, N. S. Mitochondrial reactive oxygen species regulate cellular signaling and dictate biological outcomes. *Trends Biochem. Sci.* 35, 505–513 (2010).

36. Tormos, K.V., Anso, E., Hamanaka, R.B., Eisenbart, J., Joseph, J., Kalyanaraman, B. & Chandel, N.S. Mitochondrial complex III ROS regulate adipocyte differentiation. *Cell Metab.* 14, 537–544 (2011).

37. West, A.P., Brodsky, I.E., Rahner, C., Woo, D.K., Erdjument-Bromage, H., Tempst, P., Walsh, M.C., Choi, Y., Shadel, G.S. & Ghosh, S. TLR signalling augments macrophage bactericidal activity through mitochondrial ROS. *Nature* 472, 476–480 (2011).

38. Brand, M. D. Mitochondrial generation of superoxide and hydrogen peroxide as the source of mitochondrial redox signaling. *Free Radic. Biol. Med.* 100, 14–31 (2016).

39. D'Autréaux, B. & Toledano, M. B. ROS as signalling molecules: Mechanisms that generate specificity in ROS homeostasis. *Nature Reviews Molecular Cell Biology*, 8, 813–824 (2007).

40. Orrenius, S., Gogvadze, V. & Zhivotovsky, B. Mitochondrial Oxidative Stress: Implications for Cell Death. *Annu. Rev. Pharmacol. Toxicol.* 47, 143–183 (2007).

41. Brieger, K., Schiavone, S., Miller, F. J. & Krause, K.-H. Reactive oxygen species: from health to disease. *Swiss Med. Wkly.* 142, w13659 (2012).

42. Dickinson, B. & Chang, C. J. chemistry and biology of reactive oxygen species in signaling or stress responses. *Nat. Chem. Biol.* 7, 504 (2011).

43. Cabiscol Català, E., Tamarit Sumalla, J. & Ros Salvador, J. Oxidative stress in bacteria and protein damage by reactive oxygen species. *Int. Microbiol.* 3, 3–8 (2000).

44. Simat, T. J. & Steinhart, H. Oxidation of free tryptophan and tryptophan residues in peptides and proteins. *J. Agric. Food Chem.* 46, 490–498 (1998).

45. Christman, M. F., Morgan, R. W., Jacobson, F. S. & Ames, B. N. Positive control of a regulon for defenses against oxidative stress and some heat-shock proteins in Salmonella typhimurium. *Cell* 41, 753–762 (1985).

46. Jang, S. & Imlay, J. A. Micromolar intracellular hydrogen peroxide disrupts metabolism by damaging iron-sulfur enzymes. *J. Biol. Chem.* 282, 929–937 (2007).

47. Imlay, J. A. & Linn, S. DNA damage and oxygen radical toxicity. *Science,* 240, 1302–1309 (1988).

48. Hernandez, K., Berenguer-Murcia, A., C Rodrigues, R. & Fernandez-Lafuente, R. Hydrogen peroxide in biocatalysis. A dangerous liaison. *Curr. Org. Chem.* 16, 2652–2672 (2012).

49. Hayyan, M., Ali Hashim, M. & AlNashef, I. M. Superoxide Ion: Generation and chemical implications. *Chem. Rev.* 116, 3029–3085 (2016).

50. Winterbourn, C. C. Biological chemistry of superoxide radicals. *Chemtexts* 6, 7 (2020).

51. Mailloux, R. J. & Harper, M.-E. Uncoupling proteins and the control of mitochondrial reactive oxygen species production. *Free Radic. Biol. Med.* 51, 1106–1115 (2011).

52. Flint, D. H., Tuminello, J. F. & Emptage, M. H. The inactivation of Fe-S cluster containing hydro-lyases by superoxide. *J. Biol. Chem.* 266, 22369–22376 (1993).

53. Imlay, J. A. Pathways of oxidative damage. *Annu. Rev. Microbiol.* 57, 395–418 (2003).

54. Gardner, P. R. & FridovichS, I. Superoxide sensitivity of the *Escherichia coli* Aconitase. *J. Biolo* 266, 19328–19333 (1991).

55. Stadtman, E. R. & Levine, R. L. Free radical-mediated oxidation of free amino acids and amino acid residues in proteins. *Amino Acids* 25, 207–218 (2003).

56. Finnegan, M., Linley, E., Denyer, S.P., McDonnell, G., Simons, C. & Maillard, J.Y. Mode of action of hydrogen peroxide and other oxidizing agents: differences between liquid and gas forms. *J. Antimicrob. Chemother.* 65, 2108–2115 (2010).

57. Gottfredsen, R. H., Larsen, U. G., Enghild, J. J. & Petersen, S. V. Hydrogen peroxide induce modifications of human extracellular superoxide dismutase that results in enzyme inhibition. *Redox Biol.* 1, 24–31 (2013).

58. Rosenberg, A. S. Effects of protein aggregates: an immunologic perspective. *AAPS J.* 8, E501–E507 (2006).

59. Liu, J.L., Lu, K.V., Eris, T., Katta, V., Westcott, K.R., Narhi, L.O. & Lu, H.S. In vitro methionine oxidation of recombinant human leptin. *Pharm. Res.* 15, 632–640 (1998).

60. Burek, B. O., Bormann, S., Hollmann, F., Bloh, J. Z. & Holtmann, D. Hydrogen peroxide driven biocatalysis. *Green Chem.* 21, 3232–3249 (2019).

61. Greenfield, P. F., Kittrell, J. R. & Laurence, R. L. Inactivation of immobilized glucose oxidase by hydrogen peroxide. *Anal. Biochem.* 65, 109–124 (1975).

62. Volpato, G., C Rodrigues, R. & Fernandez-Lafuente, R. Use of enzymes in the production of semi-synthetic penicillins and cephalosporins: drawbacks and perspectives. *Curr. Med. Chem.* 17, 3855–3873 (2010).

63. Holtmann, D. & Hollmann, F. The Oxygen dilemma: A severe challenge for the application of monooxygenases? *ChemBioChem* 17, 1391–1398 (2016).

64. Kim, S.J., Joo, J.C., Kim, H.S., Kwon, I., Song, B.K., Yoo, Y.J. & Kim, Y.H. Development of the radical-stable *Coprinus cinereus* peroxidase (CiP) by blocking the radical attack. *J. Biotechnol.* 189, 78–85 (2014).

65. Vasudevan, P. T. & Weiland, R. H. Deactivation of catalase by hydrogen peroxide. *Biotechnol. Bioeng.* 36, 783–789 (1990).

66. Zhang, W., Burek, B.O., Fernández-Fueyo, E., Alcalde, M., Bloh, J.Z. & Hollmann, F. Selective activation of C− H bonds in a cascade process combining photochemistry and biocatalysis. *Angew. Chemie Int. Ed.* 56, 15451–15455 (2017).

67. Takeya, R. & Sumimoto, H. Molecular mechanism for activation of superoxide-producing NADPH oxidases. *Mol. Cells* 16, 271–277 (2003).

68. Fridovich, I. Quantitative aspects of the production of superoxide anion radical by milk xanthine oxidase. *J. Biol. Chem.* 245, 4053–4057 (1970).

69. Klinman, J. P. How do enzymes activate oxygen without inactivating themselves? *Acc. Chem. Res.* 40, 325–333 (2007).

70. Gadda, G. Oxygen activation in flavoprotein oxidases: The importance of being positive. *Biochemistry* 51, 2662–2669 (2012).

71. Wongnate, T., Surawatanawong, P., Chuaboon, L., Lawan, N. & Chaiyen, P. The mechanism of sugar c−h bond oxidation by a flavoprotein oxidase occurs by a hydride transfer before proton abstraction. *Chem. – A Eur. J.* 25, 4460–4471 (2019).

72. Sucharitakul, J., Wongnate, T. & Chaiyen, P. Hydrogen Peroxide Elimination from C4a-hydroperoxyflavin in a Flavoprotein Oxidase Occurs through a Single Proton Transfer from Flavin N5 to a Peroxide Leaving Group. *J. Biol. Chem.* 286, 16900–16909 (2011).

**3**

73. Massey, V., Palmer, G. & Ballou, D. P. Oxidases and related redox systems. *TE King, HS Mason, M. Morrison, Univ. Park Press. Balt.* 1, 25 (1973).

74. Roth, J. P. & Klinman, J. P. Catalysis of electron transfer during activation of O$_2$ by the flavoprotein glucose oxidase. *Proc. Natl. Acad. Sci. U. S. A.* 100, 62–67 (2003).

75. Ayala, M., Batista, C. V & Vazquez-Duhalt, R. Heme destruction, the main molecular event during the peroxide-mediated inactivation of chloroperoxidase from *Caldariomyces fumago. JBIC J. Biol. Inorg. Chem.* 16, 63–68 (2011).

76. 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. B Enzym.* 134, 238–246 (2016).

77. Dong, J., Fernández-Fueyo, E., Hollmann, F., Paul, C.E., Pesic, M., Schmidt, S., Wang, Y., Younes, S. & Zhang, W. Biocatalytic oxidation reactions: A chemist's perspective. *Angew. Chemie Int. Ed.* 57, 9238–9261 (2018).

78. Morlock, L. K., Böttcher, D. & Bornscheuer, U. T. Simultaneous detection of NADPH consumption and H2O2 production using the Ampliflu™ Red assay for screening of P450 activities and uncoupling. *Appl. Microbiol. Biotechnol.* 102, 985–994 (2018).

79. Cryle, M. J. & De Voss, J. J. The role of the conserved threonine in p450bm3 oxygen activation: Substrate-Determined hydroxylation activity of the Thr268Ala mutant. *ChemBioChem* 9, 261–266 (2008).

80. Bernhardt, R. Cytochromes P450 as versatile biocatalysts. *J. Biotechnol.* 124, 128–145 (2006).

81. Julsing, M. K., Cornelissen, S., Bühler, B. & Schmid, A. Heme-iron oxygenases: powerful industrial biocatalysts? *Current Opinion in Chemical Biology* vol. 12 177–186 (2008).

82. Degregorio, D., Sadeghi, S. J., Di Nardo, G., Gilardi, G. & Solinas, S. P. Understanding uncoupling in the multiredox centre P450 3A4–BMR model system. *JBIC J. Biol. Inorg. Chem.* 16, 109–116 (2011).

83. Denisov, I. G., Baas, B. J., Grinkova, Y. V & Sligar, S. G. Cooperativity in cytochrome P450 3A4 linkages in substrate binding, spin state, uncoupling, and product formation. *J. Biol. Chem.* 282, 7066–7076 (2007).

84. Zangar, R. C., Davydov, D. R. & Verma, S. Mechanisms that regulate production of reactive oxygen species by cytochrome P450. *Toxicol. Appl. Pharmacol.* 199, 316–331 (2004).

85. Martin, C., Ovalle Maqueo, A., Wijma, H. J. & Fraaije, M. W. Creating a more robust 5-hydroxymethylfurfural oxidase by combining computational predictions with a novel effective library design. *Biotechnol. Biofuels* 11, 56 (2018).

86. Robinson, R. & Sobrado, P. Substrate binding modulates the activity of *mycobacterium smegmatis* G, a flavin-dependent monooxygenase involved in the biosynthesis of hydroxamate-containing siderophores. *Biochemistry* 50, 8489–8496 (2011).

87. Forneris, F., Orru, R., Binivento, D., Chiarelli, L. R. & Mattevi, A. ThermoFAD, a Thermofluor-adapted flavin ad hoc detection system for protein folding and ligand binding. *FEBS J.* 276, 2833–2840 (2009).

**3**

# IV

# Genome mining of oxidation modules in trans–acyltransferase polyketide synthases reveals a culturable source for lobatamides

Reiko Ueoka[§], Roy A. Meoded[§], **Alejandro Gran-Scheuch**, Agneya Bhushan, Marco W. Fraaije, and Jörn Piel

[§]These authors contributed equally

Contribution to this work: Characterization of LmbC-Ox and Oock

## ABSTRACT

Bacterial trans–acyltransferase polyketide synthases (trans–AT PKSs) are multimodular megaenzymes that biosynthesize many bioactive natural products. They contain a remarkable range of domains and module types that introduce different substituents into growing polyketide chains. As one such modification, we recently reported Baeyer–Villiger–type oxygen insertion into nascent polyketide backbones, thereby generating malonyl thioester intermediates. In this work, genome mining focusing on architecturally diverse oxidation modules in trans–AT PKSs led us to the culturable plant symbiont *Gynuella sunshinyii*, which harbors two distinct modules in one orphan PKS. The PKS product was revealed to be lobatamide A, a potent cytotoxin previously known from a marine tunicate. Biochemical studies show that one module generates glycolyl thioester intermediates, while the other is proposed to be involved in oxime formation. The data suggest varied roles of oxygenation modules in the biosynthesis of polyketide scaffolds and support the importance of trans–AT PKSs in the specialized metabolism of symbiotic bacteria

**Keywords:** Bacterial natural products, biosynthesis, marine natural products, polyketides, Baeyer–Villiger monooxygenases

## INTRODUCTION

Bacterial complex polyketides belong to the most important natural product classes of therapeutic value.[1] Two distinct families of multimodular polyketide synthases (PKSs), termed cis– and trans–acyltransferase (*trans*–AT) PKSs, generate most of these compounds. The textbook biosynthetic model is represented by the *cis*–AT PKS family, such as the erythromycin PKS.[1,2] These enzymes are usually composed of a limited set of functionally different biosynthetic multidomain modules that elongate and modify intermediates, and give rise to contiguous carbon chains carrying keto, hydroxy, double-bond, and carbon branch modifications. In contrast, the highly complex *trans*-AT PKSs can accommodate a remarkable diversity of modules (>150 module architectures are currently known)[3] and employ integrated domains as well as free-standing, trans-acting enzymes with often poorly understood functions.[4] The large functional range of *trans*–AT PKSs suggests high biosynthetic diversity outside the scope of canonical polyketides. Another phenomenon of *trans*-AT PKSs is their dominant role in the specialized metabolism of non-actinomycete symbiotic bacteria, with examples reported from fungi,[5] insects,[6] marine invertebrates,[7] and plants.[8] To fully access this biosynthetic potential, insights into the function of *trans*-AT PKS assembly lines and their components are needed to improve our ability to predict their products,[1] identify alternative producers to invertebrate sources,[9] and provide novel tools for biosynthetic engineering. In previous work, we identified the insertion of oxygen into growing polyketide chains as a non–canonical reaction by which *trans*-AT PKSs diversify product skeletons.[9] As shown for the trans-acting flavin adenine dinucleotide (FAD)–dependent oxygenase OocK from the oocydin pathway, the enzyme performs a Baeyer–Villiger (BV) oxidation on a β–ketothioester intermediate to generate a malonyl derivative (**Figure 1a**) that is then further elongated. Preservation or hydrolysis of the new ester moiety gives rise to oxygen atoms within polyketide backbones (oocydin B and haterumalides), carboxylate pseudostarters (oocydin A), or terminal alcohols (introduced by the OocK homologues in the pederin[6] and toblerol pathway,[8] **Figure 1b**). Mid–chain oxygen atoms also occur in other polyketides (**Figure 1c**).[10] Although these moieties could principally result from various processes,[11] the presence of OocK homologs in diverse orphan PKS biosynthetic gene clusters (BGCs; **Figures S1–S4**) suggests that nature employs oxygen insertion more widely to construct unusual polyketide scaffolds. In this work, we applied a genome-mining strategy to investigate the wider scope of flavoprotein monooxygenases acting during polyketide elongation. The data revealed an architecturally distinct oxygen insertion module that generates a glycolyl intermediate, as well as evidence for a third module type involved in oxime formation. Both new modules are used in the biosynthesis of lobatamides, potent cytotoxins that were previously known from marine tunicates, but identified here from a culturable plant symbiont.
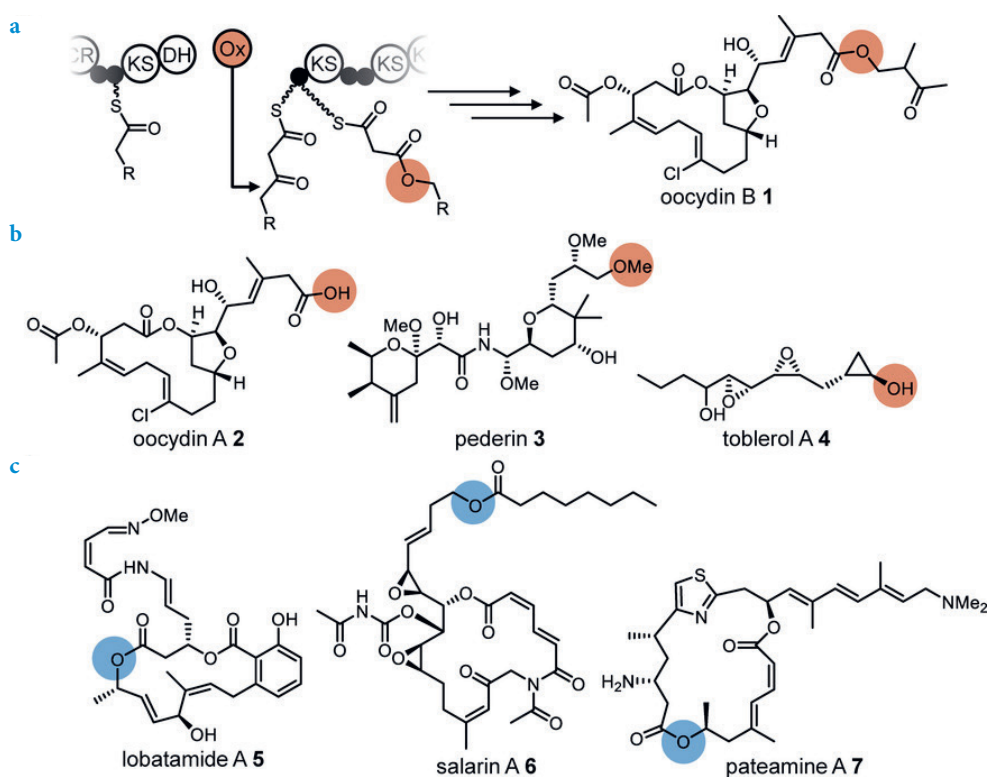
**Figure 1. Oxygen incorporation into polyketide backbones**. Moieties derived from oxygen insertion are highlighted in orange. Oxygen atoms of unknown biosynthetic origin are highlighted in blue. a) Mechanism for oxygen insertion into growing polyketide backbones, exemplified by OocK, a trans-acting BV monooxygenase from the oocydin (oocydin B, **1**) pathway. The enzyme converts a β–ketothioester substrate into a malonyl derivative. b) Oocydin A (**2**), pederin (**3**), and toblerol A (**4**) with carboxylate- and alcohol-type termini arising from oxygen insertion and ester cleavage. c) Biosynthetically unassigned polyketides harboring mid-chain oxygens: lobatamide A (**5**) from a tunicate, and salarin A (**6**) and pateamine A (**7**) from marine sponges.

## RESULTS AND DISCUSSION

We initiated our work by identifying OocK homologues and other PKS–associated flavoprotein monooxygenases encoded in orphan BGCs using GenBank, genome neighborhood,[12] phylogenetic, and manual analyses (status November 2019; **Figures S1–S4**). These revealed 69 candidate *trans*-AT PKS systems harboring such enzymes (**Figures S1–S4**). The phylogram contained a large clade containing the functionally related OocK, PedG, and TobD, as well as homologues from uncharacterized PKSs. Another functionally assigned clade bears module-integrated oxygenase (Ox) domains for which the variant from *Burkholderia* sp. FERM BP–3421 was shown to introduce an epoxide unit in spliceostatin (**Figure S1**).[8] In addition, other non-OocK clades mostly belong to Ox

domains integrated in various module types, thus suggesting further biosynthetic diversity. Focusing on uncharacterized oxygenases (**Figure S1**), an orphan *trans*–AT system (termed *lbm* PKS, **Figure 2a**) in the Gram-negative bacterium *Gynuella sunshinyii* (NZ_CP007142) appeared an intriguing candidate, since the PKS contains oxygenase modules from two different unassigned clades (**Figure S1**, **Table S1**). *G. sunshinyii* is an unusual halophilic root-associated plant symbiont[13] that was recently recognized as a talented producer of diverse natural products.[14] Its genome contains six *trans*–AT PKS BGCs, the highest known number for any organism. The two integrated Ox domains are located in the PKS proteins LmbA and LmbC, and will be referred to as LmbA–Ox and LmbC–Ox, respectively. Additional features in these modules are a methyltransferase (MT) in the LmbC-Ox module, which suggests α–C–methylation, and a predicted non-elongating KS (KS$^0$) in the LmbA–Ox module, which suggests that a moiety introduced by the upstream module is further modified. This upstream module is located at the N–terminus of LbmA and resembles an NRPS loading module with predicted[15] glycine specificity. A C–terminal TE domain is encoded on the PKS gene *lmbE* at the downstream end, which overall suggests collinearity between the PKS gene order and biosynthetic events.
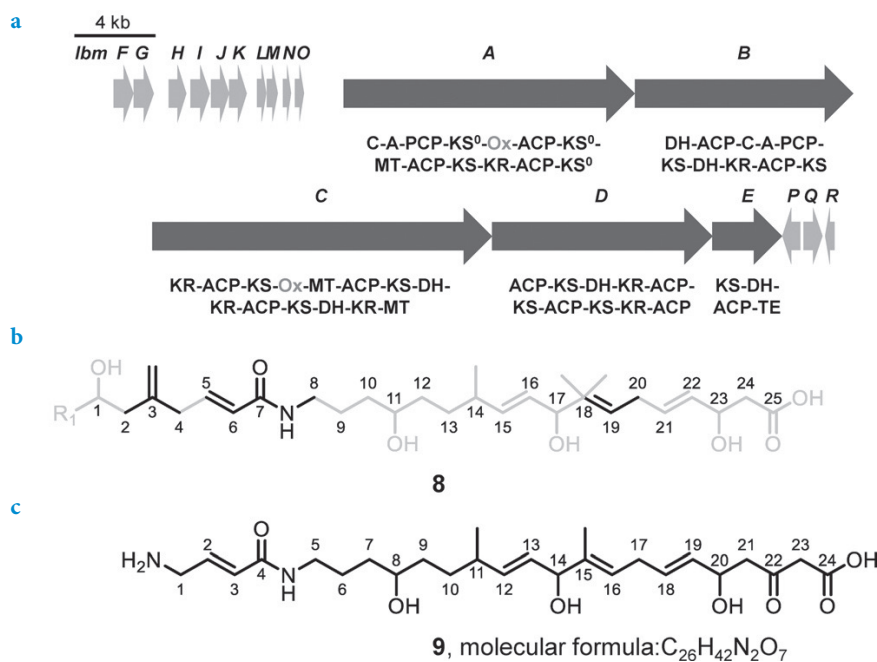


**Figure 2. The *lbm* BGC and predicted polyketide structures**. a) BGC architecture with core PKS and accessory genes marked in dark and light grey, respectively. The domain architecture of the core PKSs is shown using the following abbreviations: A=adenylation domain; C=condensation domain; DH=dehydratase; ER=enoylreductase; KR=ketoreductase; KS=ketosynthase; KS$^0$=non-elongating KS; MT=methyltransferase; Ox=oxygenase; TE=thioesterase. Ox domains are shown in grey. b) TransATor-predicted structure **8** of the *lbm* product. High- and low-confidence regions are shown in black and grey, respectively. c) Manually refined structure **9** used to find the natural product.

As guidance for the targeted isolation of the *lbm* polyketide, we analyzed the PKS using the recently developed automated prediction tool TransATor.[3] This web application suggests chemical structures for *trans*–AT PKSs from phylogenetically inferred[15] KS substrates. In agreement with the unusual PKS architecture, the TransATor output **8** (**Figure 2b**) contained extended regions of low confidence. After removal of a methyl group from the five-bonded C18, the structure was refined by closer inspection of the modular architecture. Since TransAtor can predict intermediates only for modules that have a downstream KS, no elongation was predicted for the terminal module. A terminal C2 unit was therefore added based on cis-AT PKS rules. To account for the N-terminal NRPS module of LbmA, we suspected glycine for biosynthetic initiation. This replacement also removed an exomethylene group in the TransATor structure that was unlikely because polyketide β–branching components[16] were not found in the BGC. These modifications resulted in the hypothetical structure **9** (**Figure 2c**) as a basis for the analytical work.

To search for the predicted compound, *G. sunshinyii* was cultivated, and the extract was analyzed by ultra–high–performance liquid chromatography/high–resolution mass spectrometry (UHPLC–HRMS). Manual inspection of metabolite-related MS features and the molecular formulae suggested from the high-resolution masses showed a candidate ion peak at *m/z* 513.2230 [*M* +H]$^+$ (**Figure S5**). This corresponds to a molecular formula of $C_{27}H_{32}N_2O_8$ (calc. 513.2231). This formula was closest to that of the predicted structure ($C_{26}H_{42}N_2O_7$) and appeared a good candidate for the *lbm* product. No other orphan PKSs in the genome were predicted to account for a similar chemical formula that contains two nitrogen atoms (i.e., PKSs that contain two NRPS modules).[14] To isolate the compound, bacterial pellets from a 2 L culture of *G. sunshinyii* culture were extracted with acetone. MS-guided fractionation using reversed phase–high performance liquid chromatography (RP–HPLC) provided pure compound **5**. The $^1$H NMR and HSQC data showed the presence of an aliphatic methyl coupling with a vicinal proton, one methyl connected to an sp$^2$ carbon, one methoxy group, three oxymethines, and 11 methines connected to sp$^2$ carbons (**Figure S6**). Analysis of 2D NMR data including COSY, HMBC, and HSQC revealed the structure as that of lobatamide A (**Figure 3a**, **Figures S7–S9**). Lobatamides are potent vacuolar (H$^+$)–ATPase inhibitors that interfere with tumor metastasis.[17] Since they were previously known only from a marine invertebrate, the tunicate *Aplidium lobatum* ,[18] *G. sunshinyii* represents the first culturable source for these compounds. Lobatamides belong to a diverse group of benzolactone (H$^+$)–ATPase inhibitors (**Figure 3a**) with as-yet unknown PKSs and isolated from a remarkable range of organisms, including a sponge (salicylihalamides),[19] a fungus (CJ–12,950),[20] myxobacteria (apicularens),[21] and a *Pseudomonas* strain (oximidines).[22] Their related structures and the identification of bacterial sources suggest that all of these compounds are of prokaryotic origin. For an expansion of this family, see a study on necroximes that was conducted concurrently with this work.[23]
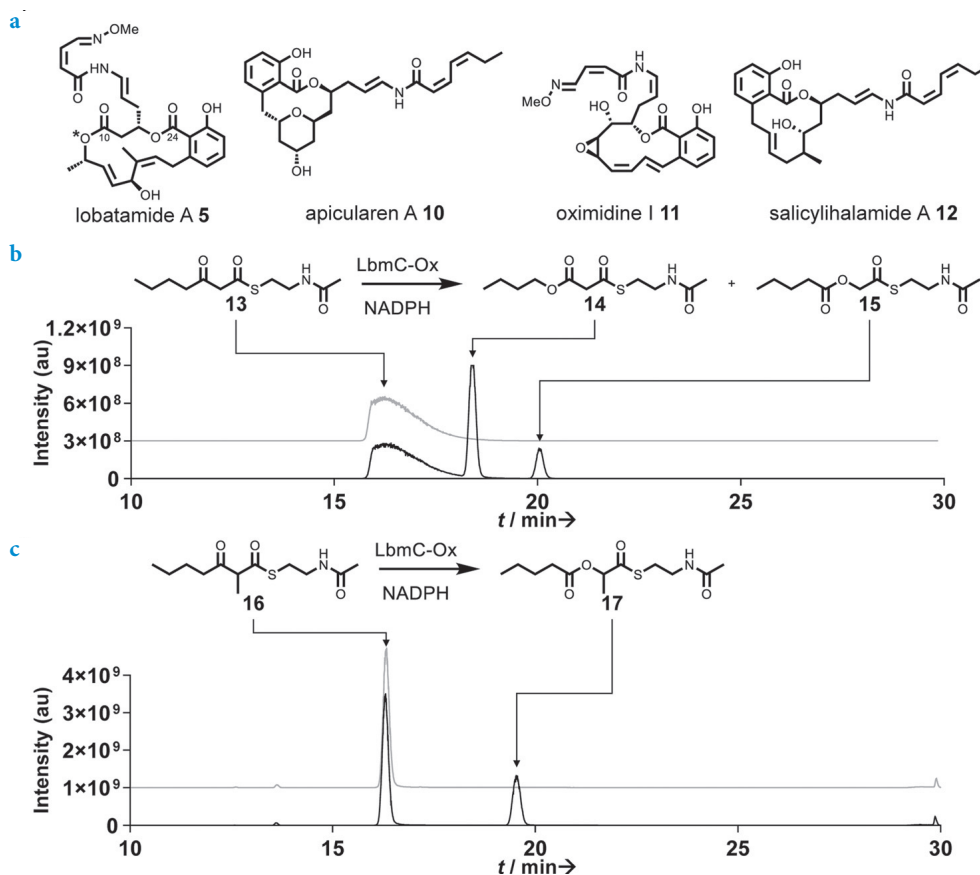
**Figure 3. Selected benzolactone enamide polyketides and enzymatic assays with LbmC-Ox**. a) Benzolactone enamides. The moiety derived from oxygen insertion is highlighted with an asterisk. b,c) UHPLC–HRMS data showing the extracted ion chromatograms (EIC) of assay mixtures, including a boiled-enzyme negative control (upper/grey) and a test reaction using all components (lower/black). b) EIC for **13** and **13** +[$^{16}$O] (calc. for $[M+H]^+$ as 246.1158, and 262.1108, respectively). For the mass spectrum of the new product **15** see Figure S11. c) EIC for **16** and **16** +[$^{16}$O] (calc. for $[M+H]^+$ as 260.1315 and 276.1264, respectively). Mass spectrum of product **17** is shown in **Figure S19**.

Structure **5** features an internal ester that is reversed in comparison to oocydin, pederin, and toblerol biosynthesis. To investigate the role of the LmbC–Ox module in its formation, we cloned the region encoding the Ox domain into a variant of the pET28a expression vector (see experimental section) for expression in *E. coli* as an N-terminally His$_6$-tagged protein (**Figure S10**). Purified LbmC–Ox was obtained as a soluble yellow protein, thus indicating bound FAD. Since the LmbC-Ox module contained an additional MT domain, we suspected that the Ox domain accepts either a β–ketothioester or an α–methyl–β–ketothioester, depending on whether LmbC–Ox acts before or after C–methylation. We

initially assayed the enzyme with the unmethylated thioester **13** as a simplified surrogate of an acyl carrier protein (ACP) –bound 4'–phosphopantetheinyl intermediate. Assay mixtures with test substrate, LbmC-Ox, and NADPH were incubated for 90 minutes at room temperature and then extracted with ethyl acetate. UHPLC–HRMS analysis suggested two new products in the assay mixture, both with a mass difference of +16 Da relative to the substrate (**Figure 3b**). Isolation and NMR–based structure elucidation of the products (**Figures S12–17**, **Table S4**) revealed that the two isomers carry the inserted oxygen either at the γ (**Figure 3b**, product **14**) or the β position (product **15**). The less abundant thioester **15** is noteworthy because it exhibits a reversed ester moiety compared to the OocK/PedG/TobD products and corresponds to the topology present in lobatamides. Suspecting that the major isomer **14** was an aberrant byproduct due to the choice of the surrogate substrate, we also synthesized thioester **16**, which harbors the predicted methyl group at the α position (**Figure 3c**, **Figure S18**). Repetition of the enzyme assay with this new test substrate produced one product peak (**Figure 3c**). Isolation and structure elucidation identified it as thioester **17** (**Figure 3c**, **Figure S19–24**, **Table S5**). The result from the biochemical study matches the internal ester topology at C–10 of lobatamides, that is, facing the same direction as the C–24 ester created by macrolactonization. In lieu of knock-out studies that were so far unsuccessful or of heterologous whole-cluster expression, the unusual function of LmbC–Ox and its location within the PKS reasonably link the *lbm* BGC to lobatamide. The results suggest an overall biosynthetic model for lobatamides as shown in **Figure 4**. A glycine starter would be the source of the oxime moiety, and oxygen insertion and macrolactonization generate the ester groups at C–10 and C–24, respectively. The salicylate group generated by the terminal four modules is also a feature of all other members of the benzolactone family, which currently lack known BGCs. However, similar tetramodular series that generate benzene moieties are also known from the otherwise unrelated psymberin and legioliulin PKSs.[24]

Comparing the PKS–associated enzymes OocK and LbmC–Ox to a wider range of previously characterized class B flavoprotein monooxygenases (**Figure S25**), BV oxidation activity appears more prevalent in the LmbC–Ox than the OocK clade.[25] To further explore their biocatalytic potential, we tested 37 substrates that do not resemble polyketide intermediates and harbor diverse functional groups (**Figures S27**, **Table S6**). When comparing the free-standing OocK with the excised LbmC–Ox, the latter exhibited greater thermostability and a broad substrate scope, accepting 12 of the 37 test compounds (**Figure S27, Figure 5**). Both enzymes accept NADPH as hydride donor, while OocK also displays activity with NADH (**Table S7**). These results suggest that the two enzymes could be used as biocatalysts for various oxidations, in addition to showing potential as PKS engineering tools to introduce non-standard moieties into polyketides.
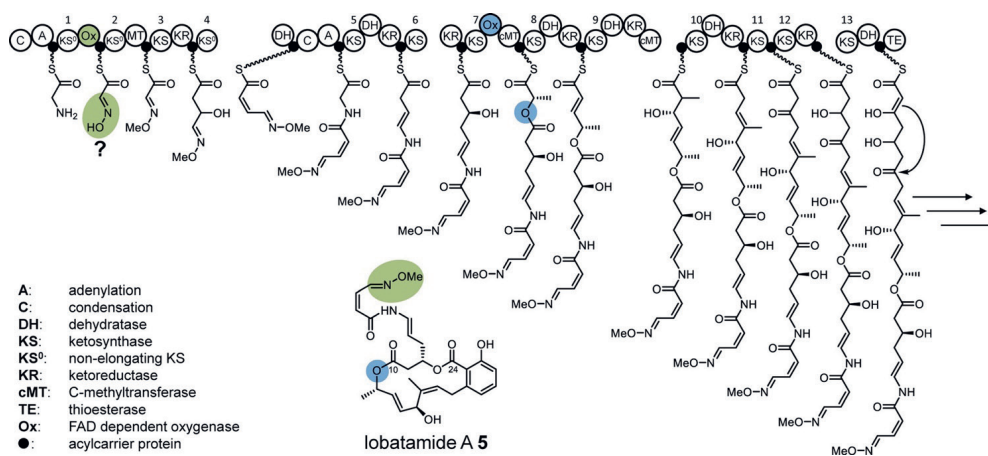
**Figure 4. Biosynthetic model for lobatamide** A. Ox domains and the corresponding polyketide moieties (proposed for the oxime) are highlighted in color. Consecutive KS numbers are shown above the domains.

A:      adenylation
C:      condensation
DH:     dehydratase
KS:     ketosynthase
KS⁰:    non-elongating KS
KR:     ketoreductase
cMT:    C-methyltransferase
TE:     thioesterase
Ox:     FAD dependent oxygenase
●:      acylcarrier protein

lobatamide A **5**



Accepted by
Oock and LbmC-Ox

Only accepted by LbmC-Ox

**18**

**19**

**20**

**21**

**22**

**23**

**24**

**25**

**26**

**27**

**28**

**29**
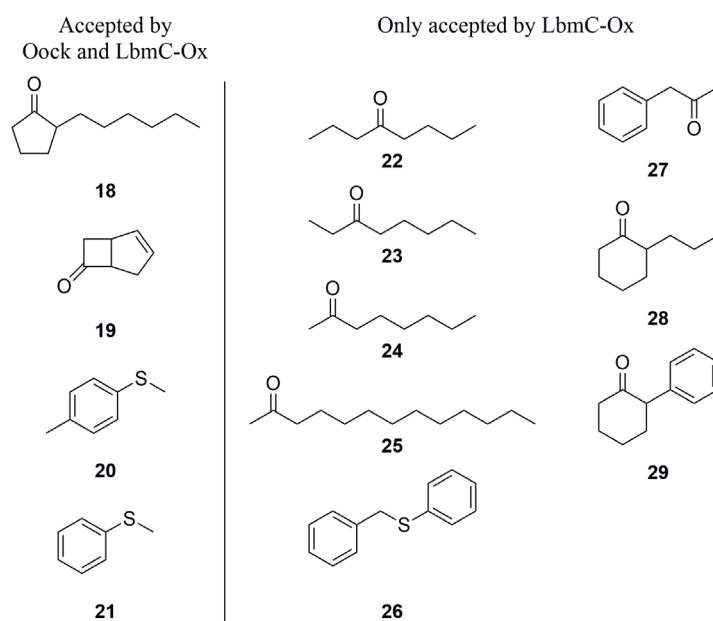
**Figure 5. Structures of compounds that deviate from polyketide intermediate but were accepted by LbmC-Ox and OocK or by LbmC-Ox alone.** For products, see **Figure S27**.

4

## CONCLUSION

In conclusion, our data support the existence of at least two further module types in *trans*–AT PKSs that permit oxidative modifications of polyketide core structures. In addition to the topologically related oxygen insertions catalyzed by modules utilizing trans-acting OocK/PedG/TobD-type enzymes (clade I modules, **Figure S1**), we identified a distinct, module-integrated Ox domain that installs esters of the opposite orientation (clade II modules, **Figure S1**). It is unknown whether all members of one oxygenase clade generate the same ester topology. However, in line with a potentially predictive uniform product pattern for clade II, an uncharacterized oxygenase from the mandelalide *trans*–AT PKS[7c] that is analogous to LmbC–Ox has been suggested to insert oxygen in the same orientation as for lobatamides.[7c] A second type of oxygenase-containing module newly identified in this work is not involved in oxygen insertion but seems to play a role in glycine oxidation to generate the unusual oxime moiety of lobatamides. This hypothesis is currently based on collinearity logic and needs to be experimentally tested. These examples demonstrate the staggering functional diversity of *trans*–AT PKSs, which not only introduce substituents but also modify the carbon backbone itself. Our results also exemplify how genome mining can pave the way to finding microbial sources for bioactive compounds that had previously been described from marine invertebrates. The production of the tunicate metabolite lobatamide A by *G. sunshinyii* further validates this plant symbiont as a rich source of bioactive specialized metabolites.

## MATERIAL AND METHODS

### General

LC-ESI mass spectrometry analyses were performed on a Thermo Scientific Q Exactive mass spectrometer coupled to a Dionex Ultimate 3000 UPLC system. NMR spectra were recorded on a Bruker Avance III spectrometer at 300 MHz or equipped with a cold probe at 500 MHz and 600 MHz for $^1$H NMR and 125 MHz and 150 MHz for $^{13}$C NMR at 298K. Chemical shifts were referenced to the solvent peaks at $\delta_H$ 2.50 and $\delta_C$ 39.51 for DMSO-$d_6$, $\delta_H$ 7.27 and $\delta_C$ 77.23 for chloroform-$d$, $\delta_H$ 3.31 and $\delta_C$ 49.15 for methanol-$d_4$, and $\delta_H$ 1.94 and $\delta_C$ 1.39 for acetonitrile-$d_3$. Sonication was performed on a Sonicator Q700 (QSonica, Newton, USA).

### Prediction of PKS products

The amino acid sequences of the core biosynthetic PKS genes from the *lbm* cluster were used as an input for *TransATor* (https://transator.ethz.ch/).[3] Regions of the structure predicted with low confidence were then manually refined based on co-linearity of domain architectures.

## Phylogenetic analysis of putative core-PKS Baeyer-Villiger monooxygenases

To analyze the putative PKS-associated flavoprotein monooxygenases, amino acid sequences of 70 domains from *trans*-AT PKS modules were selected by BLAST analysis of OocK[9] against Uniprot and internal databases. Genome neighborhood[12] was used to select for sequences associated with PKSs. The sequences were retrieved from the GenBank database and aligned using Geneious 7.1.9 and the MUSCLE algorithm.[26] The phylogenetic reconstruction was performed with Geneious Tree Builder, employing the NJ algorithm. Bootstrap analysis was performed with 100 pseudo-replicate sequences.

## Extraction and isolation of lobatamide A (5)

*Gynuella sunshinyii* YC6258 was obtained from NITE Biological Resource Center (NBRC). *G. sunshinyii* was cultured in 2 L PH-103 medium[27] at 30 °C for 3 days on an orbital shaker. The culture was centrifuged, the supernatant was extracted three times with EtOAc, and the pellet was suspended in $dH_2O$ and extracted with acetone. The pellet extract was dried and separated by RP-HPLC (Phenomenex Luna 5μ C18, φ 20 x 250 mm, 10.0 mL/min, 260 nm) with a gradient elution from 5% acetonitrile to 100% acetonitrile+0.1% formic acid to afford 20 fractions. Fraction 9 and 10 were further separated by RP-HPLC (Phenomenex Luna 5μ Phenyl-Hexyl, φ 10 x 250 mm, 2.0 mL/min, 200 nm) with 40% acetonitrile+0.1% formic acid and then separated by RP-HPLC (Phenomenex Synergi 4μ Hydro-RP, φ 10 x 250 mm, 2.0 mL/min, 200 nm) with 40% acetonitrile+0.1% formic acid to yield lobatamide A (5). HRESIMS *m/z* 535.2048 [M+Na]$^+$ (calcd. for $C_{27}H_{32}N_2O_8Na$, 535.2051) (**Figure S5**). NMR spectra matched to the reported values for lobatamide A (**Figure S6-9**).[18]

## Construction of LbmC-Ox expression plasmids

The sequence encoding LbmC-Ox was amplified from *G. sunshinyii* liquid culture using the primer pair *lbmC-Ox*_NdeI_F (CGA TCA TAT GGG CGG TCG CCA GCC ACA G) and *lbmC-Ox*_NS-NotI_R (CAT GCG GCC GCT CTG CTT TCA GGC TCG ATG AG). The gel purified fragment was digested with NdeI and NotI-HF and cloned into the pET28-derived vector with a modified tag region for C-terminal fusion, namely pET28HB-TS, yielding pET28HB-*TSlbmC-Ox*, which was transformed into *E.coli* DH5α. The plasmid was isolated and introduced into *E. coli TunerTM* (DE3) competent cells (Novagen). This strain was used for the expression of N-terminally His6-tagged *LbmC-Ox*.

## Heterologous gene expression and protein purification of LbmC-Ox

*E. coli* expression strain was inoculated in LB medium containing 50 μg/mL kanamycin at 37 °C overnight and then inoculated 1:100 into TB medium with 10% glycerol and 50 μg/mL kanamycin, and cultured at 37 °C until it reached an $OD_{600}$ of 0.5, after which the culture was cooled on ice for 15 min. Gene overexpression was induced by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) at a concentration of 0.1 mM. The

induced culture was grown for additional 16-20 h at 16 °C. The culture was harvested by centrifugation and cell pellets were either processed directly or frozen and stored at -80 °C. All purification steps were carried out at 4 °C. Cells were resuspended in lysis buffer (50 mM phosphate buffer, 300 mM NaCl, 20 mM imidazole, and 10% [v/v] glycerol) and disrupted by sonication using a Sonicator Q700 (QSonica, Newton, USA). The lysate was centrifuged for 20 min at 18,000 x g at 4 °C. The supernatant was incubated with Ni-NTA agarose (Macherey-Nagel, Oensingen, Switzerland) for 10 min at 4 °C, and then transferred to a fretted column. The resin was washed once with 3 mL lysis buffer, and then eluted with 750 µL elution buffer (50 mM phosphate buffer, 300 mM NaCl, 250 mM imidazole, and 10% glycerol at pH 8.0). Yellow elution fractions were verified to contain the eluted protein by SDS-PAGE.

### Determination of melting temperatures

The apparent melting temperatures ($T_M^{app}$) values for purified LbmC-Ox and OocK were determined using the ThermoFAD method as previously described.[29] The samples were prepared in a 96-well PCR plate at 2 mg/mL of enzyme in different buffered solutions: 50 mM Bis-Tris HCl, 50 mM Tris-HCl, or 50 mM CHES buffer, adjusted at desired pH and concentration of 1,4-dioxane. Using an RT-PCR instrument (CFX96-Touch, Bio-Rad), these samples were subjected to a temperature gradient from 20 to 95 °C, increasing by 0.5 °C every 10 s. The fluorescence signal is observed when the FAD is exposed to the solvent during the unfolding process. The fluorescence intensity was measured using a 450–490 nm excitation filter and a 515–530 nm emission filter. The $T_M^{app}$ was determined as the maximum of the derivative of the sigmoidal curve of fluorescence intensity against temperature.

### Conversion of ketones and sulfides

The biocatalytic studies were performed using 7 mixtures that contained several ketones and sulfides (**Table S6**). The reactions were prepared in 200 µL in 2 mL glass vials. These samples were prepared at 10% glycerol, 200 mM NaCl, 20 µM FAD, 10 mM $Na_2PO_3 \cdot 5H_2O$, 150 µM NADPH, 5 µM of PTDH, 5 µM of purified OocK or LbmC-Ox and 2.5% v/v of 1,4-dioxane as a co-solvent in 50 mM TrisHCl buffer pH 8.0. The final concentration of each substrate in the substrate scope analysis was 400 µM. The reactions were incubated 24 h at 24 °C and 135 r.p.m. for LbmC-Ox and 24 h at 17 °C and 135 r.p.m. for OocK. For the latter, further conversion of the accepted substrates were tested individually using either NADH or NADPH at 150 µM. The final concentration of each substrate was 1 mM (with incubations at 17 and 24 °C but obtaining statistically similar results). After the incubations, the mixtures were extracted mixing the solution with one volume of ethyl acetate containing 0.1% v/v mesitylene (as external standard), vortexed for 1 min, and centrifuged for 5 min at 13,000 x g. The organic layer was kept, and the aqueous phase was extracted two more times following the same procedure. Anhydrous sulfate magnesium

was added to the organic phase to remove residual water. The solutions were vortexed for 1 min and the supernatant obtained after centrifugation (5 min, 13,000 × g) was analyzed in a GCMS-QP2010 Ultra (Shimadzu) with electron ionization and quadrupole separation using an HP-1 column (HP-1 Agilent, 30 m x 0.25 mm x 0.25 μm). The GC program for the substrate scope analysis was: 30 °C x 5 min, 5 °C min$^{-1}$ until 70 °C; hold 5 min, 5 °C min$^{-1}$ until 130 °C; hold 5 min and 10 °C min$^{-1}$ until 325 °C; hold 5 min (using a split ratio of 5.0, and 2 μL injected).

**4**

# SUPPORTING INFORMATION CHAPTER IV

## Synthesis of substrates

### S-(2-acetamidoethyl) 3-oxoheptanethioate; β-oxoheptanoyl-SNAC (13)

Synthesis was performed according to a previously published procedure.[9]

### S-(2-acetamidoethyl) 2-methyl-3-oxoheptanethioate; methyl-beta-oxoheptanoyl-SNAC (16)

Synthesis was performed according to a previously published procedure.[32] The products were purified by RP-HPLC (Phenomenex Luna 5μ C18, φ 10 x 250 mm, 2.0 mL/min, 230 nm) with a gradient elution from 5% acetonitrile to 100% acetonitrile + 0.1% formic acid to afford **14** and **16** (**Figure 4A**).

### *In vitro* enzyme activity assay with SNACs

1μL of the synthesized SNAC substrate (80 mM in DMSO) was added to 96 μL of the phosphate buffer (200 mM phosphate buffer, pH 7.0) and 1 μL of NADPH (100 mM in the phosphate buffer). After vortex and spin down the sample, 2 μL of the purified LbmC-Ox was added and incubated overnight at room temperature. The sample was extracted with EtOAc, dried and resuspended in MeOH for LCMS.

## Preparation of 15 for NMR measurement

The biochemical assay with 13 was scaled up to 10 mL in total. The extracts were purified by RP-HPLC (Phenomenex Luna 5μ C8, φ 10 × 250 mm, 2.0 mL/min, 230 nm) with 25% acetonitrile + 0.1% formic acid to yield **15**. HRESIMS $m/z$ 262.1112 [M+H]$^+$ (calcd. for $C_{11}H_20N_1O_4S_1$, 262.1108) (**Figure S11**). $^1$H NMR (acetonitrile-$d_3$) and $^{13}$C NMR (acetonitrile-$d_3$) data, see Figure **S13-17** and **Table S4**.

## Preparation of 17 for NMR measurement

The biochemical assay with 16 was scaled up to 10 mL in total. The extracts were purified by RP-HPLC (Phenomenex Luna 5μ C8, φ 10 x 250 mm, 2.0 mL/min, 230 nm) with 35% acetonitrile + 0.1% formic acid to yield 17. HRESIMS $m/z$ 276.1266 [M+H]$^+$ (calcd for $C_{12}H_{22}N_1O_4S_1$, 276.1264) (**Figure S19**). $^1$H NMR (acetonitrile-$d_3$) and $^{13}$C NMR (acetonitrile-$d_3$) data, see **Figure S20-24** and **Table S5**.

## Tables

**Table S1. ORFs detected on the *Gynuella sunshinyii* lobatamide (*lbm*) locus and their putative functions**.

| ORF | Protein size | Proposed function | Closest homolog | Identity [%] | Accession number |
|---|---|---|---|---|---|
| LbmF *WP_044617910* | 327 | Fumarylacetoacetate hydrolase family protein | Fumarylacetoacetate hydrolase family protein, *Marinobacterium* sp. AK27 | 69 | WP_036188863 |
| LbmG *WP_044617909* | 323 | NAD-dependent epimerase/dehydratase family protein | NAD-dependent epimerase/dehydratase family protein, *Marinomonas* sp. MWYL1 | 52 | WP_012070817 |
| LbmH *WP_044617908* | 290 | Helix-turn-helix domaincontaining protein | AraC family transcriptional regulator, *Shewanella japonica* | 61 | WP_080917294 |
| LbmI *WP_044617907* | 264 | Hypothetical protein (SnoaL domains) | Hypothetical protein, *Aliivibrio logei* | 78 | WP_017021136 |
| LbmJ *WP_052830346* | 298 | Protein-ADP-ribose hydrolase | Protein-ADP-ribose hydrolase, *Enterovibrio* sp. A649 | 52 | WP_129494894 |
| LbmK *WP_044617906* | 284 | Hypothetical protein | NAD-dependent protein deacetylase of SIR2 family, *Enterovibrio* sp.A649 | 63 | WP_129494893 |
| LbmL *WP_044617905* | 153 | GNAT family Nacetyltransferase | GNAT family Nacetyltransferase, *Photobacterium halotolerans* | 41 | WP_046221096 |
| LbmM *WP_044617904* | 175 | Hypothetical Protein | Adenylate kinase, *Saccharospirillum* sp. LZ-5 | 49 | WP_127556471 |
| LbmN *WP_044617903* | 130 | VOC family protein | VOC family protein, *Microbulbifer* sp. HB161719 | 82 | WP_138236387 |
| LbmO *WP_044617902* | 143 | Acetyltransferase | Acetyltransferase *Marinomonas* sp. MWYL1 | 89 | WP_012070635 |
| LbmA *WP_044617900* | 4800 | PKS-NRPS (A PCP KS0 Ox ACP KS0 MT ACP KS KR KS0) | Amino acid adenylation domain-containing protein, *Rhizobium sp.* BK315 | 59 | *WP_132721172* |
| LbmB *WP_052830344* | 3598 | PKS-NRPS (DH ACP C A PCP KS DH KR ACP KS) | Amino acid adenylation domain-containing protein, *Rhizobium* sp. BK315 | 53 | *WP_132721174* |
| LbmC *WP_044617899* | 5591 | PKS (KR ACP KS Ox cMT ACP KS DH KR ACP KS DH KR cMT) | SDR family NAD(P)-dependent oxidoreductase, *Rhizobium* sp. BK315 | 59 | WP_132721176 |
| LbmD *WP_052830343* | 3643 | PKS (ACP KS DH KR ACP KS ACP KS KR ACP) | SDR family NAD(P)-dependent oxidoreductase, *Rhizobium* sp. BK315 | 57 | WP_132721178 |
| LbmE *WP_052830342* | 1154 | PKS (KS DH ACP TE) | Polyketide synthaselike dehydratase family protein, *Rhizobium* sp. BK315 | 59 | WP_132721180 |
| LbmP *WP_044617898* | 292 | LysR family transcriptional regulator | LysR family transcriptional regulator, *Mesorhizobium* sp. URHC0008 | 49 | WP_027040968 |
| LbmQ *WP_044617896* | 312 | DMT family transporter | DMT family transporter *Rhodoligotrophos* sp. lm1 | 60 | WP_137391615 |
| LbmR *AJQ95670* | 154 | Methyl-accepting chemotaxis protein | Methyl-accepting chemotaxis protein, *Psychromonas ossibalaenae* | 53 | WP_019615169 |

**Table S2.** Predicted KS substrates for the *lbm* locus and actually accepted groups in lobatamide biosynthesis.

| KS Domain | TransATor Prediction | Manually Refined Prediction | Actual accepted group |
|---|---|---|---|
| LbmA | Non-elongating (reduced/shifted double bond) | KS1 Non-elongating (Glycine) | Glycine |
| LbmA KS2 | Non-elongating (β-OH) | Non-elongating (Glycine) | Oxime |
| LbmA KS3 | Exomethyl/exoester | Glycine | Oxime |
| LbmA KS4 | Non-elongating (bimodule D-β-OH) | Non-elongating (bimodule D -β-OH) | Non-elongating (bimodule D-β-OH) |
| LbmB KS1 | Amino acids | Amino acid-Glycine | Glycine |
| LbmB KS2 | Reduced | Reduced | Shifted double bond |
| LbmC KS1 | L-β-OH | L-β-OH | L-β-OH |
| LbmC KS2 | α-Methyl + reduced/keto/D-β- OH | α-Methyl + reduced | α-Methyl, Oxygen inserted |
| LbmC | Double bonds (*E*-configured) | KS3 Double bonds (*E*-configured) | Double bond |
| LbmD KS1 | α-L-(di)Methyl + β-OH | α-Methyl + β-OH | α-Methyl + β-OH |
| LbmD | KS2 Shifted double bonds (some with α-methyl) | Shifted double bonds (some with α-methyl) | Shifted double bond |
| LbmD KS3 | β-keto or double bonds | Double bond | Keto |
| LbmE KS1 | *Cis*-AT/Starters or β-OH | β-OH | β-OH |

**Table S3. Chemical shifts of lobatamide A (5) from *G. sunshinyii* in CD$_3$OD.**
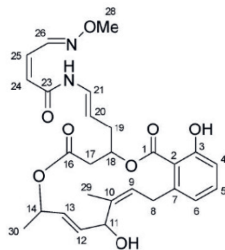


| No. | $\delta_C$ | δH, mult |
|-----|------|----------|
| 1 | 170.0 | |
| 2 | 122.1 | |
| 3 | 156.6 | |
| 4 | 114.2 | 6.69 brd |
| 5 | 131.7 | 7.15 (dd, 8.0) |
| 6 | 120.8 | 6.65 brd |
| 7 | 141.2 | |
| 8 | 33.0 | 2.95 brd<br>3.23 ovlp. |
| 9 | 125.5 | 5.18 m |
| 10 | 139.3 | |
| 11 | 73.2 | 4.79 (d, 8.6) |
| 12 | 134.9 | 5.69 (dd, 8.6, 15.2) |
| 13 | 132.6 | 5.50 (dd, 8.6, 15.2) |
| 14 | 73.7 | 5.24 (dq, 6.7, 8.6) |
| 16 | 171.7 | |
| 17 | 38.8 | 2.60 (dd, 10.7, 16.7)<br>2.70 (dd, 2.1, 16.7) |
| 18 | 72.9 | 5.60 m |
| 19 | 35.5 | 2.49 brdd |
| 20 | 109.8 | 5.34 (dt, 7.7, 14.3) |
| 21 | 126.8 | 6.83 (d, 14.3) |
| 23 | 164.2 | |
| 24 | 126.1 | 6.06 (dd, 1.1, 11.5) |
| 25 | 135.5 | 6.50 (dd, 10.3, 11.5) |
| 26 | 148.6 | 8.96 (dd, 1.1, 10.3) |
| 28 | 62.6 | 3.90 s |
| 29 | 19.5 | 1.80 brs |
| 30 | 20.1 | 1.36 (d, 6.6) |

4

**Table S4.** Chemical shifts of **15** in acetonitrile-$d_3$.

| No. | δC | δH, mult |
|-----|------|----------------------|
| 1 | 196.6 | |
| 2 | 68.4 | 4.73 s |
| 3 | 173.7 | |
| 4 | 34.0 | 2.42 (t, 7.5) |
| 5 | 27.4 | 1.62 m |
| 6 | 22.7 | 1.37 m |
| 7 | 14.0 | 0.92 (t, 7.4) |
| 8 | 28.5 | 3.01 (t, 6.6) |
| 9 | 39.3 | 3.28 (dt, 6.3, 6.6) |
| 9-NH | | 6.49 br |
| 10 | 170.8 | |
| 11 | 22.9 | 1.82 s |

**Table S5.** Chemical shifts of **17** in acetonitrile-*d*3.

| No. | $δ_c$ | δH, mult |
|-----|------|----------------|
| 1 | 200.4 | |
| 2 | 75.6 | 5.20 (q, 7.0) |
| 3 | 173.5 | |
| 4 | 34.2 | 2.40 (t, 7.4) |
| 5 | 27.5 | 1.61 m |
| 6 | 22.7 | 1.37 m |
| 7 | 13.9 | 0.92 (t, 7.4) |
| 8 | 28.5 | 2.97 m |
| 9 | 39.2 | 3.26 m |
| 9-NH | | 6.47 br |
| 10 | 170.7 | |
| 11 | 22.9 | 1.82 s |
| 12 | 18.2 | 1.42 (d, 7.0) |

**Table S6. Substrate scope analysis of OocK and LmbC-Ox.** ''+'' for <30% conversion, ''++'' for conversion between 30-80% and ''+++'' for 80-100% conversion

| Mixture | Compound | CAS | Structure | Oock | LmbC-Ox |
|---|---|---|---|---|---|
| 1 | benzyl phenyl sulfide | 831-91-4 | | | +++ |
| | phenylacetone | 103-79-7 | | | + |
| | 2-hexylcyclopentanone | 13074-65-2 | | + | +++ |
| | indole | 120-72-9 | | | |
| | cycloundecanone | 878-13-7 | | | |
| 2 | cyclopentanone | 120-92-3 | | | |
| | 3-octanone | 106-68-3 | | | +++ |
| | bicycle[3.2.0]hept-2-en-6-one | 13173-09-6 | | ++ | +++ |
| | 2-propylcyclohexanone | 94-65-5 | | | ++ |
| | cyclododecanone | 830-13-7 | | | |
| | methyl-p-tolyl sulfide | 1519-39-7 | | ++ | +++ |
| 3 | cyclopentadecanone | 502-72-7 | | | |
| | 2-phenylcyclohexanone | 1444-65-1 | | | ++ |
| | ethionamide | 536-33-4 | | | |
| | vanillyl acetone | 122-48-5 | | | |
| | nicotin | 54-11-5 | | | |
| | androst-4-ene-3,17-dione | 63-05-8 | | | |
| 4 | styrene | 100-42-5 | | | |
| | 4-phenylcyclohexanone | 4894-75-1 | | | |
| | 4-hydroxyacetophenone | 99-93-4 | | | |
| | 1,4-androstadiene-3,17-dione | 897-,6-3 | | | |

| | | | | | |
|---|---|---|---|---|---|
| 5 | cyclohexanone | 108-94-1 | | | |
| | 4-octanone | 589-63-9 | | | +++ |
| | benzyl methyl sulfide | 100-68-5 | | +++ | +++ |
| | acetophenone | 98-86-2 | | | |
| | 2-dodecanone | 6175-49-1 | | | +++ |
| | pregnenolone | 145-13-1 | | | |
| 6 | 4-methylcyclohexanone | 589-92-4 | | | |
| | 2-octanone | 111-13-7 | | | ++ |
| | cyclooctanone | 502-49-8 | | | |
| | benzoin | 11953-9 | | | |
| | stanolone | 521-18-6 | | | |
| | thiacetazone | 104-06-3 | | | |
| 7 | 3,3,5-trimethylcyclohexanone | 873-94-9 | | | |
| | methylvinylketone | 78-94-4 | | | |
| | phenyl methyl sulfide | 766-92-7 | | | |
| | isophorone | 78-59-1 | | | |
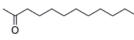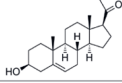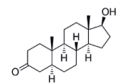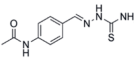
**Table S7.** Effect of NAD(P)H in conversions using OocK

| Substrate[a] | Conversion [%] NADH [b] | Conversion [%] NADPH [c] |
|---|---|---|
| methyl phenyl sulfide | 25 | 95 |
| methyl-p-tolyl sulfide | 5 | 60 |
| bicyclo[3.2.0]hept-2-en-6-one | 5 | 70 |
| 2-hexylcyclopentanone | 5 | 35 |

[a] The final concentration of each substrate was 1 mM. The reactions samples were prepared in 50 mM TrisHCl buffer pH 8.0, 10% glycerol, 200 mM NaCl, 20 μM FAD, 10 mM $Na_2PO_3.5H_2O$, 5 μM of PTDH, 5 μM of purified OocK, 2.5% v/v of 1,4-dioxane and 150 μM NADH[b] or NADPH[c]

# Figures
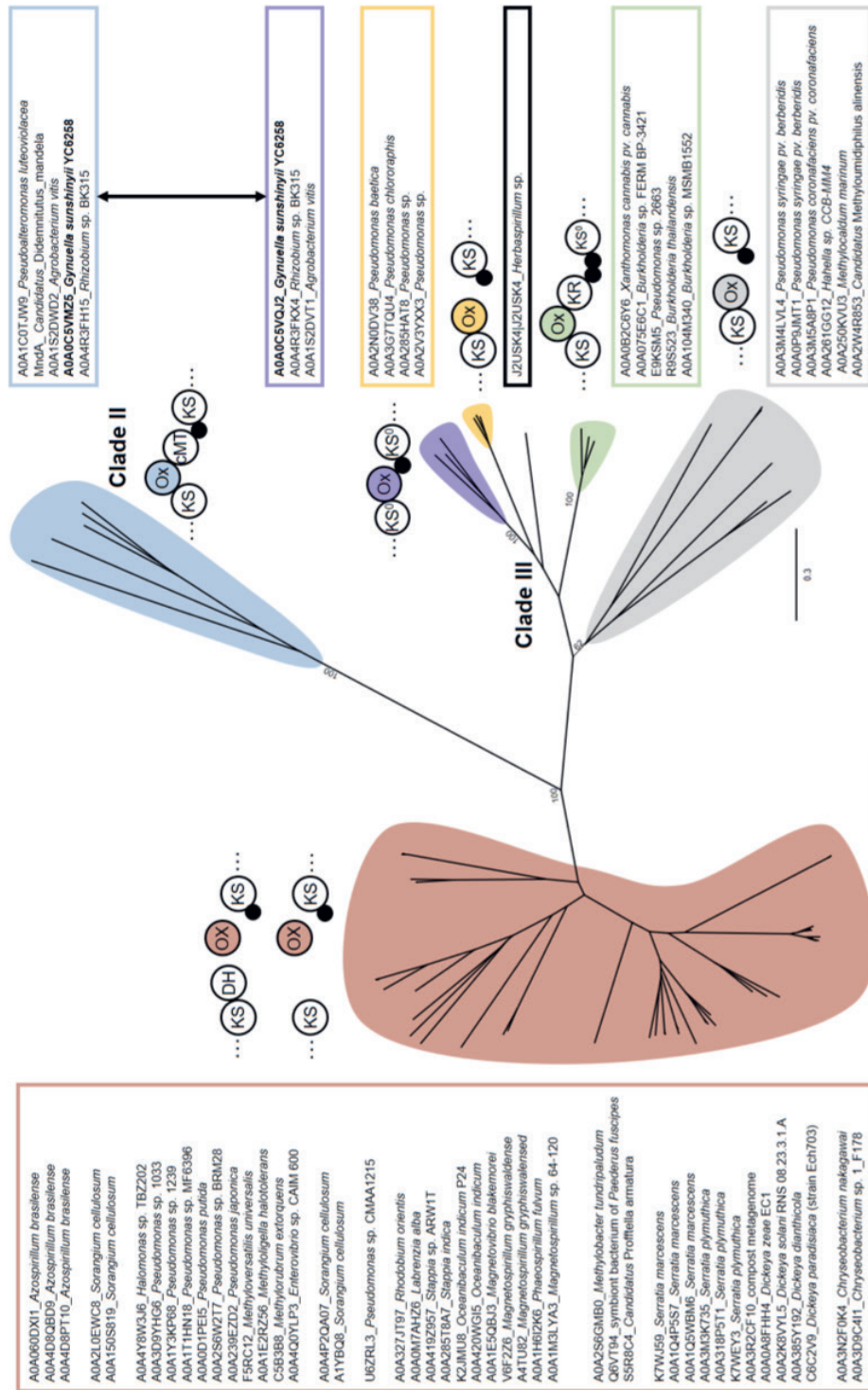


**Figure S1. Phylogram of various flavin adenine dinucleotide (FAD)-dependent monooxygenases associated with *trans*-AT PKS pathways.** Colored boxes next to tip labels consist of the protein accession number and organism. Representative schemes of the oxygenase domains within modular context are depicted for several clades with the corresponding color denoting the oxygenase domains.
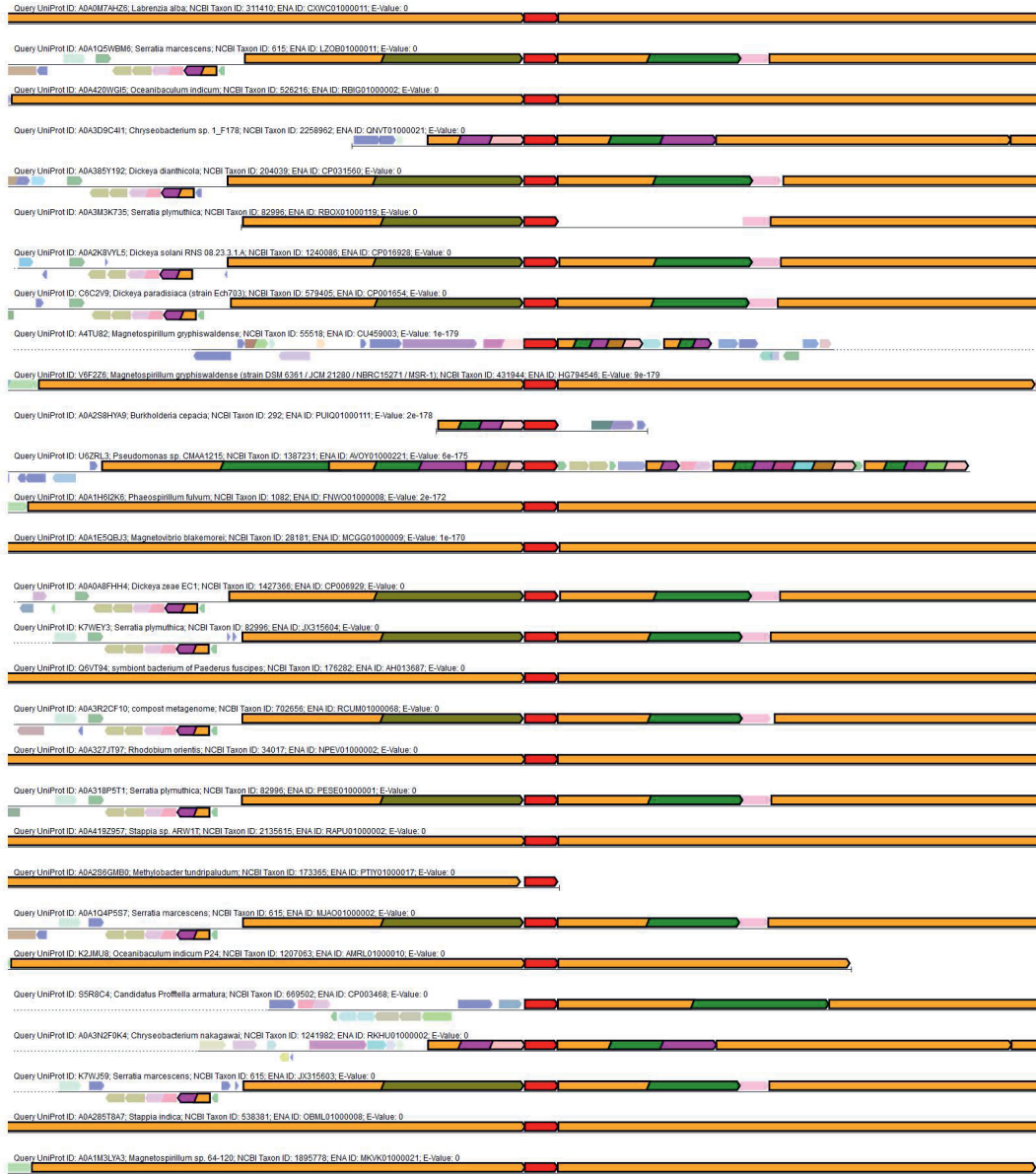
**Figure S2. Output from the Enzyme Function Initiative-Genome Neighborhood Tool (EFI-GNT).**[12] The query sequence was the OocK amino acid sequence. The top hits presenting both a protein family of FMO-like (PF00732) and ketoacyl-synthase (PF00109) are shown.

**Figure S2.** Continued.

**Figure S3. Output from EFI-GNT.**[12] The query sequence was the LbmA-Ox amino acid sequence. The top hits presenting both a protein family of "FMO-like" (PF00732) and "ketoacyl-synthase" (PF00109) are shown.



**Figure S4. Output from EFI-GNT.**[12] The query sequence was the LbmC-Ox amino acid sequence. The top hits presenting both a protein family of "FMO-like" (PF00732) and "ketoacyl-synthase" (PF00109) are shown.
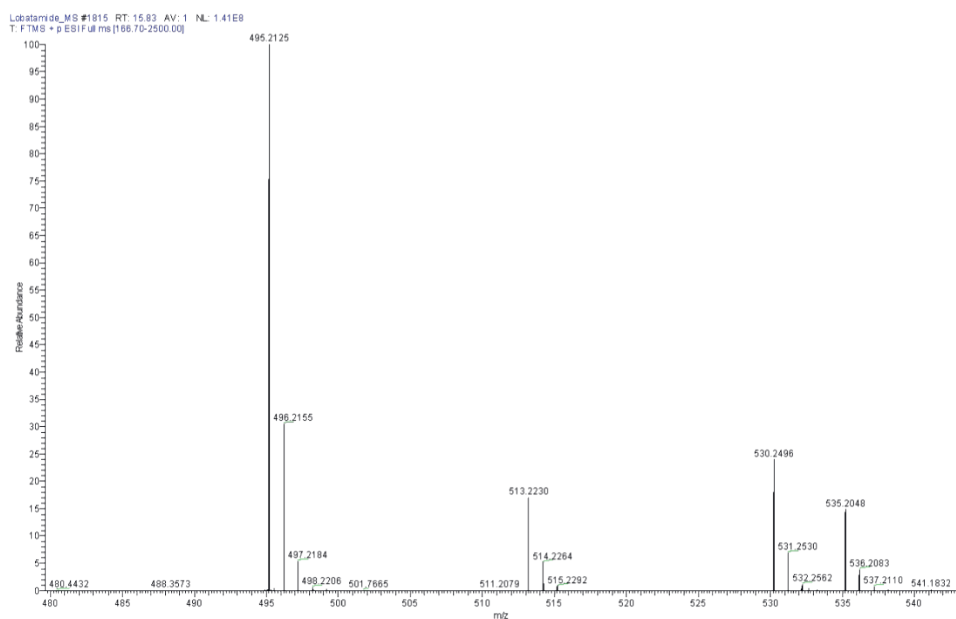
**Figure S5. HR-LC-ESIMS data of lobatamide A (5).** Mass spectrum of the HPLC peak at 15.83 min. The major peaks were the following: $m/z$ 495.2125 [M+H-$H_2O$]$^+$, $m/z$ 513.2230 [M+H]$^+$, $m/z$ 530.2496 [M+NH4]+ and $m/z$ 535.2048 [M+Na]$^+$.
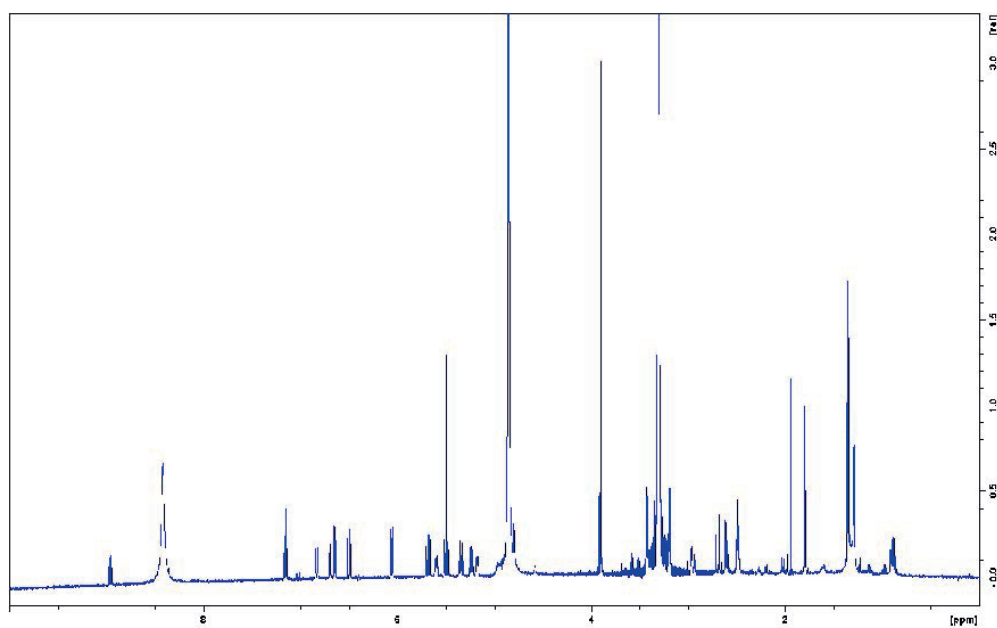


**Figure S6.** $^1$H NMR spectrum of lobatamide A (5) from *G. sunshinyii* in CD$_3$OD at 600 MHz.
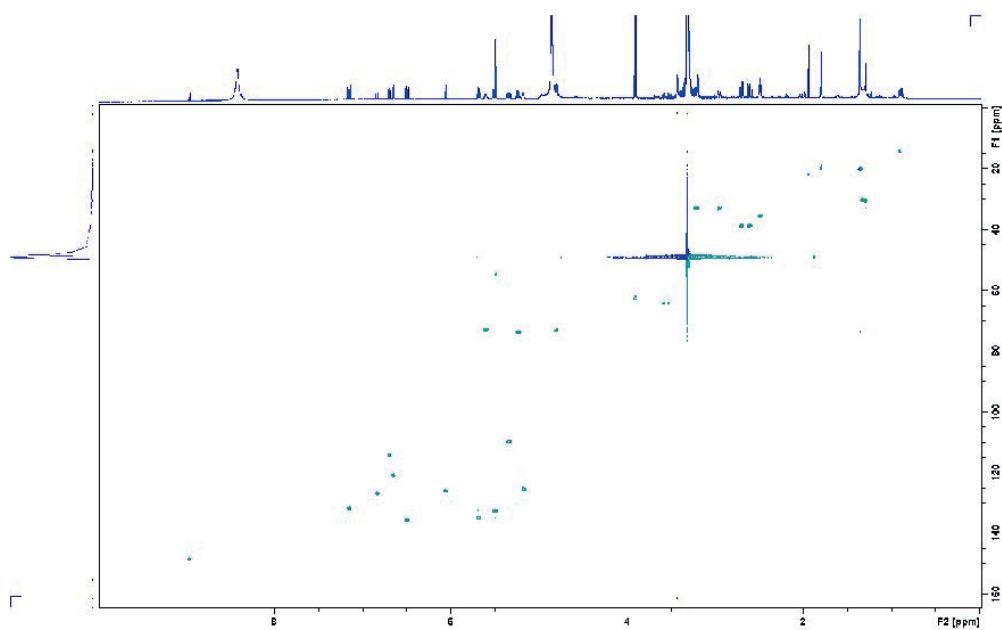
**Figure S7.** HSQC spectrum of lobatamide A (5) from *G. sunshinyii* in CD₃OD at 600 MHz
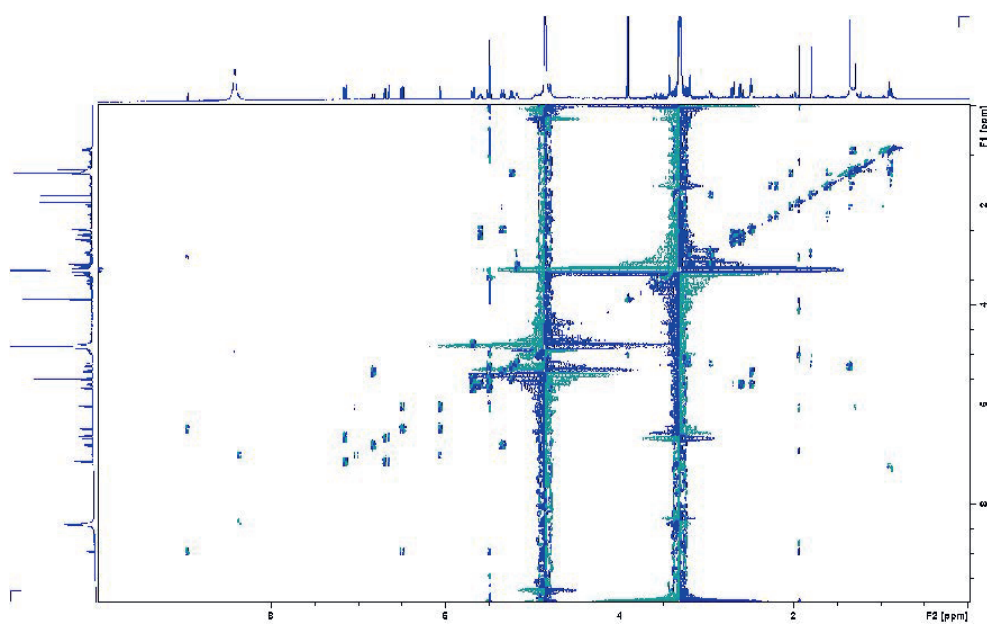


**Figure S8.** COSY spectrum of lobatamide A (5) from *G. sunshinyii* in CD₃OD at 600 MHz.
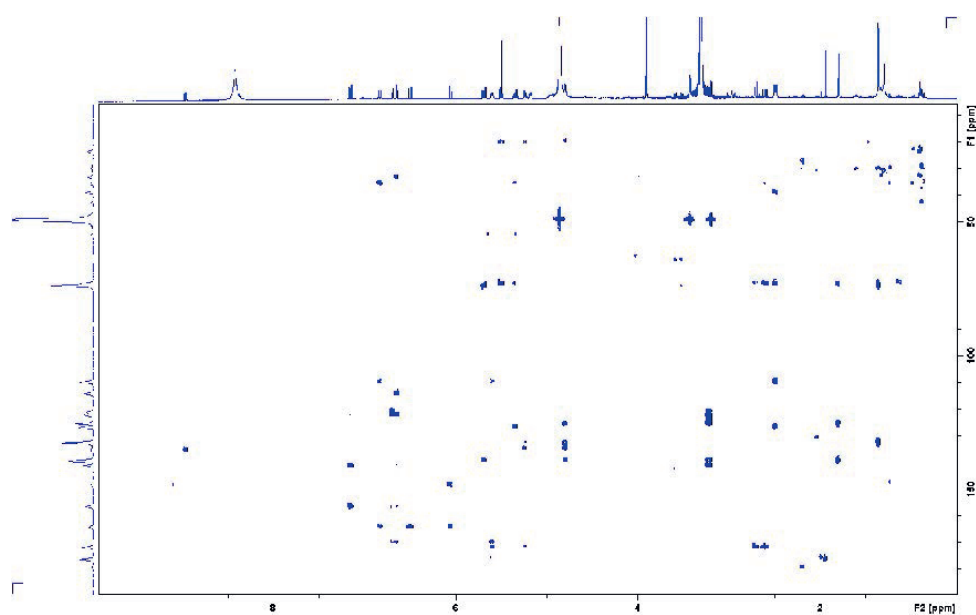
**Figure S9.** HMBC spectrum of lobatamide A (5) from *G. sunshinyii* in CD$_3$OD at 600 MHz.
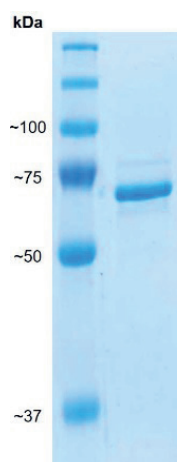


**Figure S10.** 12% SDS-PAGE gel of LbmC-Ox (expected molecular weight: 67.6 kDa) post His6-tag purification
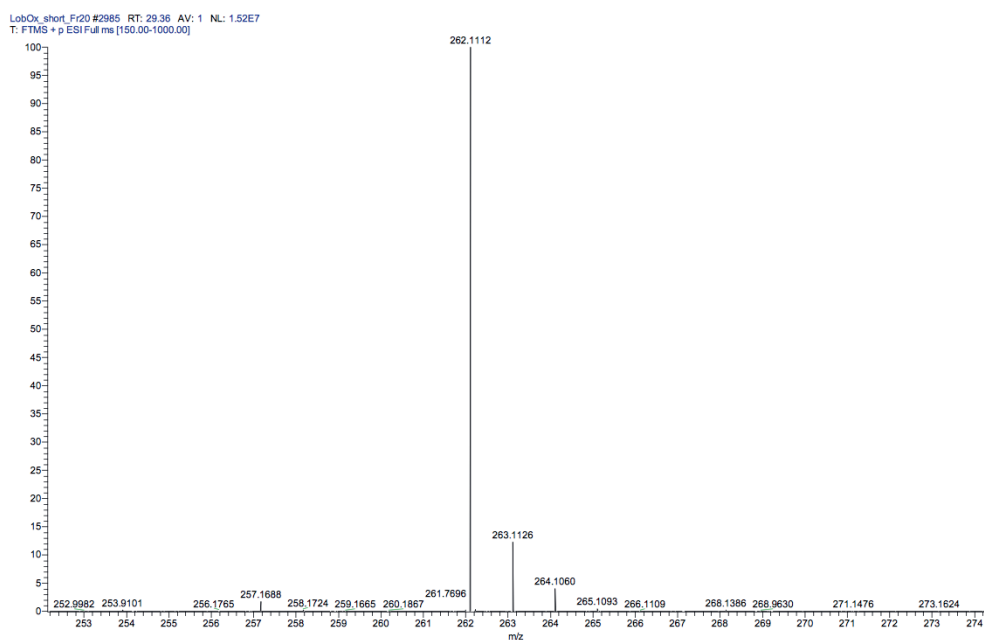
**Figure S11:** **HR-LC-ESIMS data of 15 (*m/z* 262.1112 [M+H]+).** Mass spectrum of the peak at 29.36 min.
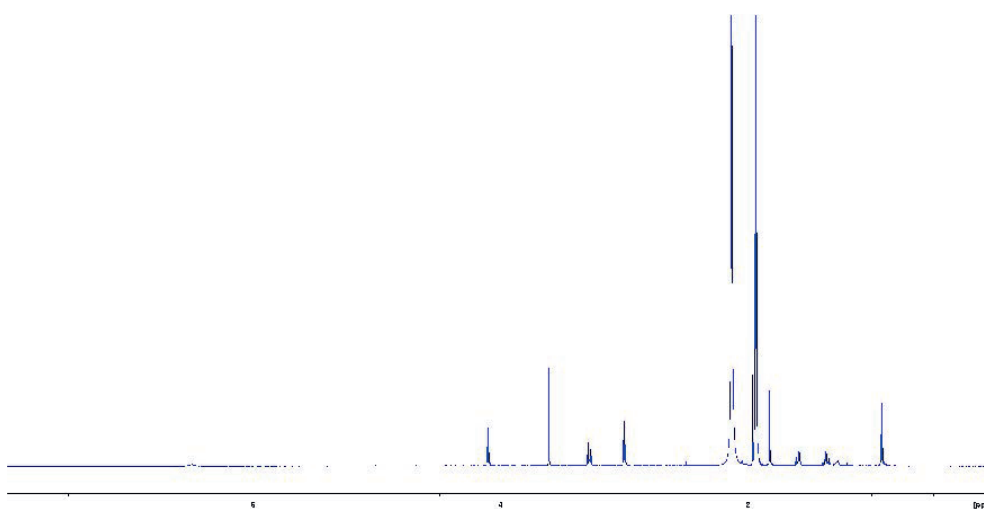


**Figure S12:** **¹H NMR spectrum of 14 acetonitrile-*d*₃ at 600 MHz**. The spectrum matches to the reported values for **14** that were described in previous work.[9]
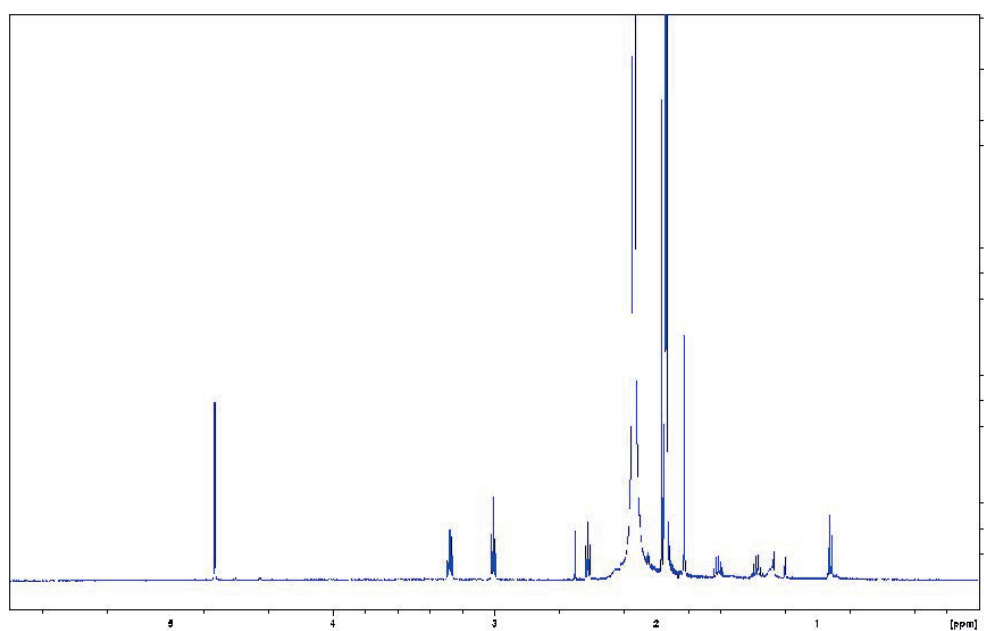
**Figure S13.** ¹H NMR spectrum of 15 in acetonitrile-$d_3$ at 600 MHz.
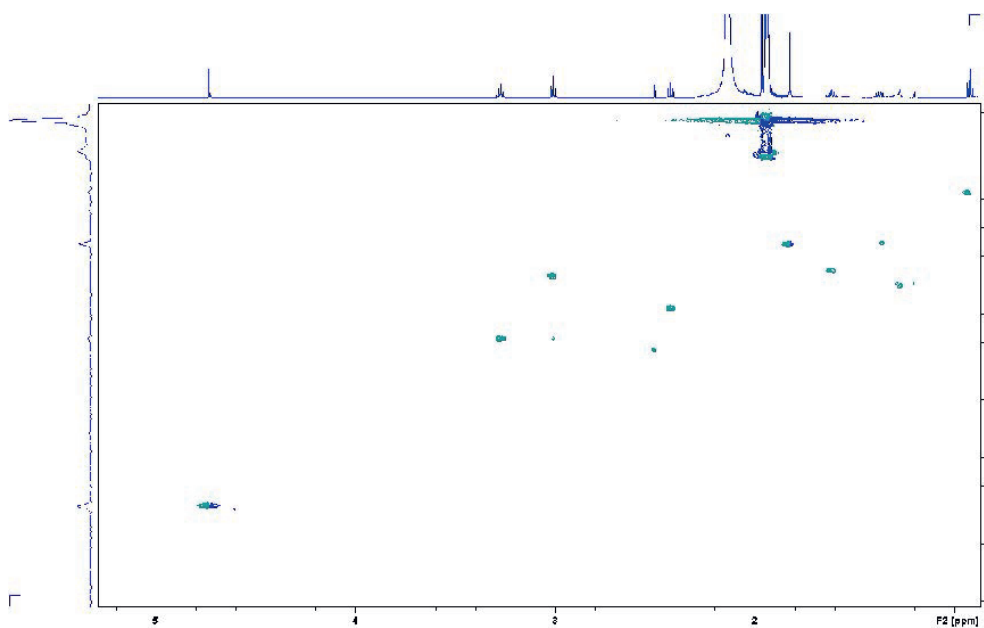


**Figure S14.** HSQC spectrum of 15 in acetonitrile-$d_3$ at 600 MHz.
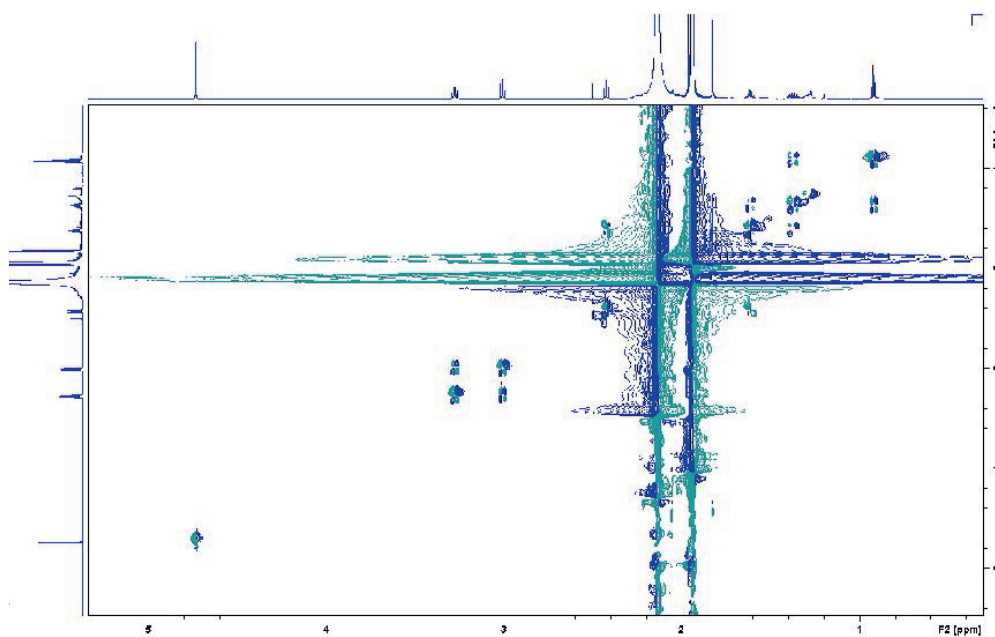
**Figure S15.** COSY spectrum of 15 in acetonitrile-$d_3$ at 600 MHz.
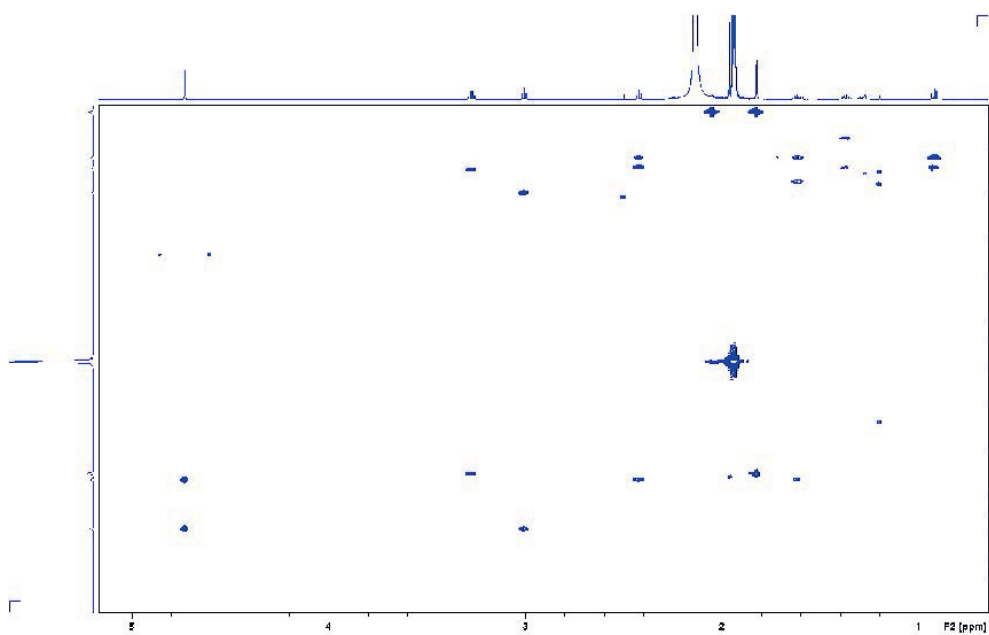


**Figure S16.** HMBC spectrum of 15 in acetonitrile-$d_3$ at 600 MHz
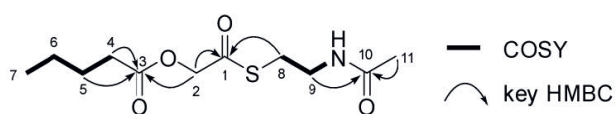
**Figure S17.** **Structure elucidation of 15. COSY and key HMBC correlations**
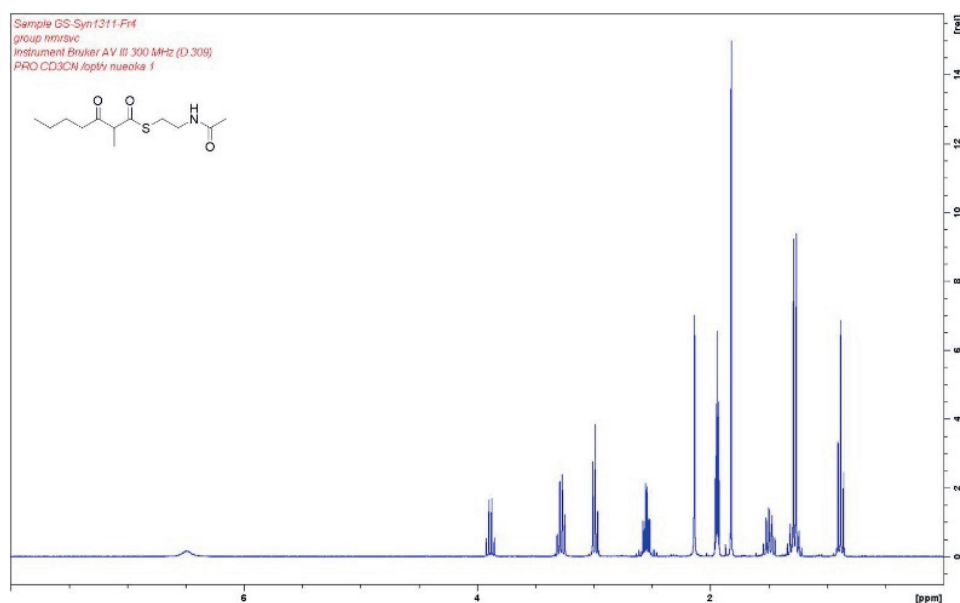

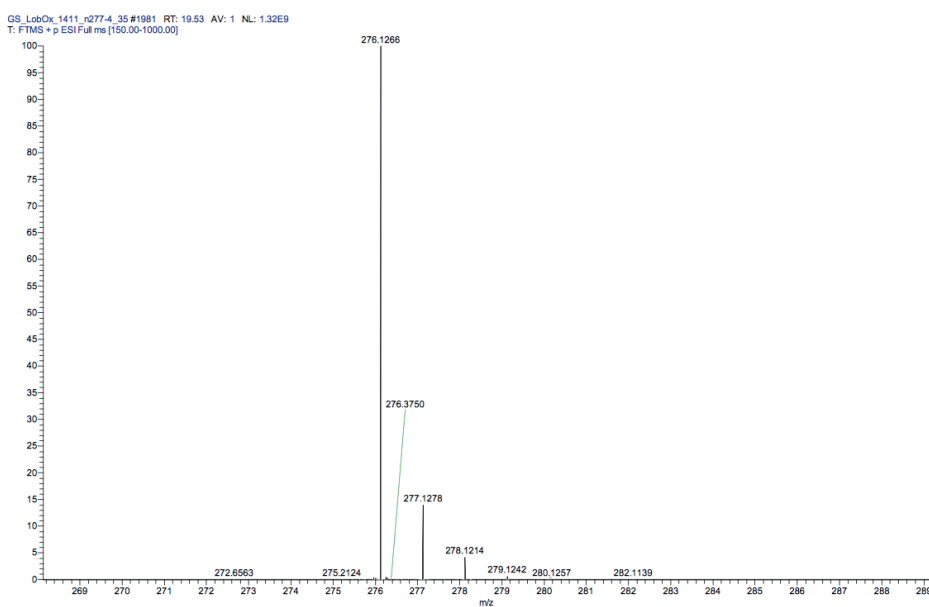
**Figure S18.** $^1$H NMR spectrum of **16** in acetonitrile-$d_3$ at 300 MHz.



**Figure S19:** **HR-LC-ESIMS data of 17 ($m/z$ 276.1266 [M+H]$^+$).** Mass spectrum of the HPLC peak at 19.53 min.
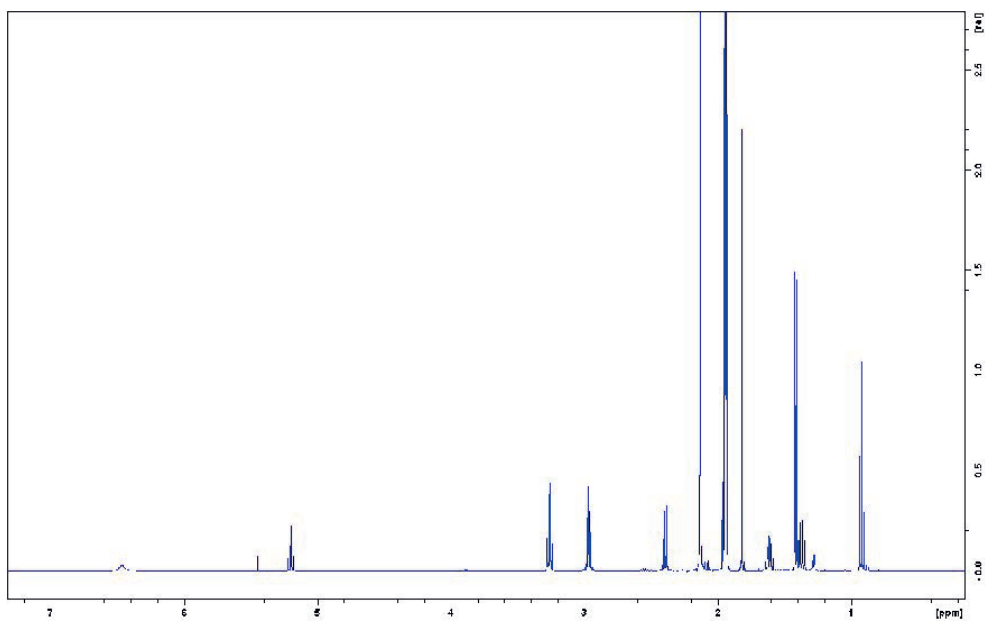
**Figure S20.** $^{1}$H NMR spectrum of **17** in acetonitrile-$d_3$ at 500 MHz.
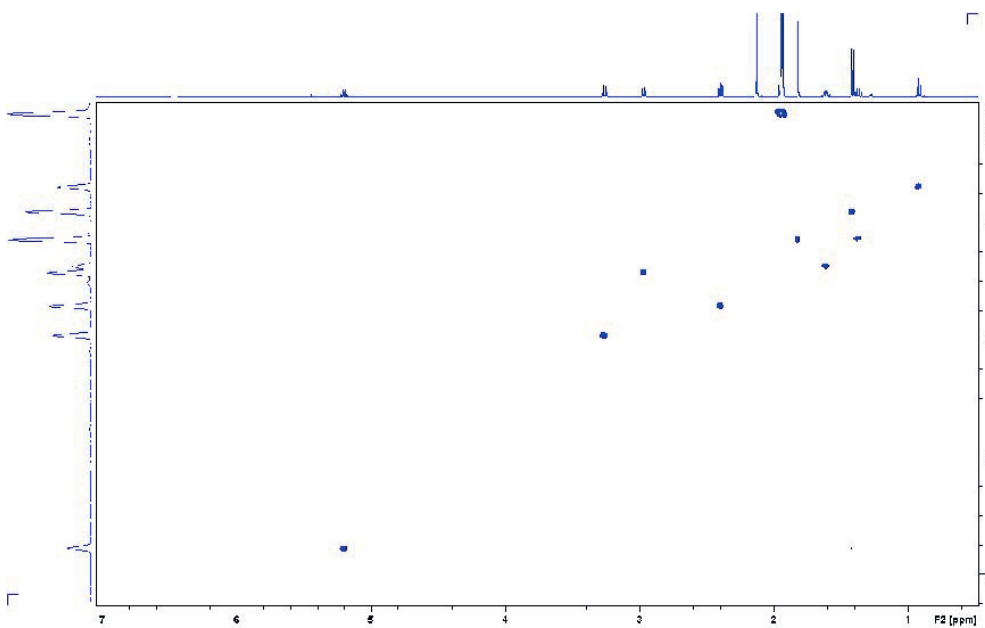


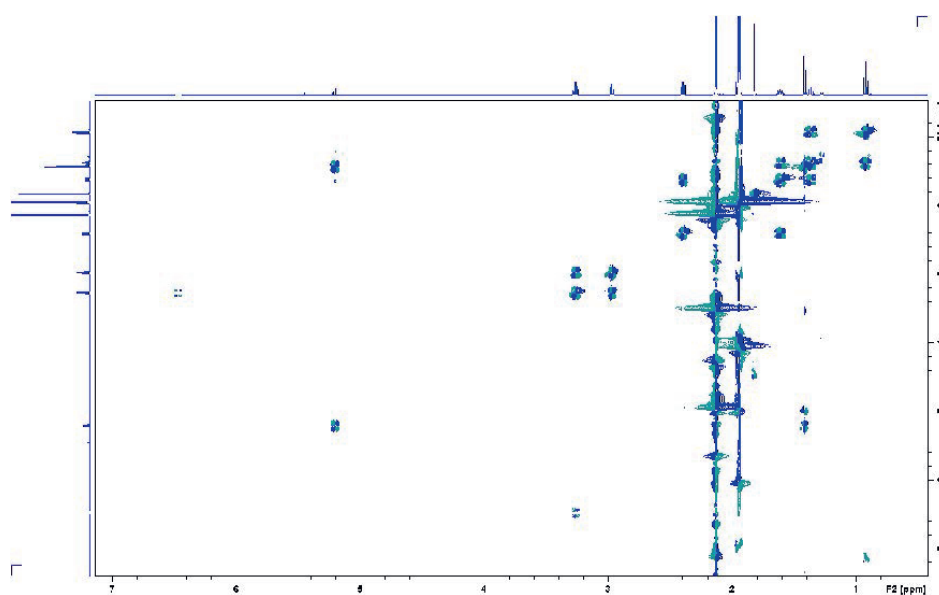**Figure S21.** HSQC spectrum of **17** in acetonitrile-$d$3 at 500 MHz.

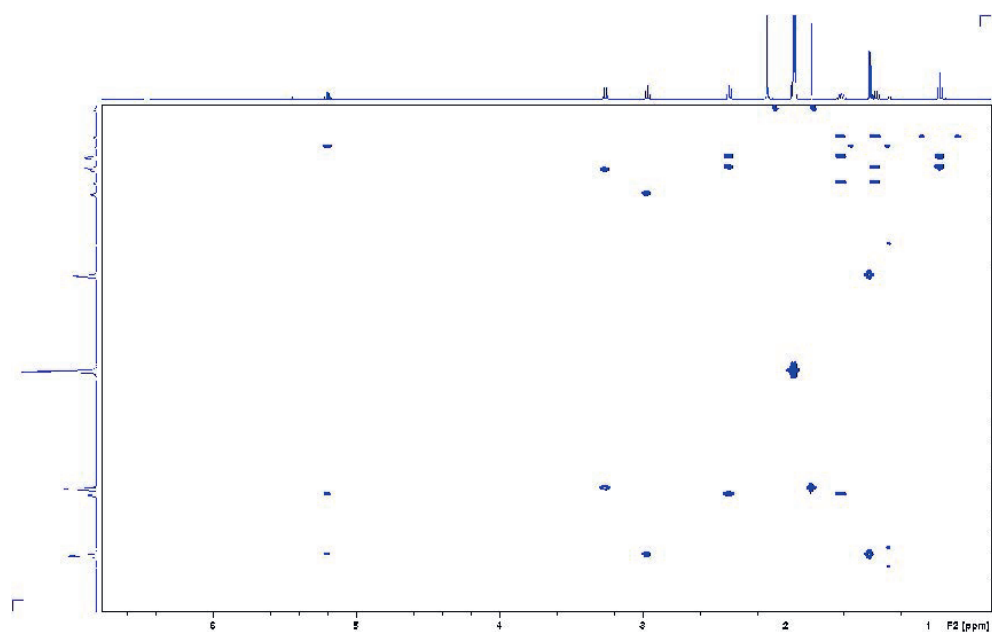**Figure S22. COSY spectrum of 17 in acetonitrile-*d*₃ at 500 MHz.**

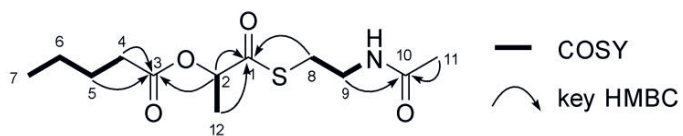**Figure S23. HMBC spectrum of 17 in acetonitrile-*d*₃ at 500 MHz.**



**Figure S24. Structure elucidation of 17**. COSY and key HMBC correlations are shown.
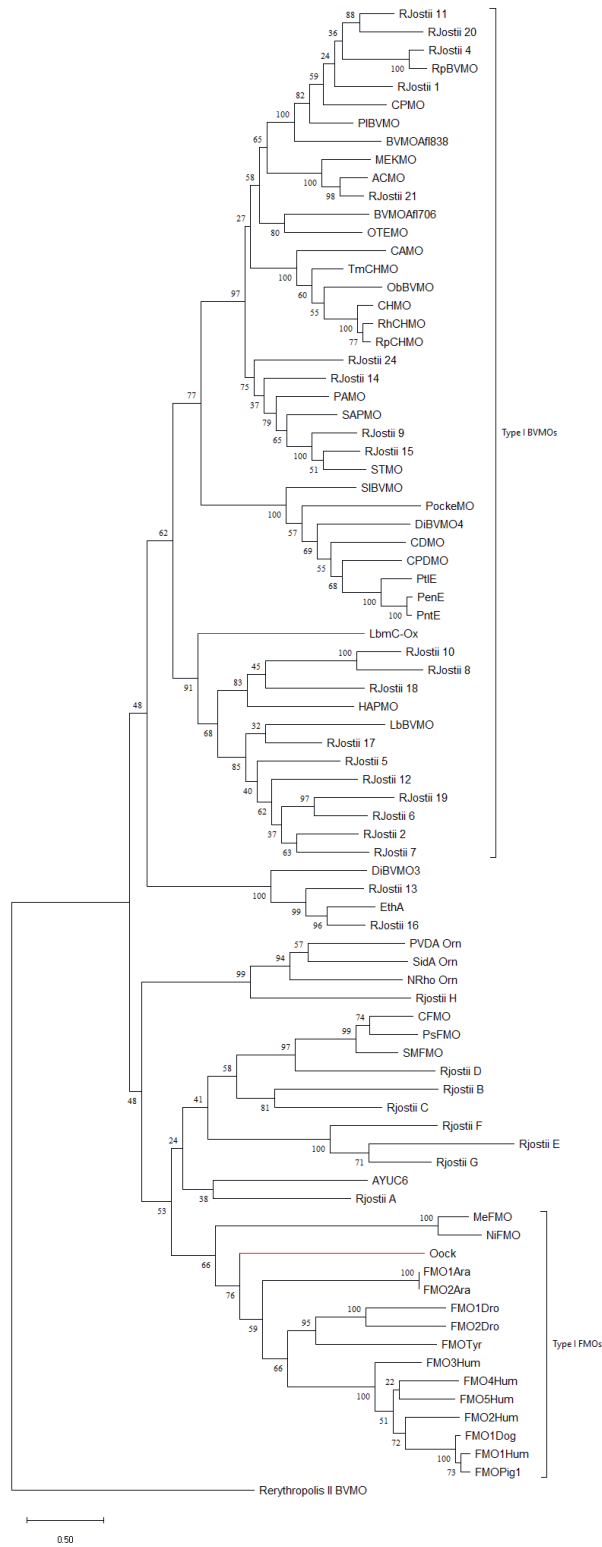
**Figure S25. Cladogram of class B flavoproteins.** Multiple sequence alignment was prepared using 82 amino acid sequences in the MUSCLE software (v3.8.31) configured with default settings for highest accuracy and employing the UPGMB clustering method. The evolutionary history was inferred by using the Maximum Likelihood (ML) method implemented in MEGA X (500 bootstrap replications).[30] The tree with the highest log likelihood (-80182.57) is shown. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 82 amino acid sequences. The default substitution model was selected assuming an estimated proportion of invariant sites and 4 gamma-distributed rate categories to account for rate heterogeneity across sites (WAG model). Nearest-Neighbor Interchange (NNI) ML heuristic method was chosen. Initial tree(s) for the heuristic search were obtained by applying the BioNJ method to a matrix of pairwise distances estimated using a JTT model.[31] OocK and LbmC-Ox highlighted with arrows. The tree indicates that OocK is a type I FMO while LbmC-Ox is related to type I Baeyer-Villiger monooxygenases.[33]
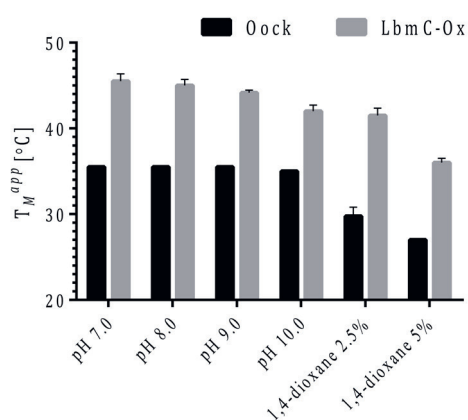
**Figure S26. Apparent melting temperatures profile of OocK and LbmC-Ox.** The thermostability of OocK (black bars) and LbmC-Ox (gray bars), were measured at different pH values (100 mM NaCl, 50 mM Tris-HCl or 50 mM CHES) and at 2.5 or 5% v/v of 1,4-dioxane



**Figure S27. MS spectra of obtained products.** List of MS spectra of the products detected using OocK (panels b, d, e and n) or LbmC-Ox (a-m) as catalyst. *Library does not contain the structure, product could be the normal or abnormal ester.

# REFERENCES

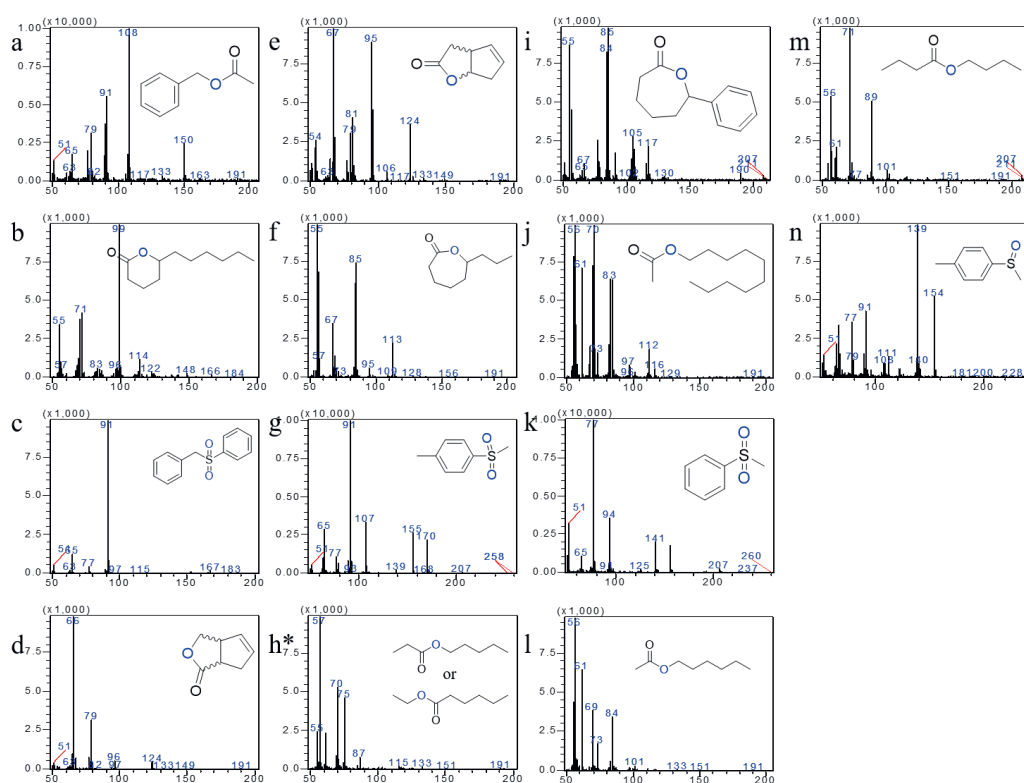1.  Hertweck, C. The biosynthetic logic of polyketide diversity. *Angewandte Chemie International Edition* 2009, 48 (26), 4688-716.

2.  Khosla, C., Tang, Y., Chen, A.Y., Schnarr, N.A., Cane, D.E. Structure and mechanism of the 6-deoxyerythronolide B synthase. *Annu. Rev. Biochem* 2007, 76, 195-221.

3.  Helfrich, E.J., Ueoka, R., Dolev, A., Rust, M., Meoded, R.A., Bhushan, A., Califano, G., Costa, R., Gugger, M. Steinbeck, C., Moreno, P. Automated structure prediction of trans-acyltransferase polyketide synthase products. *Nature chemical biology* 2019, 5 (8), 813-821.

4.  a) Piel, J. Biosynthesis of polyketides by trans-AT polyketide synthases. *Natural product reports.* 2010, 27 (7), 996-1047

    b) Helfrich, E.J., Piel, J. Biosynthesis of polyketides by trans-AT polyketide synthases. *Natural product reports* 2016, 33 (2), 231-316.

5.  Partida-Martinez, L.P., Hertweck, C. A gene cluster encoding rhizoxin biosynthesis in *Burkholderia rhizoxina*, the bacterial endosymbiont of the fungus *Rhizopus microsporus. ChemBioChem* 2007, 8 (1), 41-45.

6.  Piel, J. A polyketide synthase-peptide synthetase gene cluster from an uncultured bacterial symbiont of *Paederus beetles. Proceedings of the National Academy of Sciences* 2002, 99 (22), 14002-14007.

7.  a) Ueoka, R., Uria, A.R., Reiter, S., Mori, T., Karbaum, P., Peters, E.E., Helfrich, E.J., Morinaka, B.I., Gugger, M., Takeyama, H., Matsunaga, S. Metabolic and evolutionary origin of actin-binding polyketides from diverse organisms. *Nature chemical biology* 2015, 11(9), 705-712;

    b) Hildebrand, M., Waggoner, L.E., Liu, H., Sudek, S., Allen, S., Anderson, C., Sherman, D.H., Haygood, M. *bryA*: an unusual modular polyketide synthase gene from the uncultivated bacterial symbiont of the marine bryozoan *Bugula neritina. Chemistry & biology* 2004, 11 (11),1543-1552.

    c) Lopera, J., Miller, I.J., McPhail, K.L., Kwan, J.C. Increased biosynthetic gene dosage in a genome-reduced defensive bacterial symbiont. *Msystems* 2017, 2 (6).

8.  Ueoka, R., Bortfeld-Miller, M., Morinaka, B.I., Vorholt, J.A., Piel, J. Toblerols: Cyclopropanol-containing polyketide modulators of antibiosis in Methylobacteria. *Angewandte Chemie* 2018, 130 (4), 989-993.

9.  Meoded, R.A., Ueoka, R., Helfrich, E.J., Jensen, K., Magnus, N., Piechulla, B., Piel, J. A polyketide synthase component for oxygen insertion into polyketide backbones. *Angewandte Chemie International Edition* 2018, 57 (36), 11644-11648.

10. a) Bishara, A., Rudi, A., Aknin, M., Neumann, D., Ben-Califa, N., Kashman, Y. Salarins A and B and tulearin A: new cytotoxic sponge-derived macrolides. *Organic letters* 2008, 10 (2), 153-156.

    b) Northcote, P.T., Blunt, J.W., Munro, M.H. Pateamine: a potent cytotoxin from the New Zealand marine sponge, *Mycale* sp. *Tetrahedron letters* 1991, 32 (44), 6411-6414.

11. a) Dunn, Z.D., Wever, W.J., Economou, N.J., Bowers, A.A., Li, B. Enzymatic basis of "hybridity" in thiomarinol biosynthesis. *Angewandte Chemie* 2015, 127 (17), 5226-5230.

    b) Biggins, J.B., Ternei, M.A., Brady, S.F. Malleilactone, a polyketide synthase-derived virulence factor encoded by the cryptic secondary metabolome of *Burkholderia pseudomallei* group pathogens. *Journal of the American Chemical Society* 2012, 134 (32), 13192-13195.

    c) Franke, J., Ishida, K., Hertweck, C. Genomics-driven discovery of burkholderic acid, a noncanonical, cryptic polyketide from human pathogenic *Burkholderia* species. *Angewandte Chemie* 2012, 124 (46), 11779-11783.

12. Zallot, R., Oberg, N., Gerlt, J.A. The EFI web resource for genomic enzymology tools: leveraging protein, genome, and metagenome databases to discover novel enzymes and metabolic pathways. *Biochemistry* 2019, 58 (41), 4169-4182.

13. Chung, E.J., Park, J.A., Jeon, C.O., Chung, Y.R.. *Gynuella sunshinyii* gen. nov., sp. nov., an antifungal rhizobacterium isolated from a halophyte, *Carex scabrifolia* Steud. *International journal of systematic and evolutionary microbiology* 2015, 65 (3), 1038-1043.

14. Ueoka, R., Bhushan, A., Probst, S.I., Bray, W.M., Lokey, R.S., Linington, R.G., Piel, J. Genome-Based Identification of a Plant-Associated Marine Bacterium as a Rich Natural Product Source. *Angewandte Chemie* 2018, 130 (44),14727-14731.

15. Nguyen, T., Ishida, K., Jenke-Kodama, H., Dittmann, E., Gurgui, C., Hochmuth, T., Taudien, S., Platzer, M., Hertweck, C., Piel, J. Exploiting the mosaic structure of trans-acyltransferase polyketide synthases for natural product discovery and pathway dissection. *Nature biotechnology* 2008, 26(2), 225-233.

16. Calderone, C.T., Kowtoniuk, W.E., Kelleher, N.L., Walsh, C.T., Dorrestein, P.C. Convergence of isoprene and polyketide biosynthetic machinery: isoprenyl-S-carrier proteins in the pksX pathway of *Bacillus subtilis*. *Proceedings of the National Academy of Sciences* 2006, 103 (24), 8977-8982.

17. Pérez-Sayáns, M., Somoza-Martín, J.M., Barros-Angueira, F., Rey, J.M., García-García, A. V-ATPase inhibitors and implication in cancer treatment. *Cancer treatment reviews* 2009, 35 (8), 707-713.

18. McKee, T.C., Galinis, D.L., Pannell, L.K., Cardellina, J.H., Laakso, J., Ireland, C.M., Murray, L., Capon, R.J., Boyd, M.R. The lobatamides, novel cytotoxic macrolides from southwestern *Pacific tunicates*. The *Journal of Organic Chemistry* 1998, 63 (22), 7805-7810.

19. Xie, X.S., Padron, D., Liao, X., Wang, J., Roth, M.G., De Brabander, J.K. Salicylihalamide A inhibits the V0 sector of the V-ATPase through a mechanism distinct from bafilomycin A1. *Journal of biological chemistry* 2004, 279 (19), 19755-19763.

20. Dekker, K.A., Aiello, R.J., Hirai, H., Inagaki, T., Sakakibara, T., Suzuki, Y., Thompson, J.F., Yamauchi, Y., Kojima, N. Novel lactone compounds from *Mortierella verticillata* that induce the human low density lipoprotein receptor gene: Fermentation, isolation, structural elucidation and biological activities. *The Journal of antibiotics* 1998, 51 (1), 14-20.

21. Kunze, B., Jansen, R., Sasse, F., Höfle, G., Reichenbach, H. Apicularens A and B, new cytostatic macrolides from *Chondromyces* species (*Myxobacteria*). *The Journal of antibiotics* 1998, 51 (12), 1075-1080.

22. Kim, J.W., Shin-Ya, K., Furihata, K., Hayakawa, Y., Seto, H. Oximidines I and II: novel antitumor macrolides from *Pseudomonas* sp. *The Journal of organic chemistry* 1999 64 (1), 153-155.

23. Niehs, S.P., Dose, B., Richter, S., Pidot, S.J., Dahse, H.M., Stinear, T.P., Hertweck, C. Mining symbionts of a spider-transmitted fungus illuminates uncharted biosynthetic pathways to cytotoxic benzolactones. *Angewandte Chemie International Edition* 2020, 59 (20), 7766-7771.

24. a) Fisch, K.M., Gurgui, C., Heycke, N., van der Sar, S.A., Anderson, S.A., Webb, V.L., Taudien, S., Platzer, M., Rubio, B.K., Robinson, S.J., Crews, P., J. Piel. *Nat. Chem. Biol.* 2009 5, 494-501.
    b) Ahrendt, T., Miltenberger, M., Haneburger, I., Kirchner, F., Kronenwerth, M., Brachmann, A.O., Hilbi, H., Bode, H.B. Biosynthesis of the natural fluorophore legioliulin from l*egionella*. *ChemBioChem* 2013, 14 (12), 1415-1418.

25. Riebel, A., Fink, M.J., Mihovilovic, M.D., Fraaije, M.W. Type II flavin-containing monooxygenases: a new class of biocatalysts that harbors Baeyer–Villiger monooxygenases with a relaxed coenzyme specificity. *ChemCatChem* 2014, 6 (4), 1112-1117.

26. Edgar, R.C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004, *32*, (5), 1792-1797.

27. Acebal, C., Alcazar, R., Cañedo, L. M., De La Calle, F., Rodriguez, P., Romero, F., Puentes, J. F. Two marine Agrobacterium producers of sesbanimide antibiotics. *The Journal of Antibiotics* 1998, 51 (1), 64-67.

28. Piasecki, S. K., Taylor, C. A., Detelich, J. F., Liu, J., Zheng, J., Komsoukaniants, A., Siege, D.R., Keatinge-Clay, A.D. Employing modular polyketide synthase ketoreductases as biocatalysts in the preparative chemoenzymatic syntheses of diketide chiral building blocks. *Chemistry & biology* 2011, 18 (10), 1331-1340.

29. Forneris, F., Orru, R., Bonivento, D., Chiarelli, L. R., Mattevi, A. ThermoFAD, a Thermofluor®-adapted flavin ad hoc detection system for protein folding and ligand binding. *FEBS J.* 2009, 276, (10), 2833-2840.

30. Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular biology and evolution* 2018, 35 (6), 1547-1549.

31. Whelan, S., & Goldman, N. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Molecular biology and evolution* 2001, 18 (5), 691-699.

32. a) Riebel, A., Fink, M.J., Mihovilovic, M.D., Fraaije, M.W. Type II flavin-containing monooxygenases: a new class of biocatalysts that harbors Baeyer–Villiger monooxygenases with a relaxed coenzyme specificity. *ChemCatChem* 2014, 6 (4), 1112-1117.

33. b) Van Berkel, W. J. H., Kamerbeek, N. M., & Fraaije, M. W. Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts 2006, *Journal of biotechnology*, 124 (4), 670-689.

# Identification of two flavin-containing monooxygenases from *Chloroflexi* and *Hypsibius dujardini*

**Alejandro Gran-Scheuch,** Simone Savino, Nikola Lončar, Pimvisuth Chunkrua, Milos Trajkovic, Loreto Parra and Marco W. Fraaije

## ABSTRACT

Flavin-containing monooxygenases (FMOs) are attractive biocatalysts able to oxidize a wide spectrum of substrates with soft nucleophilic heteroatoms, such as N or S. Moreover, it has been established that some members of this group can perform Baeyer-Villiger oxidations. Recently, FMOs have been classified in two types: type I and II FMOs. Previously reported type II FMOs show affinity for both NADPH and NADH as hydride donor, while type I FMOs only accept the phosphorylated cofactor. Through genome mining we identified two putative FMOs, from *Chloroflexi* and *Hypsibius dujardini* (a tardigrade species): CbFMO (type II) and HdFMO (type I), respectively. Both enzymes were expressed fused to SUMO protein and purified as FAD-containing soluble proteins. Characterization revealed that both enzymes display different biochemical properties. CbFMO converts preferentially ketones, displays a poor thermostability ($T_M^{app}$ of 34°C), and shows a preference for NADPH as coenzyme despite its clustering within type II FMOs. HdFMO, on the other hand, was found to be active with sulfides, did only accept NADPH as hydride donor, and was more robust as evidenced by its $T_M^{app}$ of 45°C.

**Keywords:** FMOs, genome mining, biocatalysis, flavoproteins, tardigrade

# INTRODUCTION

Enzymes are commonly used as biocatalysts by organic chemists in both industry and academia. This is because enzymes are attractive for (bio)chemosynthesis, due to their remarkably high chemo-, regio- and enantiospecificity/selectivity[1–5]. One particular group of enzymes that are attractive of performing selective oxygenations, are the flavin-containing monooxygenases (FMOs), which belong to the class B flavoprotein monooxygenases[6]. This class of enzymes is interesting because these NAD(P)H-dependent and FAD-containing enzymes typically oxidize a wide spectrum of substrates and have been shown to be useful for the synthesis of chemical building blocks[6–9]. Commonly, FMOs catalyze the oxidation of soft nucleophilic heteroatoms as N or S, using molecular oxygen as oxidant. Furthermore, it has been described that some FMOs also exhibit Baeyer-Villiger monooxygenase (BVMO) activity[10,11].

Recently, based on distinct sequence features, FMOs have been grouped into two different subclasses, type I and II FMOs[11,12]. Specifically, type I FMOs show a higher affinity for NADPH than NADH, while type II FMO can use both nicotinamide cofactors with a similar efficiency. Furthermore, several type II FMOs also showed to act as BVMOs, and have been described to oxidize small cyclic ketones[12,13]. Mechanistically, like other class B flavoprotein monooxygenases, the catalytic mechanism of an FMO begins with the reduction of the fully oxidized and tightly bound FAD by NAD(P)H[6]. Then, in the so-called oxidative half-reaction, dioxygen reacts with the reduced flavin cofactor. As a consequence, a quasi-stable C4α-(hydro)peroxyflavin adduct is formed, and depending on the protonation state, this intermediate will act as nucleophile or electrophile. The hydroperoxyflavin state is involved in the electrophilic oxidation of heteroatoms, while the peroxyflavin can perform the nucleophilic attack of electron-poor ketones leading to esters or lactones[14]. While the enzyme inserts one oxygen atom into the organic substrate, and the another is reduced to water. With the dissociation of NAD(P)$^+$, the catalytic cycle is complete. The sequences of type I and type II FMOs share many features, including the sequence motifs for FAD and NAD(P)H binding. This makes it likely that all FMO share a similar catalytic mechanism.

Only a handful of type I and type II FMOs have been studied in the context of biocatalysis, all the known examples are of bacterial origin. An interesting type I FMO is NiFMO, identified from *Nitrincola lacisaponensis,* an alkaliphilic bacterium isolated from Soap Lake (USA)[15]. This FMO is a relatively thermostable dimer ($T_M$$^{app}$ of 51 °C) with 71 % sequence identity with mFMO from *Methylophaga* sp. strain SK1. NiFMO was shown to produce indigoid pigments from indoles and displayed a strong activity towards N-containing substrates. Additionally, mFMO has also been explored for the biocatalytic production of indigo from indole while it was also shown to oxidize a variety of sulfides[8,16]. A few type II FMOs have been studied as biocatalysts, which includes FMO-E from *Rhodoccocus jostii* RHA1 and

PsFMO-A from *Pimelobacter sp.*[12,13]. Both enzymes oxidize small (bi)cyclic ketones using either NADH or NADPH.

Type I and type II FMOs are of great biocatalytic interest owing to their broad substrate scope, enabling the production of sulfoxides and/or esters and lactones. Type II FMOs appear especially attractive due to their relaxed nicotinamide cofactor dependency. Yet, to be useful, enzymes must perform the desired reaction with outstanding catalytic parameters, such as a high activity and a wide substrate scope. Moreover, a *good* biocatalyst is expected to tolerate harsh conditions, such as the presence of cosolvents or high temperatures[17–19]. To satisfy such requirements, enzymes can be improved by protein engineering[20–22]. An alternative approach is to search for novel putative biocatalysts by mining the genome of extremophilic organisms[23–26]. This represents a bioinformatic approach that capitalizes on the vast amount of genetic information available to search for novel enzymes. FMOs contain characteristic sequence motifs that can be used for identifying FMO-encoding genes[23,26,27]. FMO sequences can be identified by the presence of specific fingerprints within their sequence. Initially, a FMO-specific motif (F-x-G-x-x-x-H-x-x-x-[Y/F]-[K/R]) and two Rossmann fold motifs (G-x-G-x-x-[G/A]) were described to be part of a FMO sequence[6,28]. The latter motifs reflect the two domains that bind NAD(P)H and FAD[29]. However, in recent years, for type II FMOs an extension of the N-terminus and some modifications in the 'FMO' sequence fingerprint were described (including [H] instead of [Y/F] and [D], [P], [V] or [G] instead of [K/R])[12,13]. In this work, we used the above-mentioned FMO sequence motifs as query to search for new potential FMOs in the genome (predicted proteome) database. This yielded two putatively FMO-encoding genes from the genomes of a *Chloroflexi* bacterium and the tardigrade *Hypsibius dujardini*: CbFMO (type II FMO) and HdFMO (type I FMO), respectively. These two FMOs were found to display distinct biocatalytic features.

## RESULTS

### Identification and expression of novel FMOs

Using the protein sequence of two recently described FMOs as templates (NiFMO and FMO-E), protein searches using the basic local alignment search tool (BLAST) from the NCBI database were performed. From this exploration, two putative proteins were selected: OQV24618 (HdFMO) and RLT40563 (CbFMO). Both were named according to the source of the organism from where they were identified. HdFMO, has a hypothetical molecular weight of 57.2 kDa and its corresponding gene is present in the genome of *Hypsibius dujardini*, a tardigrade (also known as water bear or moss piglet). This putative protein showed 29 % sequence identity with NiFMO (363 residues overlap out of 505). CbFMO has a hypothetical molecular weight of 65.0 kDa and was identified from a *Chloroflexi* bacterium. CbFMO exhibited 46 % sequence identity with FMO-E (572 residues overlap out

of 585). When compared, both putative proteins presented an identity of 25.5 % between each other (in only 110 residues overlap). HdFMO and CbFMO contain two Rossmann fold motifs (GxGxxG/A) within their sequences, even though for HdFMO the second motif had a variation of a lysine instead of the conserved glycine (**Table 1**). Additionally, for both sequences, the FMO-specific fingerprint was found. While HdFMO exhibited the classic type I FMO motif, CbFMO showed a natural variation of an aspartic acid residue at the [K/R] position. Subsequently, a multiple alignment sequence analysis using 48 other known flavin-dependent monooxygenases, and a phylogenetic analysis was performed. As expected, HdFMO was found to be closely related to the type I FMOs, while CbFMO clustered with type II FMOs (**Figure 1**).

**Table 1.** CbFMO and HdFMO fingerprint sequences.

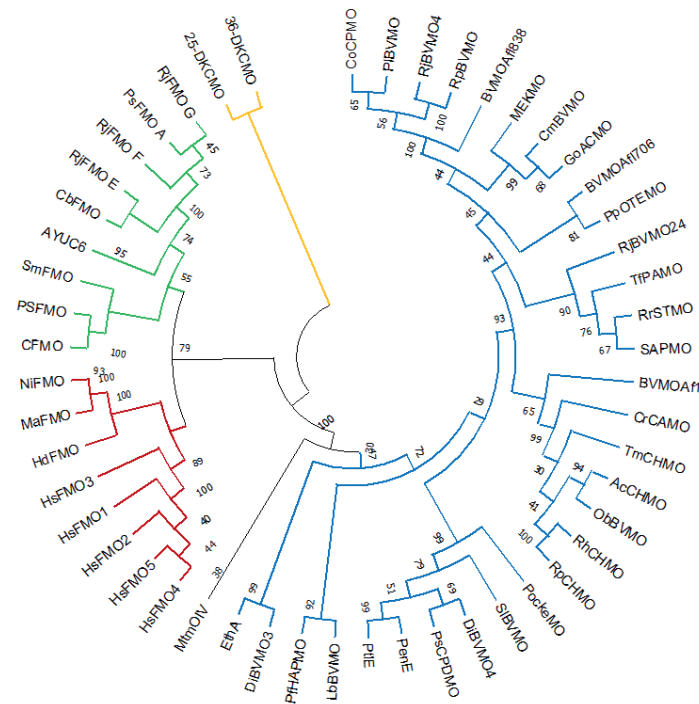|  | Rossmann fold | FMO Fingerprint | Rossmann fold |
|---|---|---|---|
| **CbFMO** | L-V-I-**G**-T-**G**-S-S-**A** | **F**-E-**G**-D-V-I-**H**-S-E-G-**F**-**D** | L-V-V-**G**-G-**A**-Q-A-**G** |
| **HdFMO** | A-V-I-**G**-A-**G**-M-T-**G** | **F**-K-**G**-T-V-M-**H**-S-C-E-**Y**-**R** | L-I-I-**G**-A-**K**-T-S-**A** |



**Figure 1. Evolutionary analysis by Maximum Likelihood method.** The cladogram was inferred by using the Maximum Likelihood method and Whelan And Goldman model[30]. The tree with the highest log likelihood (-52182) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. The color of the clade represents the flavoprotein group to which the respective flavoprotein belongs (blue for type I BVMOs, yellow for type II BVMOs, black for type O BVMOs, red for type I FMOs, and green for type II FMOs). CbFMO (type II FMO) and HdFMO (type I FMO) were clustered within the respective FMO branches.

Both enzymes lack predicted transmembrane domains according to the TMHMM server v 2.0, suggesting that these FMOs are soluble proteins. Subsequently, synthetic genes were codon optimized and ordered with BsaI recognition site overhangs for Golden Gate cloning. Both genes were cloned into a pBAD-based plasmid encoding a SUMO fusion protein with a N-terminal hexa-histidine tag for purification purposes. To obtain CbFMO as soluble protein, it was necessary to include additives during cell lysis (DTT, glycerol and IGEPAL). Moreover, to avoid aggregation, after the desalting step, CbFMO required a high salt concentration (350 mM NaCl). The expression and purification of both FMOs resulted in decent yields: 19 mg L$^{-1}$ and 15 mg L$^{-1}$ for HdFMO and CbFMO, respectively. Additionally, the spectrophotometric analysis of CbFMO and HdFMO revealed the typical UV-visible absorbance profile of flavoproteins (**Figure 2a,b**). Both proteins showed a distinctive absorbance maximum at 450 nm. However, for the second absorbance maximum, CbFMO exhibited a maximum at 386 nm, while for HdFMO it was shifted to 372 nm. These UV-vis absorbance features show that both enzymes had folded in a competent conformation to bind the FAD cofactor as prosthetic group. To assess their oligomeric state, both enzymes were incubated with SUMO protease and analyzed by gel permeation and dynamic light scattering (DLS). The results from size exclusion chromatography (**Figure 2c**) and DLS analyses (**Figure 2d,e**) suggested that both CbFMO and HdFMO form stable dimers, as has been observed for other FMOs. The DLS analysis indicated for HdFMO a molecular weight of ~100 kDa, with a polydispersity of 14 %. On the other hand, CbFMO tended to aggregate and the polydispersity was around 32 %, and a molecular weight of 125 kDa was estimated. Finally, to compare their thermostability, the ThermoFAD method was used. HdFMO seemed to be relatively stable, displaying a $T_M^{app}$ of 45 °C, while for CbFMO a $T_M^{app}$ value of only 34 °C was measured. Practically, for both proteins, the presence of 5.0 % v/v 1,4-dioxane reduced the thermostability by only 3 °C. Therefore, for further experiments that included cosolvent, samples contained at most 2.5 % v/v 1,4-dioxane.

## Analysis of substrate scope and small-scale conversions

To explore the biocatalytic potential of both FMOs, their substrate scope was analyzed through a GC-MS screening method[31]. A collection of 44 potential substrates were tested either individually or in combination. Specifically, CbFMO and HdFMO were incubated with 6 mixtures of potential substrates and 15 individual screening reactions, which included ketones (27), sulfides (14) and N-containing molecules (3), at 500 μM and 2.5 % v/v 1,4-dioxane as a cosolvent (**Table S1**). For the substrate screening, the enzymes were incubated with 150 μM NADH and 150 μM NADPH. For cofactor regeneration, 5.0 μM PTDH and 20 mM sodium phosphite were added to the reaction mix. The small-scale conversions were incubated for 24 h at 24°C for HdFMO and 24 h at 17°C for CbFMO. Substrate conversion was confirmed by verifying product formation through MS. The results showed a significant difference in the substrate acceptance profile for both enzymes (**Table 2**). CbFMO exhibited a wider acceptance for ketones, mainly converting allylic and
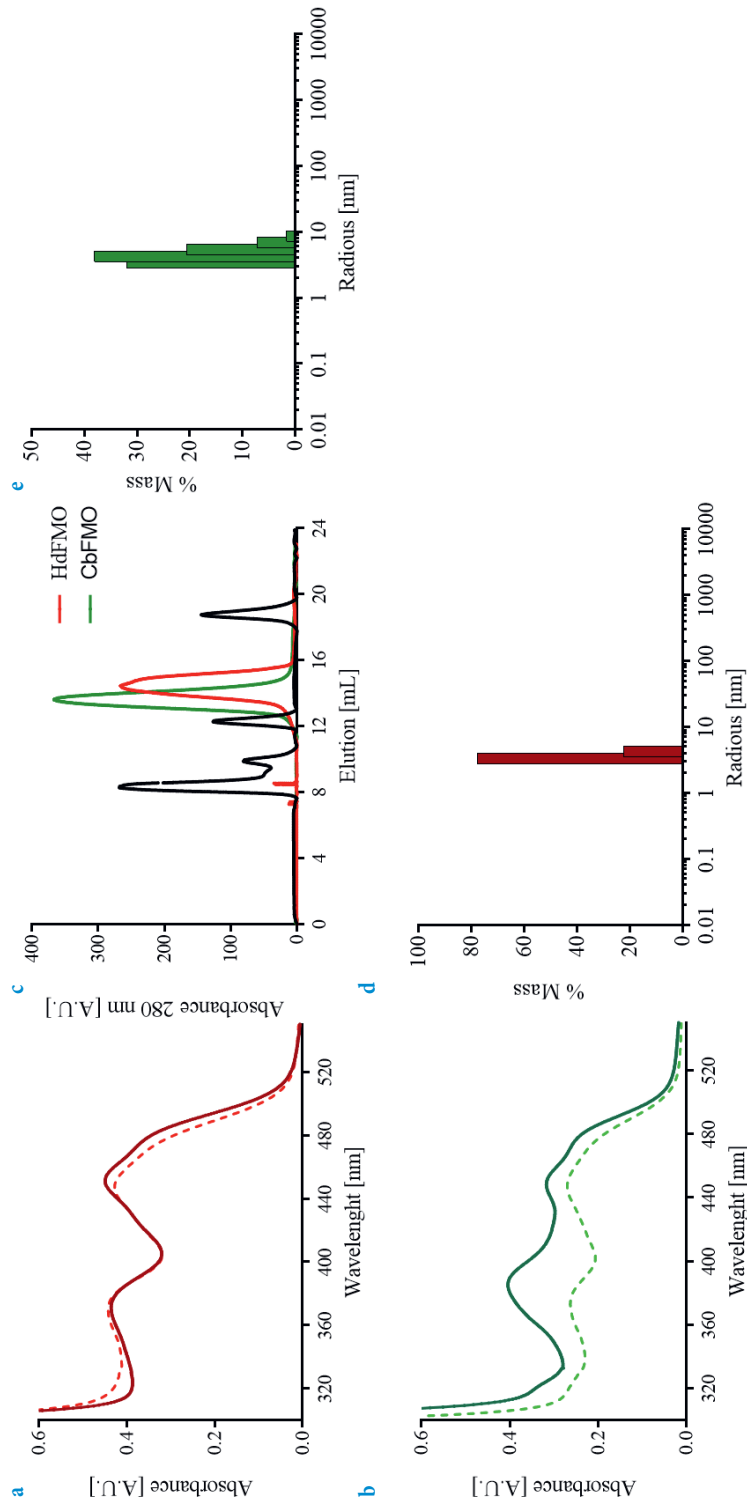
**5**



**Figure 2. UV-vis absorbance and hydrodynamic properties on the FMOs.** UV-vis absorbance spectra for (a) HdFMO and (b) CbFMO. Visible spectra of native FMOs (solid line) and FMOs after unfolding by 0.1% SDS (dotted line) were analyzed. (c) Gel permeabilization analyses of purified tag-free FMOs are shown together with the elution of marker proteins (black). (d) DLS analysis indicates a monodispersed peak corresponding to a MW of ~100 kDa for HdFMO, suggesting the presence of a dimer. (e) CbFMO showed to be less monodispersed, corresponding to a peak with a MW of ~120 kDa.

cyclic ketones, while HdFMO only accepted sulfides. Subsequently, rates of nicotinamide cofactor consumption were monitored spectrophotometrically. First, uncoupling rates were established by monitoring oxidation of NAD(P)H in the absence of any substrate. This yielded similar uncoupling rates ($k_{unc}$) for CbFMO with NADH and NADPH (0.06 s$^{-1}$). In contrast, HdFMO was only able to oxidize NADPH with an uncoupling rate of <0.01 s$^{-1}$.

**Table 2. Substrate scope analysis.** Accepted substrates by CbFMO or HdFMO are shown. The level of conversion is symbolized by: - (no product detected), + (below 40 %), ++ (40-70 %) and +++ (70-100 %). *The TIC area of the substrate was diminished (compared to control), but no product was detected.

| Accepted substrate | HdFMO | CbFMO |
|---|---|---|
| phenylacetone | - | +++ |
| bicyclo[3.2.0]hept-2-en-6-one | - | +++ |
| cyclohexanone | - | + |
| 2-hexylcyclopentanone | - | ++ |
| 2-propylcyclohexanone | - | ++ |
| 4-phenylcyclohexanone | - | + |
| 4-methylcyclohexanone | - | ++ |
| cyclopentanone | - | * |
| 4-hydroxyacetophenone | - | +++ |
| allyl phenyl sulfide | + | - |
| benzyl methyl sulfide | + | + |
| ethyl phenyl sulfide | + | - |
| methyl-*p*-tolyl sulfide | ++ | - |
| thioanisole | +++ | ++ |
| 4-chloro thioanisole | + | - |
| benzyl ethyl sulfide | +++ | + |

Next, the NADPH oxidation activities for both FMOs were tested using 1.0 mM of substrate (soluble accepted substrates from table 2). Additionally, cyclobutanone was included in this analysis. This cyclic ketone has been described to be accepted by type II FMOs but could not be identified by GC-MS. Thus, it was not tested in the screening analysis but included when measuring activity using the spectrophotometric assay. In general, the results were mostly disappointing as both FMOs exhibit rather low oxidation rates (**Table 3**). When using NADPH, CbFMO displayed relatively low oxidation rates for 2-hexylpentanone, phenyl-acetone, *rac*-bicyclo[3.2.0]hept-2-en-6-one and cyclobutanon: $k_{obs}$ of 0.13, 0.25. 0.35 and 0.6 s$^{-1}$, respectively. For cyclobutanone, CbFMO exhibited a $k_{cat}$ and K$_M$ of 0.6 s$^{-1}$ and 300 μM (**Figure S1**). For HdFMO, of the 7 tested sulfide substrates, showed highest activity for allyl phenyl sulfide and 4-chlorothioanisole (0.08 and 0.15 s$^{-1}$, respectively). Interestingly, the rate of NADPH oxidation using thioanisole —which was fully converted in the screening test— was rather low (0.03 s$^{-1}$). This may be due to inhibition issues.

**Table 3. Cofactor oxidation rates.** NADPH oxidation rates were spectrophotometrically monitored at $340_{nm}$. Activities of CbFMO and HdFMO with 1.0 mM substrate and 100 µM NADPH are shown.

| Substrate | HdFMO [s⁻¹] | CbFMO [s⁻¹] |
|---|---|---|
| uncoupling | <0.01 | 0.06 |
| phenylacetone | n.d. | 0.25 |
| *rac*-bicyclo[3.2.0]hept-2-en-6-one | n.d. | 0.35 |
| cyclohexanone | n.d. | 0.03 |
| 2-hexylcyclopentanone | n.d. | 0.13 |
| 2-propylcyclohexanone | n.d. | 0.05 |
| 4-phenylcyclohexanone | n.d. | 0.04 |
| 4-methylcyclohexanone | n.d. | 0.06 |
| cyclobutanone | n.d. | 0.60 |
| cyclopentanone | n.d. | 0.05 |
| 2-butanone | n.d. | 0.03 |
| 3-methylcyclopentanone | n.d. | 0.07 |
| allyl phenyl sulphide | 0.08 | n.d. |
| benzyl methyl sulfide | 0.02 | n.d. |
| ethyl phenyl sulfide | 0.02 | n.d. |
| methyl-*p*-tolyl sulfide | 0.02 | n.d. |
| thioanisole | 0.03 | n.d. |
| 4-chlorothioanisole | 0.15 | n.d. |
| benzyl ethyl sulfide | 0.03 | n.d. |

Finally, the two FMOs were tested for small-scale conversions using some selected substrates. Like the substrate scope analysis, HdFMO-catalyzed conversions were performed for 24 h at 24 °C, using PTDH as regeneration system and NADPH as hydride donor. For HdFMO conversions analysis, it was only tested using NADPH because this enzyme showed no activity with NADH: it was unable to convert thioanisole in the presence of NADH (data not shown). Interestingly, even when the $k_{obs}$ for 4-chlorothioanisole was almost five times higher than for thioanisole, the conversion reached only 66 % while thioanisole was fully converted, producing only the respective sulfoxide (**Table 4**). For ethyl benzyl sulfide, methyl-*p*-tolyl sulfide, ethyl phenyl sulfide, allyl phenyl sulfide and benzyl methyl sulfide, HdFMO showed conversion of 95, 80, 50, 40 and 40 %, respectively.

In the case of CbFMO, several reaction conditions were explored using 1.0 mM phenylacetone. At 17 °C for 24 h, this substrate was 99 % converted using NADPH, while in the presence of NADH the oxidation reached only 30 %. Due to its lower stability and the tendency to aggregate at low ionic strength, the reaction was also analyzed at 24 and 4 °C, in the presence of 350 mM NaCl and in phosphate buffer. As expected, the reaction at higher temperature reduced the conversion to 20 % when using NADPH while no conversion was obtained with NADH. For the other conditions, the conversions using

NADPH were complete, while using NADH conversions of 13-30 % were observed (**Table S2**). All subsequent conversions were performed at 17 °C, revealing that in the presence of NADH, CbFMO was only able to oxidize *rac*-bicyclo[3.2.0]hept-2-en-6-one (28 %), cyclohexanone (≤ 5 %) and 4-hydroxyacetophenone (10 %). In the presence of NADPH, the conversion levels were higher. CbFMO fully converted *rac*-bicyclo[3.2.0]hept-2-en-6-one, phenylacetone, and 4-hydroxyacetophenone while 2-hexylcyclopentanone, 4-methylcyclohexanone, cyclobutanone, thioanisole, 2-propylcyclohexanone, cyclo-hexanone and 4-phenylcyclohexanone gave conversions of 40-80 % (**Table 4**). For the conversion of cyclobutanone, the reaction was analyzed by ^{1}H-NMR (Figure S2). For the latter, CbFMO converted 70 % of the substrate. Importantly, for the sulfide thioanisole, only the sulfoxide and not the sulfone was produced. For the Baeyer-Villiger oxidations, the so-called normal esters were formed. In the case of 4-hydroxyacetophenone conversions, low levels of hydroquinone were also found (**Figure S3a**), indicating that the formed ester was hydrolyzed to some extent. Similarly, for phenylacetone, formation of benzylic alcohol was detected (**Figure S3b**). Finally, the enantio- and regioselectivity of CbFMO was tested using *rac*-bicyclo[3.2.0]hept-2-en-6-one as substrate. Chiral GC-analysis was used to detect formation of all possible enantiomeric products. Using NADPH as cofactor, (1*S*,5*R*)-2-oxabicyclo[3.3.0]oct-6-en-3-one (normal product) and (1*R*,5*S*)-3-oxabicyclo[3.3.0]oct-6-en-2-one (abnormal product) were predominantly formed out of four possible products with an *e.e.* of 80 and 99 %, respectively. As a positive control, a conversion of bicyclo[3.2.0]hept-2-en-6-one by TmCHMO was performed. This yielded, as expected, similar products as with CbFMO[35,37](**Figure S4**).

**Table 4.** **Small-scale conversions using FMOs as catalysts.** Products were identified by MS or ^{1}H-NMR and conversion levels are shown. *For ^{1}H-NMR analysis the small-scale conversion was performed in presence of 10 mM substrate.

| Substrate | Structure | Product | HdFMO | CbFMO | |
|---|---|---|---|---|---|
| | | | NADPH [%] | NADH [%] | NADPH [%] |
| 4-chlorothioanisole | | | 66 | - | - |
| thioanisole | | | 99 | - | 60 |
| ethyl benzyl sulfide | | | 95 | - | - |

| Substrate | Substrate structure | Product structure(s) | | | |
|---|---|---|---|---|---|
| methyl-p-tolyl sulfide | | | 80 | - | - |
| ethyl phenyl sulfide | | | 50 | - | - |
| allyl phenyl sulfide | | | 40 | - | - |
| benzyl methyl sulfide | | | 40 | - | - |
| bicyclo[3.2.0]hept-2-en-6-one | | | - | 30 | 99 |
| 4-hydroxyacetophenone | | | - | 10 | 99 |
| 2-hexylcyclopentanone | | | - | - | 80 |
| 4-methylcyclohexanone | | | - | - | 75 |
| cyclobutanone | | | - | - | 70* |
| 2-propylcyclohexanone | | | - | - | 50 |
| cyclohexanone | | | - | 5 | 40 |
| 4-phenylcyclohexanone | | | - | - | 40 |
| phenylacetone | | | - | 30 | 99 |

## DISCUSSION

In the present study we describe the identification, cloning and characterization of two novel FMOs: CbFMO and HdFMO. CbFMO exhibits 46% sequence identity with FMO-E and was identified as part of an analysis of a fresh water metagenome, and originated from an unclassified *Chloroflexi* bacterium[32]. Representatives of the *Chloroflexi* or *Chlorobacteria* phylum includes flagellated, aerobic, photoheterotrophic bacteria, which play a major role in demineralization of nitrogen-rich dissolved organic matter in the hypolimnion[32–34]. These microbes have been found to form a consistently large fraction of the prokaryotic communities in deep lake hypolimnia all over the world.

HdFMO, which exhibited 30 % identity with NiFMO, was identified in the genome of *Hypsibius dujardini*, a tardigrade. Tardigrades (colloquially known as water bears or moss piglets) are eight-legged extremely small animals and have been identified in a variety of extreme environments such as deep sea, mud volcanoes, Arctic and Antarctic samples. They can survive in extreme environments. While the genome has been reported in 2015, no enzymes from this organism had been studied in the context of biocatalysis[35]. When analyzing the genome of *H. dujardini*, eight other putative FMO-encoding genes can be identified. It is well described that FMOs are ubiquitous in all domains of life, acting mainly as membrane-associated metabolizers of xenobiotics such as toxins, pesticides and drugs[36,37]. Additionally, it has been suggested that the FMOs were already present in the last universal common ancestor (LUCA)[37]. Later, for tetrapods and mammals, these FMOs evolved into four to six different isoforms. For *Hypsibius dujardini*, it is unclear whether these eight putative FMOs could be playing the conserved biological role of the isoforms described in vertebrates.

Using synthetic genes, both above-mentioned enzymes could be overexpressed in *E. coli* at a moderate level (around 15 mg L$^{-1}$). To obtain CbFMO in a soluble form, it had to be purified in the presence of additives and a detergent. Furthermore, to prevent aggregation of CbFMO, 350 mM NaCl was required during its purification, as well as 30 μM FAD. This indicated that this bacterial enzyme is rather difficult to obtain and handle. This is also reflected in its relatively low $T_M^{app}$ value of 34 °C. CbFMO exhibited a relatively narrow substrate scope, mainly accepting cyclic ketones. Taking in consideration the phylogenetic clustering with FMO-E, its low activity with NADH was somewhat unexpected. Yet, using NADPH, CbFMO could fully oxidize phenylacetone, *rac*-bicyclo[3.2.0]hept-2-en-6-one and 4-hydroxyacetophenone while also some other ketones could be converted (40-80 %). Relatively low total turnover numbers were obtained (40-100), except for cyclobutanone, where CbFMO reached a total turnover number of 1,400 (the reaction was performed with 10 mM substrate). When considering the relatively low oxidation rates and its poor stability, CbFMO compares poorly with other known BVMOs such as PAMO, TmCHMO or PockeMO[26,38–40]. It is interesting to note that the other studied type II FMO, FMO-E, also was suffering from a relatively low stability and activity[7,12,41].

HdFMO was somewhat easier to handle when compared with CbFMO. It exhibited a $T_M^{app}$ of 45 °C, which is close to the value found for the sequence-related bacterial NiFMO[15]. Still, compared with other described class B flavoprotein monooxygenases, its thermostability is rather low[26,40]. HdFMO displayed a narrow substrate scope and low oxidation rates. Interestingly, only sulfide could be identified as substrates. Their conversions resulted in the formation of the respective sulfoxides as product while no overoxidation (formation of sulfones) was observed. It is worth noting that no formation of indigo blue was observed in the cell culture during expression of HdFMO (or CdFMO) as has been reported for other FMOs such as NiFMO[12,42]. Besides, a substrate screen using ten N-containing molecules did not show any activity (data not shown).

The motivation for studying CbFMO was partly because of the relaxed nicotinamide cofactor specificity of previously reported type II FMOs. Yet, CbFMO was found to have a strong preference for NADPH. When compared with known class B flavoprotein monooxygenases, CbFMO is most closely related in sequence with FMO-E. It also contains the type II-identifying motif. When comparing the sequence region of this motif for CbFMO and FMO-E, there is only one difference: CbFMO has a phenylalanine (Phe329) instead of a tyrosine. Nevertheless, based on a modeled structure, this residue does not interact with either NAD(P)H or FAD. Interestingly, CbFMO shows at the NADPH binding site a tyrosine (Tyr541) and an arginine (Arg452) close to the position of where the 2'-phosphate of the cofactor is expected to bind (**Figure 3a**). In this position, FMO-E contains a phenylalanine and a methionine, while PsFMO-A (another known type II FMO) contains a isoleucine and leucine[12,13,28]. Evidently, the positive charge of Arg452 could interact with the phosphate moiety of NADPH, resulting in a preference for NADPH over NADH.

HdFMO shows significant sequence homology with several previously studied type I FMOs. Using mFMO, a homology model could be prepared. The modeled structure suggests that the active site should be able to accommodate relatively large substrates (**Figure 3b**). This is mainly due to the fact that the enzyme contains two relatively small residues in the substrate binding pocket when compared with the mFMO crystal structure (Thr249 and Ala360 in HdFMO versus Tyr207 and Trp319 in mFMO). However, the substrate (sulfides) profile was similar to mFMO but with lower activities[8]. For mFMO, Trp319 and Trp400 were found to limit the size of the substrate binding pocket[43]. The role of the tyrosine in mFMO has also been studied. Replacing it with an asparagine resulted in a diminished activity against indole. Clearly, the substrate binding pocket of HdFMO is different when compared with other FMOs, which may hint to a fundamentally different substrate specificity. It may be that the tested set of compounds did not resemble any of its natural substrates which would explain the poor catalytic performance on the identified substrates.
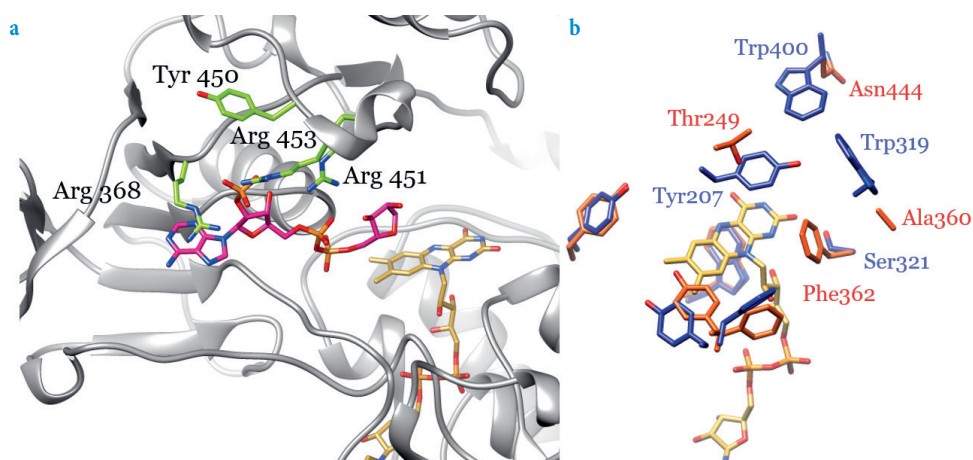
**5**

**Figure 3. Homology models of CbFMO and HdFMO.** (a) CbFMO NAD(P)H binding pocket is shown. (b) Active site of mFMO (blue) and HdFMO (red) are aligned.

While this study has resulted in the identification of two novel dimeric FMOs, these monooxygenases do not appear appealing for biocatalysis. Especially CbFMO was found to be a labile enzyme and showed poor Baeyer-Villiger oxidation activity with a number of ketones. HdFMO displayed somewhat better (thermo)stability features, but was restricted to low sulfoxidation activities. Protein engineering could possibly improve the biocataytic properties of both enzymes. Yet, such approach would ideally make use of structural information. While homology models can be prepared, accurate crystal structure would be more reliable. Unfortunately, crystallization trails with both FMOs failed despite many efforts. Therefore, in order to enlarge the biocatalytic toolbox of FMOs, it is probably more effective to explore other genomes for potential candidates. Or, alternatively, one could opt for engineering one of the known robust FMOs.

## MATERIAL AND METHODS

### Genome mining and sequence analysis

For the identification of the putative FMOs, the NCBI server was used by a protein-protein BLAST search. The sequences of NiFMO and FMO-E were used as templates to find putative homologs. Sequences with high query and identity were subsequently aligned. Multiple sequence alignments (MSA) were prepared using MEGA X software by MUSCLE protein alignment. The MSA was configured with default settings for highest accuracy and employing the UPGMA clustering method. A phylogenetic tree was reconstructed using the maximum likelihood (ML) method implemented in MEGA X (500 bootstrap replications). Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-

Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (4 categories (+G, parameter = 2.7354)). This analysis involved 50 amino acid sequences. There was a total of 1254 positions in the final dataset. Evolutionary analyses were conducted in MEGA X[44].

## Cloning, expression and purification

CbFMO- and HdFMO-encoding DNA sequences were codon optimized for *E. coli* and obtained from Integrated DNA technologies (IDT). The synthetic genes were cloned by Golden Gate Assembly in pBAD vectors encoding an N-terminal SUMO fusion protein equipped with histidine tag. The obtained plasmids were transformed into chemo-competent *E. coli* NEB 10β cells (New England Biolabs) by heat shock. For purification, cells harboring the verified plasmids were inoculated in LB media supplemented with 50 μg mL$^{-1}$ ampicillin at 37 °C overnight. An aliquot of the preculture (1:100) was used to inoculate 50 mL TB media (50 μg mL$^{-1}$ ampicillin). Cultures were grown at 37 °C with continuous shaking until OD$_{600nm}$ 0.7 was reached. For expression, conditions were optimized employing different L-arabinose concentrations (0.2, 0.02 or 0.002 % w/v) and temperatures (17, 24, 30 or 37 °C) for 16, 24, 48 and 72 h. Subsequently, cells were centrifuged at 6,000 r.p.m. for 20' at 4 °C. The supernatant was discarded and cells were suspended in lysis buffer (50 mM TRIS pH 8.0 buffer, 500 mM NaCl, 1.0 mM PMSF, 30 μM FAD and 1.5 mg mL$^{-1}$ lysozyme). After sonication (amplitude 70 %, 10 s on and 10 s off for 10' on ice), cell debris was removed by centrifugation (10,000 RPM for 1h at 4 °C). Expression was verified by SDS-PAGE analysis of cell free extracts (CFE) and insoluble fractions. For CbFMO, high expression levels in the insoluble fraction were observed and additives were added to test for the best condition for obtaining soluble protein (10% v/v glycerol, 1 mM DTT and 100 μM IGEPAL). After obtaining optimized expression conditions according to SDS-PAGE analysis, the expressions were up-scaled to 400 mL. After sonication, lysates were filtered at 0.45 μm and loaded on a nickel-sepharose HP resin (GE Health care) pre-equilibrated with 50 mM TRIS pH 8.0, 500 mM NaCl. Then, filtered CFEs were incubated with the resin for 1 h at 4 °C in a rotating system. The resin was washed with 10 CV of buffer (50 mM TRIS pH 8.0 and 500 mM NaCl), followed by 2 CV of buffer supplemented with 20 mM imidazole. Finally, bound proteins were eluted with 400 mM imidazole. Yellow fractions were collected and desalted in 50 mM TRIS, pH 8.0 with 150 mM NaCl for HdFMO and 350 mM NaCl for CbFMO (EconoPac 10DG columns, Bio-Rad). Desalted proteins were flash frozen in liquid nitrogen and stored at -70 °C.

## Enzymatic characterization

To identify the redox state of the flavoenzymes, UV-visible spectra were recorded. Extinction coefficients were calculated by recording the UV-visible spectra of a control sample and a solution of the flavoproteins incubated with SDS 0.1% w/v for 10 min. Then, UV-vis spectra were recorded and protein concentration was determined using the FAD molar extinction

coefficient ($\varepsilon_{450nm}$=11,300 M$^{-1}$ cm$^{-1}$). For CbFMO and HdFMO, $\varepsilon_{450nm}$= 14.4 and 12.5 mM$^{-1}$ cm$^{-1}$ were calculated, respectively.

The apparent melting temperatures ($T_M{}^{app}$) of SUMO-FMOs were analyzed through intrinsic flavoprotein fluorescence at increasing temperatures. In duplicate, a mixture at 1.0 mg mL$^{-1}$ of enzyme was prepared in a 96-well PCR plate. Using an RT-PCR instrument (CFX96-Touch, Bio-Rad), the solutions were exposed to a temperature ramp (25 to 99 °C, by 0.5 °C every 10 s). The fluorescence was measured using a 450–490 nm excitation filter and a 515–530 nm emission filter[45]. The unfolding temperature was calculated as the maximum of the negative derivative of fluorescence against temperature.

The oligomerization state was studied after proteolytic digestion with SUMO protease (1.0 mg mL$^{-1}$ incubated overnight on a rotating system at 4 °C). Protease and SUMO were removed and flavoproteins were desalted. Then, 500 µL of sample was injected into a Superdex 75 10/300 column connected to an AKTA purifier (GE Health care). A calibration curve was prepared using commercial gel permeation standards (Bio-Rad). Subsequently, dynamic lights scattering (DLS) analysis was done in a DynaPro NanoStar instrument (Wyatt Technology). The DLS results were analyzed with dynamics software, version 7 (Wyatt Technology).

Enzyme activity was measured following the oxidation of NAD(P)H by measuring absorbance ($\varepsilon_{340nm}$=6,220 M$^{-1}$ cm$^{-1}$) at 25 °C. Depletion of NAD(P)H was monitored using a microplate reader (Synergy H1 Hybrid multi-mode reader, Biotek) and the reactions were prepared in 96-well plates at 300 µL final volume. For the analysis, the mixtures contained 150 µM NAD(P)H µM, 2.0 µM SUMO-FMO and 1.0 mM substrate, solubilized at final 2.5 % v/v 1,4-dioxane in 50 mM TRIS buffer at pH 8.0 with 150 mM NaCl. All data were analyzed using GraphPad Prism (GraphPad Software) and kinetic parameters were obtained by fitting the obtained results to the Michaelis–Menten equation.

## Bioconversions using HdFMO and CbFMO

Substrate screening for the FMOs was carried out using 6 different substrate mixtures at 500 µM final concentration in 2.5% v/v 1,4-dioxane. Reactions were prepared in 500 µL in 20 mL vials and the mixture contained 30 µM FAD, 150 µM NADH, 150 µM NADPH, 10 mM Na$_2$PO$_3$·5H$_2$O, 5 µM PTDH and 10 µM SUMO-FMOs in 50 mM TRIS and 150 mM NaCl buffer at pH 8.0. The solutions were incubated at 24 or 17 °C at 150 r.p.m. for 24 h. After incubation, the mixtures were extracted three times by mixing one volume of ethyl acetate for 60 s. Subsequently, anhydrous sulfate magnesium was added to the organic solution to remove residual wa.ter. Identification of the reaction performance was done by GC-MS (GCMS-QP2010 Ultra Shimadzu with electron ionization and quadrupole). Separation was achieved using a HP-1MS column. Chromatograms and MS spectra were analyzed

using GCMS solution Postrun Analysis 4.11 (Shimadzu). The library employed for the MS spectra identification was NIST11.

For further bioconversions analysis, 1.0 mM of substrate (2.5 % v/v 1,4-dioxane) was used in the presence of 150 μM NADH or NADPH. SUMO-HdFMO at 2.0 μM and SUMO-CbFMO at 10 μM were used. Reactions were incubated at 4, 17 or 24 °C for 24 h at 150 r.p.m. As external standard, 0.025 % v/v mesitylene was used. GC analysis was carried out for conversion determination using a GCMS-QP2010 Ultra Shimadzu with a HP-1MS column. Chiral analysis was done using a 7890A GC System (Agilent Technologies) equipped with a CP-Chirasil-Dex CB column.

For ^{1}H-NMR analysis, the conversions were carried out at 0.4 mL using 5 μM SUMO-CbFMO, 10 mM cyclobutanone, 150 μM NADPH, 10 mM $Na_2PO_3 \cdot 5H_2O$, 5 μM PTDH and 30 μM FAD in potassium phosphate buffer (pH 8.0). $D_2O$ (1:5) was added to the reactions after 24 h. Subsequently, $D_2O$-treated samples were analyzed by NMR (400 MHz Varian Unity Plus spectrometer).

5

# SUPPORTING INFORMATION CHAPTER V

## Tables

**Table S1.** Substrate scope analysis



**Table S2.** Conversion of phenylacetone using CbFMO at different conditions

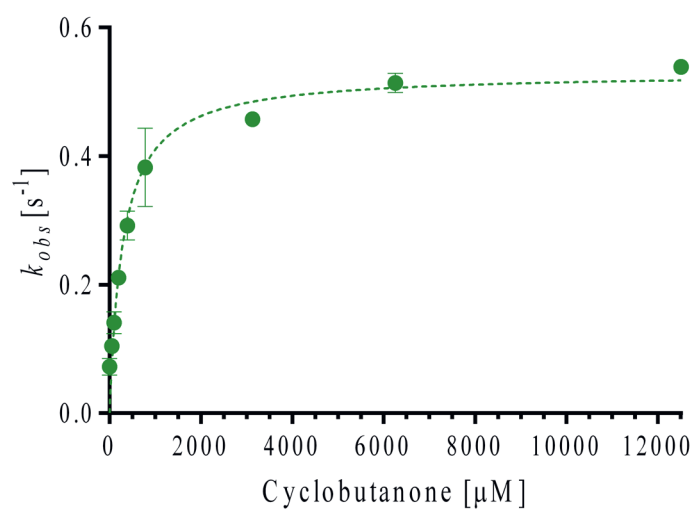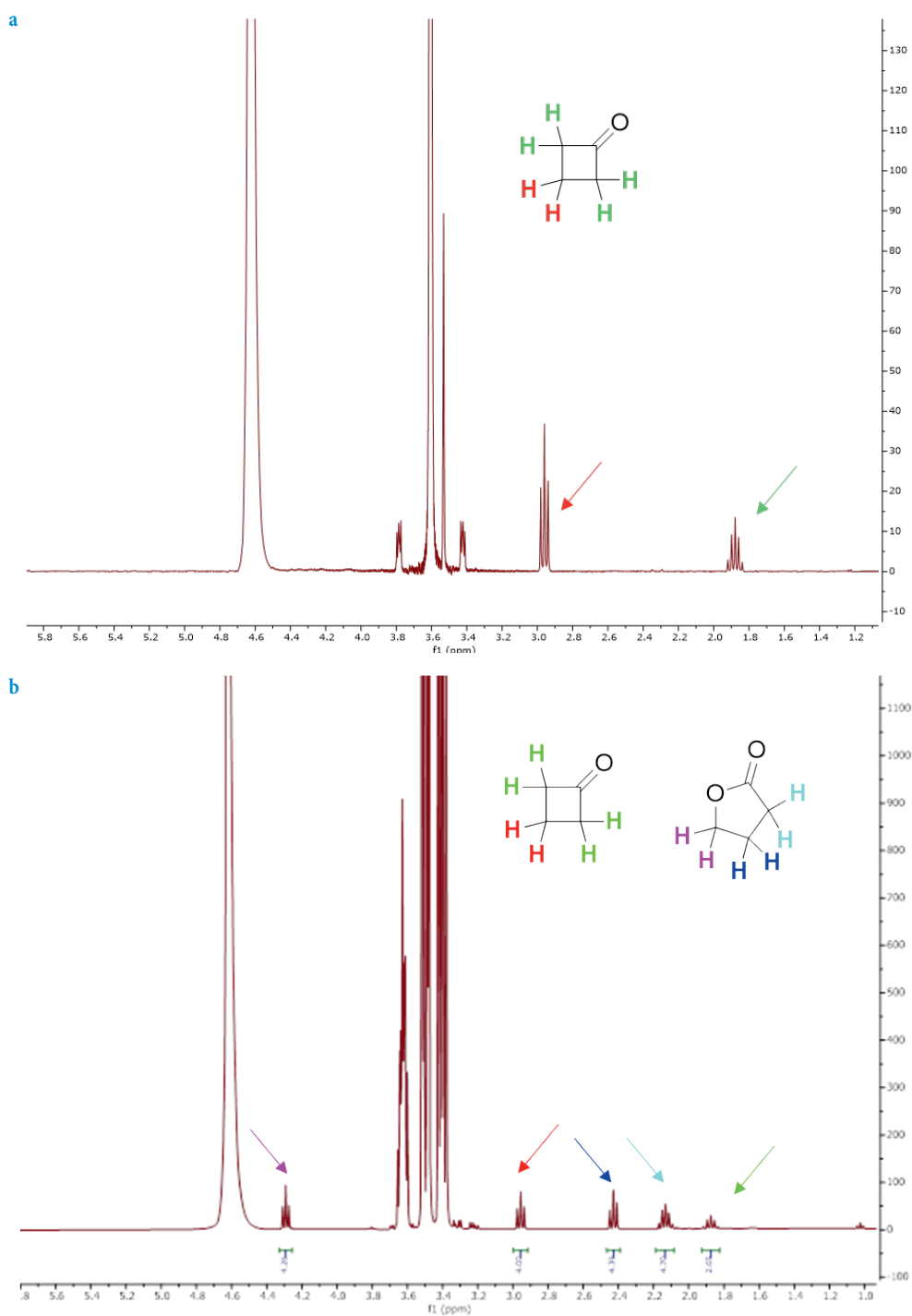| Condition | NADPH [%] | NADH [%] |
|---|---|---|
| 50 mM TRIS pH 8.0, 150 mM NaCl, 17 °C | 99 | 30 |
| 50 mM TRIS pH 8.0, 150 mM NaCl, 24 °C | 20 | - |
| 50 mM TRIS pH 8.0, 150 mM NaCl, 4 °C | 99 | 13 |
| 50 mM TRIS pH 8.0, 350 mM NaCl, 17 °C | 99 | 16 |
| 50 mM KPi pH 8.0, 150 mM NaCl, 17 °C | 99 | 30 |

# Figures



**Figure S1. NADPH consumption rates for CbFMO at increasing cyclobutanone concentrations.** SUMO-CbFMO (2.0 µM), NADPH (150 µM) and substrate were mixed in a microplate containing air-saturated 50 mM TRIS buffer pH 8.0 at 25 °C as described in materials and methods. The data obtained were fit using a Michaelis-Menten equation in GraphPad Prism to determine steady-state kinetic parameters.
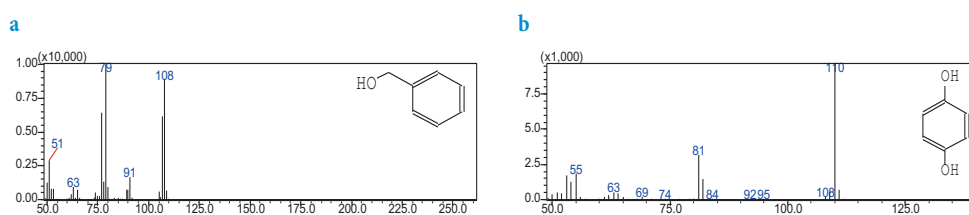
**Figure S2. Conversion of cyclobutanone using CbFMO as biocatalyst.** (a) ¹H-NMR analysis in CDCl₃ at 400 MHz of cyclobutanone as substrate (without enzyme) (b) ¹H-NMR analysis of CbFMO reaction using cyclobutanone in CDCl₃ at 400 MHz.

**Figure S3. MS identification.** (a) Benzylic alcohol and (b) hydroquinone were identified by GC-MS in small scale conversion of 4-hydroxyacetophenone and phenylacetone, respectively.
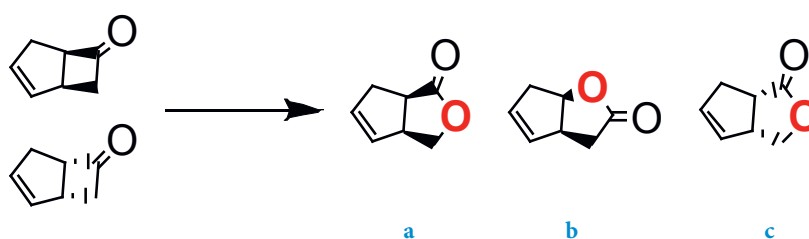


**Figure S4. Enantioselectivity profile of CbFMO.** Bioconversion of *rac*-bicyclo[3.2.0]hept-2-en-6-one using 10 μM CbFMO as catalyst in presence of 150 μM NADPH were performed in 50 mM TRIS buffer pH 8.0, 10 mM Na$_2$PO$_3$•5H$_2$O, 5.0 μM PTDH, 30 μM FAD and 150 mM NaCl. Reactions were analyzed after 24 h at 17 °C. CbFMO mainly produced (a) abnormal (1*R*,5*S*), (c) normal (1*S*,5*R*) and a smaller amount of (b) normal (1*R*,5*S*)

# REFERENCES

1. Savoca, M. P., Tonoli, E., Atobatele, A. G. & Verderio, E. A. M. Biocatalysis by transglutaminases: a review of biotechnological applications. *Micromachines* 9, 562 (2018).

2. Lin, H., Liu, J.-Y., Wang, H.-B., Ahmed, A. A. Q. & Wu, Z.-L. Biocatalysis as an alternative for the production of chiral epoxides: a comparative review. *J. Mol. Catal. B Enzym.* 72, 77–89 (2011).

3. Davis, B. G. & Boyer, V. Biocatalysis and enzymes in organic synthesis. *Nat. Prod. Rep.* 18, 618–640 (2001).

4. Woodley, J. M. New opportunities for biocatalysis: making pharmaceutical processes greener. *Trends Biotechnol.* 26, 321–327 (2008).

5. Bommarius, A. S. Biocatalysis: a status report. *Annu. Rev. Chem. Biomol. Eng.* 6, 319–345 (2015).

6. Van Berkel, W. J. H., Kamerbeek, N. M. & Fraaije, M. W. Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts. *J. Biotechnol.* 124, 670–689 (2006).

7. Huang, L., Romero, E., Ressmann, A.K., Rudroff, F., Hollmann, F., Fraaije, M.W. and Kara, S. Nicotinamide adenine dinucleotide-dependent redox-neutral convergent cascade for lactonizations with type II flavin-containing monooxygenase. *Adv. Synth. Catal.* 359, 2142–2148 (2017).

8. Rioz-Martínez, A., Kopacz, M., De Gonzalo, G., Pazmino, D.E.T., Gotor, V. and Fraaije, M.W. Exploring the biocatalytic scope of a bacterial flavin-containing monooxygenase. *Org. Biomol. Chem.* 9, 1337–1341 (2011).

9. Riebel, A., de Gonzalo, G. & Fraaije, M. W. Expanding the biocatalytic toolbox of flavoprotein monooxygenases from *Rhodococcus jostii* RHA1. *J. Mol. Catal. B Enzym.* 88, 20–25 (2013).

10. Fiorentini, F., Geier, M., Binda, C., Winkler, M., Faber, K., Hall, M. and Mattevi, A. Biocatalytic characterization of human FMO5: Unearthing Baeyer-Villiger reactions in humans. *ACS Chem. Biol.* 11, 1039–1048 (2016).

11. Jensen, C.N., Cartwright, J., Ward, J., Hart, S., Turkenburg, J.P., Ali, S.T., Allen, M.J. and Grogan, G. A flavoprotein monooxygenase that catalyses a Baeyer–Villiger reaction and thioether oxidation using NADH as the nicotinamide cofactor. *ChemBioChem* 13, 872–878 (2012).

12. Riebel, A., Fink, M. J., Mihovilovic, M. D. & Fraaije, M. W. Type II flavin-containing monooxygenases: A new class of biocatalysts that harbors Baeyer–Villiger monooxygenases with a relaxed coenzyme specificity. *ChemCatChem* 6, 1112–1117 (2014).

13. Löwe, J., Blifernez-Klassen, O., Baier, T., Wobbe, L., Kruse, O. and Gröger, H. Type II flavoprotein monooxygenase PsFMO_A from the bacterium *Pimelobacter* sp. Bb-B catalyzes enantioselective Baeyer-Villiger oxidations with a relaxed cofactor specificity. *J. Biotechnol.* 294, 81–87 (2019).

14. Romero, E., Gómez Castellanos, J. R., Gadda, G., Fraaije, M. W. & Mattevi, A. Same substrate, many reactions: Oxygen activation in flavoenzymes. *Chem. Rev.* 118, 1742–1769 (2018).

15. Lončar, N., Fiorentini, F., Bailleul, G., Savino, S., Romero, E., Mattevi, A. and Fraaije, M.W. Characterization of a thermostable flavin-containing monooxygenase from *Nitrincola lacisaponensis* (NiFMO). *Appl. Microbiol. Biotechnol.* 103, 1755–1764 (2019).

16. Choi, H.S., Kim, J.K., Cho, E.H., Kim, Y.C., Kim, J.I. and Kim, S.W. A novel flavin-containing monooxygenase from *Methylophaga* sp. strain SK1 and its indigo synthesis in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* 306, 930–936 (2003).

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

18. Yeoman, C.J., Han, Y., Dodd, D., Schroeder, C.M., Mackie, R.I. and Cann, I.K. Thermostable enzymes as biocatalysts in the biofuel industry. in *Advances in applied microbiology* 70, 1–55 (Academic Press).

19. Sheldon, R. A. & Pereira, P. C. Biocatalysis engineering: the big picture. *Chem. Soc. Rev* 46, 2678 (2017).

20. Cirino, P. C. & Arnold, F. H. Protein engineering of oxygenases for biocatalysis. *Curr. Opin. Chem. Biol.* 6, 130–135 (2002).

21. Kourist, R., Brundiek, H. & Bornscheuer, U. T. Protein engineering and discovery of lipases. *Eur. J. lipid Sci. Technol.* 112, 64–74 (2010).

22. Güven, G., Prodanovic, R. & Schwaneberg, U. Protein engineering–an option for enzymatic biofuel cell design. *Electroanal. An Int. J. Devoted to Fundam. Pract. Asp. Electroanal.* 22, 765–775 (2010).

23. Furuya, T. & Kino, K. Genome mining approach for the discovery of novel cytochrome P450 biocatalysts. *Appl. Microbiol. Biotechnol.* 86, 991–1002 (2010).

24. Furuya, T. & Kino, K. Discovery of 2-naphthoic acid monooxygenases by genome mining and their use as biocatalysts. *ChemSusChem Chem. Sustain. Energy Mater.* 2, 645–649 (2009).

25. Guérard-Hélaine, C., de Berardinis, V., Besnard-Gonnet, M., Darii, E., Debacker, M., Debard, A., Fernandes, C., Hélaine, V., Mariage, A., Pellouin, V. and Perret, A. Genome mining for innovative biocatalysts: new dihydroxyacetone aldolases for the chemist's toolbox. *ChemCatChem* 7, 1871–1879 (2015)

26. Fraaije, M.W., Wu, J., Heuts, D.P., Van Hellemond, E.W., Spelberg, J.H.L. and Janssen, D.B. Discovery of a thermostable Baeyer-Villiger monooxygenase by genome mining. *Appl. Microbiol. Biotechnol.* 66, 393–400 (2005).

27. Gong, J.S., Lu, Z.M., Li, H., Zhou, Z.M., Shi, J.S. and Xu, Z.H. Metagenomic technology and genome mining: emerging areas for exploring novel nitrilases. *Appl. Microbiol. Biotechnol.* 97, 6603–6611 (2013).

28. Riebel, A., Dudek, H.M., De Gonzalo, G., Stepniak, P., Rychlewski, L. and Fraaije, M.W. Expanding the set of rhodococcal Baeyer–Villiger monooxygenases by high-throughput cloning, expression and substrate screening. *Appl. Microbiol. Biotechnol.* 95, 1479–1489 (2012).

29. Kleiger, G. & Eisenberg, D. GXXXG and GXXXA motifs stabilize FAD and NAD (P)-binding Rossmann folds through Cα–H··· O hydrogen bonds and van der Waals interactions. *J. Mol. Biol.* 323, 69–76 (2002).

30. Whelan, S. & Goldman, N. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol.* 18, 691–699 (2001).

31. de Gonzalo, G., Fürst, M. & Fraaije, M. Polycyclic ketone monooxygenase (PockeMO): A robust biocatalyst for the synthesis of optically active sulfoxides. *Catalysts* 7, 288 (2017).

32. Okazaki, Y., Hodoki, Y. & Nakano, S. Seasonal dominance of CL500-11 bacterioplankton (phylum Chloroflexi) in the oxygenated hypolimnion of Lake Biwa, Japan. *FEMS Microbiol. Ecol.* 83, 82–92 (2013).

33. Mehrshad, M., Salcher, M.M., Okazaki, Y., Nakano, S.I., Šimek, K., Andrei, A.S. and Ghai, R. Hidden in plain sight—highly abundant and diverse planktonic freshwater *Chloroflexi. Microbiome* 6, 1–13 (2018).

34. Okazaki, Y., Fujinaga, S., Tanaka, A., Kohzu, A., Oyagi, H. and Nakano, S.I. Ubiquity and quantitative significance of bacterioplankton lineages inhabiting the oxygenated hypolimnion of deep freshwater lakes. *ISME J.* 11, 2279–2293 (2017).

35. Boothby, T.C., Tenlen, J.R., Smith, F.W., Wang, J.R., Patanella, K.A., Nishimura, E.O., Tintori, S.C., Li, Q., Jones, C.D., Yandell, M. and Messina, D.N. Evidence for extensive horizontal gene transfer from the draft genome of a tardigrade. *Proc. Natl. Acad. Sci.* 112, 15976–15981 (2015).

36. Mascotti, M. L., Lapadula, W. J. & Juri Ayub, M. The origin and evolution of Baeyer-Villiger monooxygenases (BVMOs): An ancestral family of flavin monooxygenases. *PLoS One* 10, e0132689 (2015).

37. Huijbers, M. M. E., Montersino, S., Westphal, A. H., Tischler, D. & van Berkel, W. J. H. Flavin dependent monooxygenases. *Arch. Biochem. Biophys.* 544, 2–17 (2014).

38. Fürst, M.J., Savino, S., Dudek, H.M., Gómez Castellanos, J.R., Gutiérrez de Souza, C., Rovida, S., Fraaije, M.W. and Mattevi, A. Polycyclic ketone monooxygenase from the thermophilic fungus *Thermothelomyces thermophila*: a structurally distinct biocatalyst for bulky substrates. *J. Am. Chem. Soc.* 139, 627–630 (2016).

39. Gran-Scheuch, A., Trajkovic, M., Parra, L. & Fraaije, M. W. Mining the genome of *Streptomyces leeuwenhoekii*: Two new type I Baeyer-Villiger monooxygenases from Atacama desert. *Front. Microbiol.* 9, 1–10 (2018).

40. Romero, E., Castellanos, J. R. G., Mattevi, A. & Fraaije, M. W. Characterization and crystal structure of a robust cyclohexanone monooxygenase. *Angew. Chem. Int. Ed. Engl.* 55, 15852–15855 (2016).

41. Huang, L., Aalbers, F.S., Tang, W., Röllig, R., Fraaije, M.W. and Kara, S. Convergent cascade catalyzed by monooxygenase–alcohol dehydrogenase fusion applied in organic media. *ChemBioChem* 20, 1653–1658 (2019)

**5**

42. Fabara, A. N. & Fraaije, M. W. An overview of microbial indigo-forming enzymes. *Appl. Microbiol. Biotechnol.* 104, 925–933 (2020).

43. Lončar, N., van Beek, H. L. & Fraaije, M. W. Structure-based redesign of a self-sufficient flavin-containing monooxygenase towards indigo production. *Int. J. Mol. Sci.* 20, 6148 (2019).

44. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549 (2018).

45. Forneris, F., Orru, R., Bonivento, D., Chiarelli, L.R. and Mattevi, A. ThermoFAD, a Thermofluor®-adapted flavin ad hoc detection system for protein folding and ligand binding. *The FEBS journal*, *276*(10), 2833-2840 (2009)

5

# VI

# Mining the genome of *Streptomyces leeuwenhoekii*: Two new Type I Baeyer-Villiger monooxygenases from Atacama Desert

**Alejandro Gran-Scheuch**, Milos Trajkovic, Loreto Parra and Marco W. Fraaije

# ABSTRACT

Actinobacteria are an important source of commercial (bio)compounds for the biotechnological and pharmaceutical industry. They have also been successfully exploited in the search of novel biocatalysts. We set out to explore a recently identified actinomycete, *Streptomyces leeuwenhoekii* C34, isolated from a hyper-arid region, the Atacama Desert, for Baeyer-Villiger monooxygenases (BVMOs). Such oxidative enzymes are known for their broad applicability as biocatalysts by being able to perform various chemical reactions with high chemo-, regio- and/or enantioselectivity. By choosing this specific Actinobacterium, which comes from an extreme environment, the respective enzymes are also expected to display attractive features by tolerating harsh conditions. In this work, we identified two genes in the genome of *S. leeuwenhoekiii* (*sle_13190 and sle_62070*) that were predicted to encode for Type I BVMOs, the respective flavoproteins share 49 % sequence identity. The two genes were cloned, overexpressed in *E. coli* with phosphite dehydrogenase as fusion partner and successfully purified. Both flavin-containing proteins showed NADPH-dependent Baeyer-Villiger oxidation activity for various ketones and sulfoxidation activity with some sulfides. Gratifyingly, both enzymes were found to be rather robust by displaying a relatively high apparent melting temperature (45 °C) and tolerating water-miscible cosolvents. Specifically, Sle_62070 was found to be highly active with cyclic ketones and displayed a high regioselectivity by producing only one lactone from 2-phenylcyclohexanone, and high enantioselectivity by producing only normal (-)-1*S*,5*R* and abnormal (-)-1*R*,5*S* lactones (*ee* >99 %) from bicyclo[3.2.0]hept-2-en-6-one. These two newly discovered BVMOs add two new potent biocatalysts to the known collection of BVMOs.

**Keywords:** Atacama, Actinobacteria, Baeyer-Villiger monooxygenase, Flavoprotein, Biocatalysis.

# INTRODUCTION

Enzymes are attractive catalysts for several industrial processes by being biodegradable, non-toxic, efficient and selective. These biocatalysts can offer a high level of safety, low energy consumption and a global environmentally friendly process[1]. Enzyme-based approaches often fulfill all the requirements for ecological and economical viable processes[2–5]. The recognition that enzymes can be used in industrially relevant processes is reflected in a predicted growing market for biocatalysts[5,6]. A particular example of enzymes that show industrial potential are Baeyer-Villiger monooxygenases (BVMOs). BVMOs are well-studied enzymes (EC. 1.14.13.XX) that can be used for the production of (enantiopure) esters, lactones and sulfoxides by incorporating an atom of oxygen in an organic substrate releasing a molecule of water using NADPH as cofactor. These enzymes typically display a high chemo-, regio-, and enantioselectivity while operating at mild reaction conditions[7]. In the last years some new Type I BVMOs have been discovered and characterized from different organisms, such as YMOB (*Yarrowia* monooxygenase) from the yeast *Yarrowia lipolytica*, which shows activity on some ketones and sulfides[8], the $BVMO_{AFL706}$ and $BVMO_{AFL334}$ from the fungus *Aspergillus flavus*, which showed a broad substrate acceptance including substituted cyclic, aliphatic and aromatic ketones[9], $BVMO_{Lepto}$ from *Leptospira biflexa*, which was used in whole-cell reactions conversions of various ketones[10] and PockeMO from *Thermothelomyces thermophile*, which displays a high thermostability and shows activity on bulky substrates[11]. However, the BVMOs require special conditions that challenge the application of these biocatalysts on a large scale, like the expensive nicotinamide adenine dinucleotide phosphate (NADPH) as coenzyme. To reduce the costs related to the coenzyme usage, efficient external regeneration systems have been developed. For example, the thermostable phosphite dehydrogenase (PTDH) from *Pseudomonas stutzeri* WM88 can be used to regenerate NAD(P)H[12,13]. Another major issue concerning the application of BVMOs is the poor stability they often display at industrial conditions, like the presence of cosolvent, high temperature and, in some cases, high salinity[3,5]. Generally, enzymes isolated from mesophilic organisms do not tolerate such conditions. Extremozymes, which are enzymes derived from extremophilic organisms, are typically more suited to withstand harsh environments[14]. Currently, there is only one BVMO that can tolerate harsh conditions: phenylacetone monooxygenase (PAMO) from *Thermobifida fusca*[15]. This biocatalyst was obtained by a genome mining approach specifically targeting this mesothermophilic actinobacterium. PAMO was found to be rather thermostable while tolerating cosolvents[16,17]. Recently, other moderately stable BVMOs, isolated from meso-thermophilic microbes, were reported[11,18]. Inspired by these, we considered performing genome mining to another extremophilic actinobacterium to search for novel BVMOs.

The Atacama Desert is a hyper-arid area in the north of Chile, characterized by: a) a large daily temperature variation, where in some areas it ranges from -8 °C to 50 °C[19]; b) low water availability, the area is considered the driest desert in the world[20]; c) exposition to high ultraviolet (UV) light, this zone is characterized by its high altitude, prevalent cloudless conditions and relatively low total ozone column, making this desert one of the highest UV radiation sites on Earth[19,21] and; d) high salinity, the desert contains extremely large natural deposits of anions (as Cl, $ClO_3^-$, $SO_4^{2-}$, $ClO_4^-$ and others). These geographic and environmental characteristics make the microorganisms thriving in the Atacama Desert unique, comprising a genetic and molecular treasure that could lead to novel (bio)chemistry[22–24]. Several microorganisms have been isolated from the Atacama Desert, being an interesting environment to search bacteria with different adaptive qualities to be exploited for biotechnological applications. Among the microorganisms isolated, numerous Actinomycetes have been characterized, including a particular species, *Streptomyces leeuwenhoekii* C34, found to produce novel natural products[25]. This actinobacterium is a Gram-positive mycelial bacterium rich in novel pharmaceutical compounds[26]. *S. leeuwenhoekii* was found in a soil sample, grows from 4 to 50 °C, optimally at 30 °C, from pH 6.0 to 11.0, optimally at 7.0, and in the presence of 10 % w/v sodium chloride. Because of its highly biotechnological potential, this bacterium was sequenced after its discovery[27]. Genomic analysis revealed a 72 % G+C content, the presence of a linear chromosome (8 Mb) and two extrachromosomal replicons, the circular pSLE1 (86 kb) and the linear pSLE2 (132 kb). The *S. leeuwenhoekii* genome contains 35 gene clusters apparently encoding for the biosynthesis of specialized metabolites with potent antibiotic activity such as chaxamycins and chaxalactins[26,28]. Genome mining in *Streptomyces* isolates has already been reported for the identification and characterization of novel BVMOs, including: i) MtmOIV, a BVMO isolated from *S. argillaceus* which is a key enzyme in the mithramycin biosynthesis pathway[29], ii) the BVMOs PenE and PntE forming pentalenolactone precursors in the pathway of antibiotic biosynthesis in *S. exfoliates* and *S. arenae* respectively[30], iii) two BVMOs from *S. coelicolor* acting on thioanisole and a heptanone[31], and iv) PtlE from *S. avermitilis* has also been described to be involved in a pentalenolactone biosynthetic pathway[32]. We performed a search in the predicted proteome of *S. leeuwenhoekii* using the sequence motifs described for Type I BVMOs[33,34]. In this work, we report the discovery, expression and characterization of two novel Type I BVMOs fused to PTDH.

## MATERIALS AND METHODS

### Genome analysis

The GenomeNet server (www.genome.jp/tools/motif/MOTIF2.html) was used for searching proteins that harbor specific sequence motifs (Rossmann fold G-x-G-x-x-

[G/A] and the Type I BVMOs fingerprints [A/G]-G-x-W-x-x-x-x-[F/Y]-[G/M]-x-x-x-D and F-x-G-x-x-x-H-x-x-x-W-[P/D]) using the predicted proteome of *S. leeuwenhoekii* (code: Actinobacteria, Streptomyces, sle). The Uniprot server was used for the identification of the proteins (www.uniprot.org) and the NCBI server for the BLAST searches and identity sequence confirmation (blast.ncbi.nlm.nih.gov/Blast.cgi). Multiple sequence alignments were prepared using 45 protein sequences in MUSCLE software (v3.8.31) configured with default settings for highest accuracy and employing the UPGMB clustering method. The phylogenetic tree was reconstructed using the maximum likelihood (ML) method implemented in MEGA 7.0 (500 bootstrap replications). The default substitution model was selected assuming an estimated proportion of invariant sites and 4 gamma-distributed rate categories to account for rate heterogeneity across sites (WAG model). Nearest-Neighbor Interchange (NNI) ML heuristic method was chosen. Initial tree(s) for the heuristic search were obtained by applying the BioNJ method to a matrix of pairwise distances estimated using a JTT model[35,36].

### Reagents, bacterial strains and plasmids

All chemical reagents were purchased from Sigma-Aldrich, Difco or Merck, unless otherwise stated. Oligonucleotide primers synthesis and DNA sequencing were performed by Macrogen. The genes were amplified by PCR from genomic DNA isolated from *S. leeuwenhoekii* C34. *Escherichia coli* TOP10 (Thermo Fischer) and *E. coli* NEB 10β (New England Biolabs) were used as host strain for recombinant DNA. The pCRE2 vector was used for expressing the target proteins fused to phosphite dehydrogenase (PTDH) equipped with an N-terminal His-tag[12]. Lysogenic broth (LB), terrific broth (TB), and mannitol soya flour media (SFM) were used for bacterial growth[37]. PTDH-PAMO (phenylacetone monooxygenase fused to phosphite dehydrogenase) and PTDH-TmCHMO (*Thermocrispum municipale* cyclohexanone monooxygenase fused to phosphite dehydrogenase) were from GECCO-Biotech.

### Cloning, expression and purification

*S. leeuwenhoekii* was grown in SFM and its genomic DNA was isolated and purified using Purelink® Genomic DNA kit (Invitrogen) according to the recommendations of the manufacturer. Genes encoding the putative enzymes were amplified by PCR using Phusion High-Fidelity DNA Polymerase with the same program: 95 °C-420 s, [95 °C-40 s, 59 °C-40 s-73 °C-120 s] x 35 cycles, 73 °C-600 s and 4 °C-overnight. For the *sle_13190* gene the forward and reverse primers were: 5'-CCT GCG GCT GAC TCG AGA TCT GCA GCT GGT ATG GCC CGC GCC GAA and 5'-TTT TGT TCG GGC CCA AGC TTG GTA ATC TAT GTA TCC TGG TCA GCG CAG TTC GAG GCC, respectively. For the *sle_62070* gene the forward and reverse primers were 5'-CCT GCG GCT GAC TCG AGA TCT GCA GCT GGT ATG ACA CAA GGT CAG ACG TTG TCC and 5'-TTT TGT TCG GGC CCA AGC TTG GTA ATC TAT GTA TCC TGG TCA GCT CAC CGT GGA GCC, respectively. For the *sle_41160*

gene the forward and reverse primers were: 5'-CCT GCG GCT GAC TCG AGA TCT GCA GCT GGT ATG GCC GAG CAC GAG CAT and 5'-TTT TGT TCG GGC CCA AGC TTG GTA ATC TAT GTA TCC TGG TCA CGC GGT CAC CCC . For the pCRE2 amplification the primers were 5'-CCA GGA TAC ATA GAT TAC CAA GCT TGG GCC CGA ACA AAA AC and 5'-ACC AGC TGC AGA TCT CGA GT. The PCR conditions were optimized to a final concentration of 3 % DMSO and 125 nM of each primer. Purified PCR products were cloned into the pCRE2 vector by Gibson assembly[38]. Products were used directly for transformation of competent *E. coli* TOP10 cells. Colonies were grown on LB-agar plates supplemented with ampicillin at 37 °C. Plasmids were isolated (Wizard Plus SV Minipreps DNA Purification System) and sequenced for cloning confirmation (Macrogen). Verified plasmids were transformed in competent *E. coli* NEB 10β used for protein expression. For purification, a single colony was taken for growing a preculture in LB at 37 °C overnight. An aliquot of the preculture (1:100) was used to inoculate fresh TB medium supplemented with 50 µg mL$^{-1}$ ampicillin. Cultures were incubated at 37 °C with shaking until an OD$_{600}$ of 0.7 was reached after which expression was induced by adding L-arabinose. To optimize the expression, different inducer concentrations (0.002 %, 0.02 % and 0.2 %) and temperatures (17 °C, 24 °C, 30 °C and 37 °C) for 16, 24 and 48 h were tested. Cells were harvested by centrifugation (6,000 x 20' at 4 °C using a JLA-9100 rotor, Beckman Coulter) and suspended in lysis buffer (50 mM Tris HCl pH 7.0, 10 % w/v glycerol, 1.5 mg mL$^{-1}$ lysozyme, 10 µM FAD and 1 mM PMSF). Cell extracts (CE) were obtained by sonication (Vibra cell, Sonics & materials) for 10' (amplitude 70 %, 7 s on and 7 s off). The cleared cell extracts (CCE) were obtained by centrifugation at 10,000 rpm for 1 h at 4 °C (Centrifuge 5810R, Eppendorf). CCE, CE and insoluble fraction were analyzed by SDS-PAGE to verify expression of the respective BVMOs. After establishing proper expression conditions, CCE was prepared, filtered (0.45 µM) and loaded on 3 mL of nickel sepharose HP (GE Health Care) pre-equilibrated with buffer and incubated for 1 h at 4 °C in a rotating system. Then, the column was washed with ten column volumes of buffer (50 mM Tris HCl pH 7.0, 10 % glycerol, 0.5 M NaCl) followed by two column volumes of 50 mM Tris HCl pH 7.0, 10 % glycerol, 0.5 M NaCl and 5 mM imidazole. The protein was eluted using buffer with 500 mM imidazole. Fractions containing yellow protein were loaded on a pre-equilibrated EconoPac 10DG desalting column (Bio-Rad). The final sample was flash frozen with liquid nitrogen and stored at -80 °C. The purity of each purified enzyme batch was confirmed by SDS-PAGE analysis.

## Fluorescence and spectrophotometric analysis

To determine the protein concentration based on FAD content, samples were diluted until an absorbance of around 0.5 at 440 nm. After collecting a full UV-vis spectrum, sodium dodecyl sulfate (SDS) was added to a final concentration of 0.1 % w/v. An UV-vis spectrum was recorded again after 10 min. The spectrum obtained with SDS was used to determine the FAD concentration ($\varepsilon$=11,300 M$^{-1}$ cm$^{-1}$), and the extinction coefficient for the protein by comparison with the spectrum of the native protein[15,18].

The apparent melting temperatures ($T_M$') were determinate by using the ThermoFAD method[39]. For this, 20 µl samples were prepared in a 96-well PCR plate. The samples contained 1 mg mL$^{-1}$ enzyme in different buffered solutions: 50 mM BisTris HCl, 50 mM Tris HCl or 10 mM CAPS NaOH buffer adjusted at desired pH, cosolvents, and other additives. The plate was heated from 20 °C to 99 °C, increasing the temperature by 0.5 °C every 10 seconds, using an RT-PCR instrument (CFX96-Touch, Bio-Rad). By measuring fluorescence using a 450–490 nm excitation filter and a 515–530 nm emission filter, the $T_M$' or unfolding temperature was determined as the maximum of the derivative of the sigmoidal curve.

Enzyme activity was screened by measuring the oxidation of NADPH at 340 nm (ε=62,220 M$^{-1}$ cm$^{-1}$) in 96-well F-bottom plates (Greiner BioOne GmbH) at 25 °C using a SynergyMX micro-plate reader (BioTek). The reactions were performed in 200 µL containing 10 % glycerol, 50 mM BisTris HCl, 50 mM Tris HCl or 10 mM CAPS at the desired pH, cosolvent concentration, and additive, 0.45 µM of purified BVMO, 150 µM NADPH and 5.0 mM phenylacetone. As a control, to measure NADPH consumption in the absence of substrate (uncoupling activity), phenylacetone was omitted. For determining kinetic parameters, activity was analyzed using a V-660 spectrophotometer (Jasco) using a 100 µL quartz cuvette at 25 °C, and NADPH oxidation at 340 nm was followed. The reaction mixture contained 50 mM Tris HCl pH 8.0, 10 % glycerol, 100 mM NaCl, 0.45 µM of purified BVMO, 150 µM NADPH and substrate solubilized in 1,4-dioxane (2.5 % v/v final concentration). The control reaction contained no substrate and the same concentration of cosolvent. The reaction was started by adding the nicotinamide cofactor and mixing, after which the absorbance was measured for 60 s.

## Biotransformation studies

The substrate scope analysis was performed using 6 different substrate mixtures (400 µM final concentration of each substrate and 2.5 % v/v 1,4-dioxane). For transformations, the reaction mixture was prepared in 500 µL containing Tris HCl pH 7.0, 10 % glycerol, 100 mM NaCl, 30 µM FAD, 10 mM Na$_2$PO$_3$•5H$_2$O, 150 µM NADPH and 2.7 µM of purified enzyme in a 20 mL glass vial. The mixture was subsequently shaken at 150 rpm and 24 °C for 24 h. The mixture was extracted three times by mixing one volume of ethyl acetate for 60 s. Subsequently, anhydrous sulfate magnesium was added to the organic solution to remove residual water. Analysis was carried out using a GC-MS QP2010 ultra (Shimadzu) with electron ionization and quadrupole separation with a HP-1 column. The temperature program, column data and injection volumes are in the supplementary data file (**Table S1**).

After identifying the positive substrates and optimizing the conditions, biotransformations with single substrates were carried out in buffer Tris HCl pH 8.0 using the same concentration for the additives, enzyme, cofactor and phosphite described above. The conversion of

racemic bicyclo[3.2.0]hept-2-en-6-one, phenylacetone, 2-phenylcyclohexanone and 4-phenylcyclohexanone was performed using a final substrate concentration of 5.0 mM. For benzoin, a concentration of 1.0 mM was used. The reaction mixtures were incubated at 24 °C and 150 rpm for 2 and 24 h. After incubation, samples were extracted three times with one volume of tert-butyl methyl ether containing 0.1 % v/v mesitylene as an external standard and vortexed for 1 minute. Then, the GC analysis was performed in the same instrument described above, for chiral analysis of bicyclo[3.2.0]hept-2-en-6-one, 2-phenylcyclohexanone and 4-phenylcyclohexanone the substrates were analyzed using a 7890A GC System (Agilent Technologies) equipped with a CP Chiralsil Dex CB column. The enantiomers were identified by comparing with reported retention times and biocatalytic preference[18,40–42].

### ¹H-NMR analysis

For ¹H-NMR analysis, the reactions were carried out at 4 mL for 2-phenylcyclohexanone (10 mM) and 10 mL for benzoin (1.0 mM). Extraction was performed three times with ethyl acetate, dried over anhydrous sulfate magnesium and concentrated by rota-evaporation. The extracts were suspended in 1 mL $CDCl_3$ and NMR analysis was performed in a 400 MHz Varian Unity Plus spectrometer.

### Statistical analysis

All analyses were performed using GraphPad Prism v6.05 for Windows (GraphPad Software, La Jolla, USA). To assess statistical significant differences between more than two groups of data, a two-way ANOVA test was used, with the Tukey post-test used to compare each different group, using a $p < 0.05$. Kinetic parameters were obtained by fitting the obtained data to the Michaelis-Menten equation. Chromatograms and MS spectra were analyzed using GCMSsolution Postrun Analysis 4.11 (Shimadzu). The library for the MS spectra was NIST11.

## RESULTS

### Genome analysis and molecular cloning

By using the sequence motif for Rossmann fold (GxGxxG/A) and two previously described Type I BVMO-specific sequence motifs ([A/G]GxWxxxx[F/Y]P[G/M]xxxD and FxGxxxHxxxWP/D)[33,34], we could identify three putative BVMOs in the predicted proteome of *S. leeuwenhoekii* C34: Sle_41160, Sle_13190 and Sle_62070 (Uniprot codes A0A0F7VV32, A0A0F7VUW7 and A0A0F7W6X7, respectively) (**Table S2**). A sequence alignment analysis revealed that Sle_13190 displays 92 % sequence identity with PntE (pentalenolactone D synthase from *S. arenae*)[30] while, Sle_62070 only has 50 % sequence identity with PntE. Another known BVMO that is closely related in sequence with

Sle_62070 is PockeMO (Polycyclic Ketone Monooxygenase, 39 % seq. ident.) from the fungus *T. thermophile*[11]. These two putative BVMOs share 49 % of sequence identity. On the other hand, Sle_41160, has 36 % sequence identity with HAPMO (4-hydroxyacetophenone monooxygenase from *P. fluorescens*)[33] and shares around 30 % sequence identity with the other two predicted BVMOs. A phylogenetic molecular analysis was inferred using Maximum likelihood[35]. The resultant tree revealed that Sle_13190 and Sle_62070 belong to a distinct clade of Type I BVMOs (**Figure S1**). Based on the recently elucidated crystal structure of PockeMO, it has been reported that this group of BVMOs have a special structure feature in common that allows them to accommodate relatively large substrates in their active site[11]. This suggests that Sle_13190 and Sle_62070 should have the capacity to deal with bulky compounds. On the other hand, Sle_41160 was clustered close to HAPMO, a Type I BVMO described to catalyze the reaction of 4-hydroxyacetophenone to the corresponding acetate ester[33].

The *sle_13190*, *sle_62070* and *sle_41160* genes have 69, 71 and 73 % of G+C % content which is similar to the chromosomal DNA of *S. leeuwenhoekii*[27]. We amplified the three genes from the isolated genomic DNA after optimizing PCR conditions. Thereupon, we cloned them into a pCRE2 vector that harbors a NADPH-recycling phosphite dehydrogenase as a N-terminal fusion partner with a N-terminal histidine tag. The generated expression plasmids were used to transform *E. coli* TOP10, the subsequent expression of soluble protein was tested at various temperatures. The best results for Sle_13190 and Sle_62070 were obtained when expression was performed at 17 °C for 48 h using 0.02 % of L-arabinose. For Sle_41160, no expression could be obtained at any of the tested conditions and therefore was discarded for further experiments. The proteins were purified through $Ni^{+2}$-affinity chromatography in one step, a clear yellow color was indicative of proper folding and FAD binding. The purified proteins displayed UV-vis spectra that are characteristic for flavin-containing proteins displaying a maxima absorbance at 385 nm and 440 nm (**Figure S2**). Using SDS as unfolding agent, the extinction coefficient at 450 nm of each flavoprotein was determined: 14.1 and 15.7 $mM^{-1} cm^{-1}$ for Sle_13190 and Sle_62070, respectively.

## Characterization of the Atacama BVMOs

To obtain a better view on the biochemical properties of the two purified flavoenzymes, their thermostability and tolerance towards cosolvents were studied. The ThermoFAD method was used to probe their thermostability. This method determines the apparent melting temperature ($T_M$') of a flavoprotein based on the increase in flavin fluorescence when the flavin cofactor is released upon protein unfolding[39]. First, we determined the $T_M$' of the enzymes at various pH values using different buffers (Tris HCl, Bis Tris HCl or CAPS) in the presence of 100 mM NaCl and 10 % w/v glycerol. Interestingly, both flavoproteins display a similar pH-dependent unfolding profile (**Figure 1a**). For Sle_13190 and Sle_62070, the $T_M$' was around 45 °C between pH 7.5 and 10.0, showing that the two enzymes are
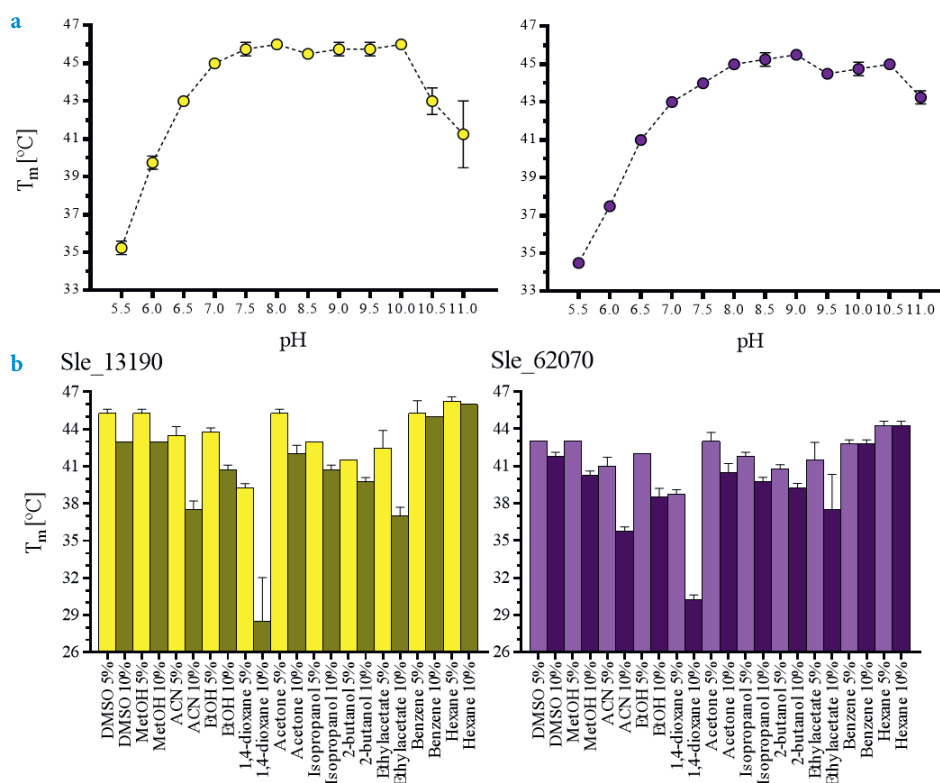
**Figure 1. Determination of apparent melting temperature of *S. leeuwenhoekii* Type I BVMOs.** (a) The thermostability of Sle_13190 (yellow dots) and Sle_62070 (purple dots) were measured at different pH (5.0-11.0). (b) The effect in the $T_M$' by the presence of cosolvents at 5 % v/v (light column) and 10 % v/v (dark column) was analyzed for Sle_13190 and Sle_62070 in Tris HCl pH 7.0.

relatively thermostable. To discard a possible buffer composition effect, HEPES at pH 7.0-7.5 and citrate buffer at pH 6 were also tested, resulting in highly similar $T_M$' values. Subsequently, the effect of several commonly used cosolvents on the thermostability was explored by analyzing samples containing 5 and 10 % v/v of DMSO, methanol, acetonitrile (ACN), ethanol, 1,4-dioxane, acetone, isopropanol, 2-butanol, ethyl acetate, benzene or hexane (**Figure 1b**). Again, both flavoenzymes displayed similar patterns of solvent tolerance. For both enzymes, a major deleterious effect was observed with 10 % ACN, ethyl acetate and 1,4-dioxane, resulting in a drop of 7-8 °C with ACN and ethyl acetate, and a drop of >10 °C with 1,4-dioxane. The data suggest that the enzymes can be employed in the presence of various solvents. The effect of increasing concentrations of NaCl, ectoine, 5-hydroxyectoine and proline was also analyzed. These additives were chosen because the BVMOs may have evolved to operate in the presence of high concentrations of these

compounds as *S. leeuwenhoekii* thrives in a highly salty environment. While Sle_13190 was rather insensitive to increasing amounts of NaCl, Sle_62070 was more stable at lower concentrations (**Figure S3**). For the other additives, no significant differences in effects on the $T_M$' values were observed. While increasing amounts of ectoine resulted in lower $T_M$' values, 5-hydroxyectoine (0-200 mM) and proline (0-4 mM) did not have a significant effect (**Figure S3**).

## Substrate profiling of the Atacama's BVMOs

After the thermostability analysis, the substrate scopes for both flavoenzymes were studied through a high-throughput GC-MS analysis approach by verifying product formation for each potential substrate. Each enzyme was incubated with 6 different mixtures containing 3-6 distinct ketones and thioethers at a final concentration of 400 µM each and 2.5 % 1,4-dioxane as cosolvent. In this way, a total of 30 different potential substrates were tested with each BVMO (Fig. S4). For regeneration of the nicotinamide coenzyme, the fusion partner of the recombinant enzymes, phosphite dehydrogenase, was exploited by including phosphite and a catalytic amount of the coenzyme. The conversions were incubated for 24 h at 24 °C and after extraction the analysis revealed a broad substrate acceptance for both enzymes. For Sle_13190, 15 compounds were identified as substrate through detection of formed substrate, while Sle_62070 was found to convert 17 of the 30 tested compounds (**Table 1**). The substrate profiles include several cyclic aliphatic ketones, aromatic ketones and sulfides, linear aliphatic ketones and also a steroid, stanolone. The two BVMOs shared most of the uncovered substrates but also some striking differences were noted. For example, of the tested octanones, Sle_13190 converted 2-octanone, and Sle_62070 converted 3- and 4-octanone (the predicted products are included in the **Figure S5**)

To determine the optimal conditions and robustness of the potential biocatalysts, phenylacetone was selected as model substrate as it was found to be well-accepted substrate (**Table 1**). We measured the rate of NADPH consumption in the absence or presence of 5.0 mM of this ketone at various pH values (**Figure 2**). The data confirm that Sle_62070 has a good activity with phenylacetone as substrate, and a significantly higher activity with the substrate over the uncoupling activity (consumption of NADPH in absence of substrate). The highest activity (around 2.4 s$^{-1}$) was observed at pH 7.5–10. Sle_13190 has a relatively low activity on phenylacetone (around 0.24 s$^{-1}$ at pH 7.0-9.5), which is only slightly higher when compared with its uncoupling activity. Based on these activity results we chose a pH value of 8.0 for the subsequent experiments. We also compared the effect of seven water-miscible solvents at 2.5, 5 and 10 % v/v (**Figure S6**). This revealed that DMSO is probably a substrate for both BVMOs. For Sle_62070 the rate of NADPH consumption increased significantly at higher DMSO concentrations while for Sle_13190 the observed rate with DMSO were even equal to the rates in the presence phenylacetone. BVMO

activity on DMSO has also been noted before and therefore is not a suitable cosolvent[8]. Interestingly, up to 10 % v/v of the other six cosolvents seem to be compatible with both BVMOs, even at the highest concentration these cosolvents did not significantly affect the observed activities. While the BVMO activities were in the same range when compared in buffer, only some modest increase in uncoupling activity was seen in the presence of ACN, methanol and isopropanol. This demonstrates that both biocatalysts are rather tolerant towards regularly used cosolvents.

**Table 1. Substrate scope analysis of Type I BVMOs**. The 30 compounds tested in the substrate scope analysis are shown with their respective CAS number. Conversion was determined semi-quantitatively for each substrate by analysis of the peak areas in the GC chromatograms normalized by the compound(s) of the mixture that were not accepted by the enzyme. The results are categorized by the observed degree of conversion as 100-81 %, +++++; 80-61 %; ++++; 60-41 %, +++; 40-21 %, ++; < 20, + or <1 %, - .

| Mix | Substrate | CAS number | | Sle_13190 | Sle_62070 |
|---|---|---|---|---|---|
| 1 | 2-hexylcyclopentanone | 13074 | 65-2 | +++++ | ++++ |
| 1 | 3-methyl-2,4-pentanedione | 815 | 57-6 | - | - |
| 1 | benzylphenyl sulfide | 831 | 91-4 | - | +++ |
| 1 | cycloundecanone | 878 | 13-7 | - | - |
| 1 | indole | 120 | 72-9 | - | - |
| 1 | phenylacetone | 103 | 79-7 | +++++ | +++++ |
| 2 | 2-propylcyclohexanone | 94 | 65-5 | +++++ | ++ |
| 2 | 3-octanone | 106 | 68-3 | - | ++ |
| 2 | bicyclo[3.2.0]hept-2-en-6-one | 13173 | 09-6 | +++++ | +++++ |
| 2 | cyclododecanone | 830 | 13-7 | - | - |
| 2 | cyclopentanone | 120 | 92-3 | - | - |
| 2 | methyl-p-tolyl sulfide | 1519 | 39-7 | ++++ | ++++ |
| 3 | 2-phenylcyclohexanone | 1444 | 65-1 | ++++ | +++++ |
| 3 | androst-1,4-diene-3,17-dione | 897 | 06-3 | - | - |
| 3 | cyclopentadecanone | 502 | 72-7 | + | + |
| 3 | nicotin | 54 | 11-5 | - | - |
| 3 | vanillylaceton | 122 | 48-5 | - | - |
| 4 | 4-hydroxyacetophenone | 99 | 93-4 | - | - |
| 4 | 4-phenylcyclohexanone | 4894 | 75-1 | +++++ | +++++ |
| 4 | androst-4-ene-3,17-dione | 63 | 05-8 | ++++ | ++ |
| 5 | 4-octanone | 589 | 63-9 | - | + |
| 5 | acetophenone | 98 | 86-2 | - | - |
| 5 | cyclohexanone | 108 | 94-1 | +++ | + |
| 5 | pregnalone | 145 | 13-1 | - | - |
| 5 | thioanisole | 100 | 68-5 | +++ | ++++ |
| 6 | 2-octanone | 111 | 13-7 | + | - |
| 6 | 4-methylcyclohexanone | 589 | 92-4 | ++ | +++ |
| 6 | benzoin | 119 | 53-9 | ++ | +++++ |
| 6 | cyclooctanone | 502 | 49-8 | - | - |
| 6 | stanolone | 521 | 18-6 | ++ | +++++ |
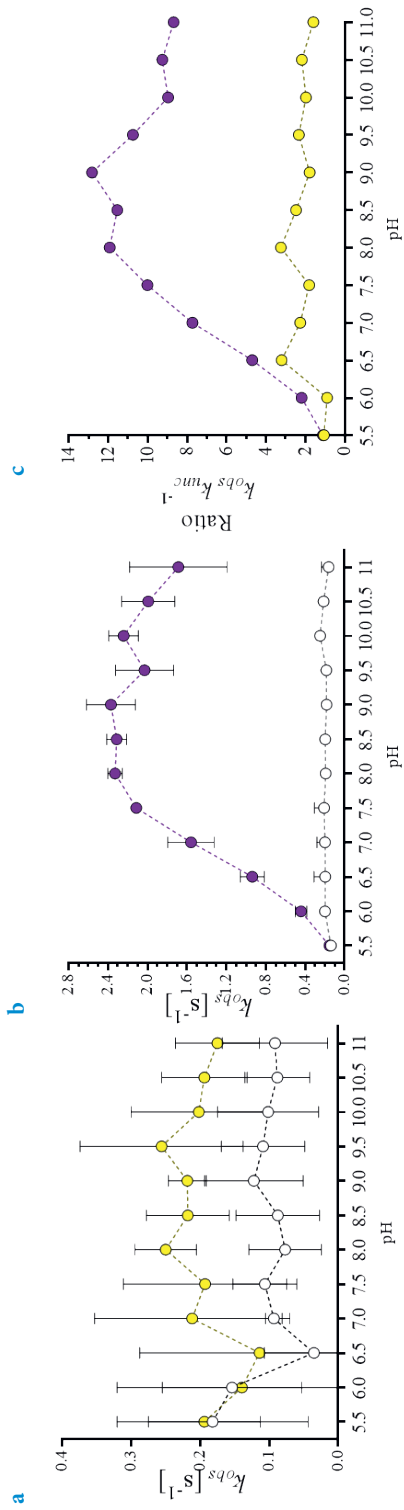
**Figure 2. Effects of pH on BVMO and NADPH oxidase activities.** (a,b) The $k_{obs}$ in the presence (color dots) and absence of 5.0 mM phenylacetone (white dots) was analyzed at pH 5.0–11.0 for Sle_13190 (a) and Sle_62070 (b). (c) The ratio $k_{obs}$ $k_{unc}$ $^{-1}$ at different pH (5.0–11.0) was calculated for Sle_13190 (yellow dots) and Sle_62070 (purple dots).

In order to test the effects of known microbial osmoprotectants we studied their effects on Sle_62070 because this BVMO seemed to exhibit better kinetic properties. We chose 1,4-dioxane (2.5 % v/v) as cosolvent. The effect of increasing concentrations of ectoin, 5-hydroxyectoin (0-200 mM) and proline (0-4 M) on BVMO activity of Sle_62070 was studied (**Figure S7**). None of the osmoprotectants exerted a dramatic effect on the BVMO or uncoupling activities. A slight boost (20 %) on BVMO activity was observed with 200 mM ectoin, while 4 M proline decreased the BVMO activity by 30 %. To have a better view on the kinetic properties of Sle_62070, the initial rates of NADPH consumption with a set of ketones and sulfides were determined (**Figure S8a**). These kinetic data revealed a relatively high $k_{obs}$ with bicyclo[3.2.0]hept-2-en-6-one, phenylacetone, 2-phenylcyclohexanone and 4-phenylcyclohexanone, and a lower activity with the thioethers methyl-p-tolyl sulfide, benzylphenyl sulfide and thioanisole and the cyclic ketones 2-hexylcyclopentanone and cyclopentadecanone. The other tested compounds did not shown a significant difference in activity when compared with the uncoupling rate. The steady-state kinetic parameters for four substrates on which Sle_62070 displayed a relatively high activity were determined (**Table 2**). As it was found for other Type I BVMOs, Sle_62070 displays a relatively high activity with bicyclo[3.2.0]hept-2-en-6-one and phenylacetone, showing a $k_{cat}$ of 4.0 s$^{-1}$ for both compounds, and $K_M$ values of 0.2 and 3 mM, respectively (**Figure S8b,c**). In addition, Sle_62070 shows high activity with 2-phenylcyclohexanone and 4-phenylcyclohexanone, displaying $k_{cat}$ values of >4.0 s$^{-1}$ and $K_M$ values of >3.0 mM for both phenylcyclohexanones (**Figure S8d,e**). We also attempted to determine kinetic parameters for Sle_13190, but the observed rates were too low for an accurate kinetic analysis. Clearly, using the applied conditions, Sle_62070 is a superior biocatalyst concerning its kinetic properties.

**Table 2.** **Kinetic parameters of Sle_62070**. NADPH oxidation rates were spectrophotometrically followed in 50 mM Tris-HCl at pH 8.0, 10 % glycerol and 100 mM NaCl at 25 °C (enzyme 0.45 µM, NADPH 150 µM) with four ketones in increasing concentrations. The data obtained were fit using a Michaelis-Menten equation in GraphPad Prism.

| Substrate | $k_{cat}$ [s$^{-1}$] | $K_M$ [mM] | $k_{cat} K_M^{-1}$ [s$^{-1}$ mM$^{-1}$] |
|---|---|---|---|
| bicyclo[3.2.0]hept-2-en-6-one | 4.0 ± 0.06 | 0.20 ± 0.01 | 20 |
| phenylacetone | 4.1 ± 0.2 | 3.0 ± 0.3 | 1.3 |
| 2-phenylcyclohexanone | >4.0 | >3.0 | 1.3 |
| 4-phenylcyclohexanone | >4.0 | >3.0 | 1.3 |

After establishing the substrate profiles and kinetic properties of the two newly discovered BVMOs, some substrates were selected as candidates to perform conversions at a larger scale. Racemic bicyclo[3.2.0]hept-2-en-6-one was chosen as a hallmark BVMO substrate for enantio- and regioselectivity, phenylacetone as a well described ketone for BVMOs, and thioanisole to include a thioether for testing a sulfoxidation reaction.

2-Phenylcyclohexanone, 4-phenylcyclohexanone and benzoin were selected as relatively unexplored BVMO substrates. All the compounds were tested at 5.0 mM except for the conversion of benzoin; due its poor solubility it was used at a concentration of 1.0 mM. Upon incubating the substrates with 2.7 μM Sle_62070-PTDH for 2 h, complete conversion was observed with most targeted compounds (**Table 3**). For 2-phenylcyclohexanone merely 69 % was converted and for thioanisole only 18 % conversion was obtained. Extending the incubation to 24 h only resulted in a minor improvement (83 % and 22 % conversion for 2-phenylcyclohexanone and thioanisole, respectively). Sle_13190 was only tested for the conversion of benzoin, resulting in 40 % conversion after 2 h incubation. By GC-MS analysis it was found that both enzymes produce benzaldehyde when converting benzoin. NMR analysis revealed that, except for benzaldehyde, also benzoic acid is formed upon conversion of benzoin (**Figure S9a**). Also, the conversion of 2-phenylcyclohexanone was subjected to $^1$H-NMR analysis. The NMR spectral data revealed the production of the proximal lactone when 2-phenylcyclohexanone was converted by Sle_62070 (**Figure S9b**).

**Table 3. Biocatalysis performance of Sle_62070.** Biotransformation assays were carried out for 2 h at 24 °C for 6 different substrates, consumption percentages are shown. Enantioselectivity was determined for bicyclo[3.2.0]hept-2-en-6-one. N:A, ratio normal and abnormal product, *ee*, enantiomeric excess of the product calculated as $ee = \frac{(Amajor - Aminor)}{(Amajor + Aminor)} x\ 100$ .

| bicyclo[3.2.0]hept-2-en-6-one | $ee_N$ | $ee_A$ | phenyl-acetone | 2-phenyl-cyclohexanone | 4-phenyl-cyclohexanone | benzoin | thioanisole |
|---|---|---|---|---|---|---|---|
| >99 % | >99 % | >99 % | >99 % | 69 % | >99 % | >99 % | 18 % |

To probe the enantioselectivity of Sle_62070, chiral GC analyses were performed. First, the conversion of 4-phenylcyclohexanone was studied. As reference the conversion of 4-phenylcyclohexanone with TmCHMO-PTDH was performed. The reaction using CHMOs is described to produce preferably the *S* lactone[40,43]. The results revealed that Sle_62070 is highly enantioselective towards this ketone as only the *S* lactone was formed (**Figure S10**). Using 2-phenylcyclohexanone as racemic substrate, Sle_62070 was found to be highly enantioselective for convert the *R* enantiomer substrate. The reaction was compared with TmCHMO-PTDH which is described to display a higher preference for the same enantiomer[41,42] (**Figure S10**). The enantioselective properties of Sle_62070 with racemic bicyclo[3.2.0]hept-2-en-6-one as substrate was also analyzed. As reference reaction, the biotransformation of the racemic prochiral cyclic ketone was also performed with PAMO-PTDH resulting in the formation of all four possible lactone products[18]. After 2 h conversion, two of the four possible lactone products, in equal amounts, were observed when using Sle_62070: the (-)-1*S*,5*R* normal lactone and the (-)-1*R*,5*S* abnormal lactone. The enantiomeric excess for both products were determined as >99 % (**Figure S11**). Finally,

the reaction was analyzed in time (**Figure S12ab**) which revealed that Sle_62070 has no preference for one of the two substrate enantiomers.

## DISCUSSION

By genome sequence analysis, we have identified two new actinobacterial BVMOs. As far as we know, these are the first BVMOs described from an Atacama Desert's microorganism. Both BVMOs were shown to be rather robust by tolerating cosolvents up to 10 % v/v and by displaying relatively high melting temperatures. Compared with CHMO from *Acinetobacter* sp. or STMO from *Rhodococcus rhodochrous* (both having a $T_M$' of 39 °C), these two new BVMOs showed a higher $T_M$' (5-6 °C higher). The two BVMOs are similar in thermostability when compared with the recently reported PockeMO ($T_M$' of 47 °C) which was identified from a thermophilic fungus (Fürst et al., 2016). Both uncovered enzymes showed activity on a wide variety of ketones and sulfides, a typical feature of Type I BVMOs. As already can be deduced from the clustering of the sequences based on sequence homology with other BVMOs that are known to act on bulky compounds, Sle_13190 and Sle_62070 also accept rather complex compounds as substrate, including biphenyls and a steroid. Remarkably, even though the substrate scope is similar between these two enzymes, Sle_62070 showed to be more promising by acting on more compounds and by exhibiting higher activities. Sle_62070 also revealed to be highly enantio- and regioselective in converting bicyclo[3.2.0]hept-2-en-6-one into two enantiopure lactones. Interestingly, it is also efficient in converting benzoin which, as far we know, was only reported as substrate for CPMO without any product identification[34]. We have identified for the first time the products formed by BVMO-catalyzed benzoin oxidation: benzaldehyde and benzoic acid. For the production of these compounds, we suggest the formation of a labile ester product which decomposes to from the two aromatic products (**Figure 3**). Overall, this work has delivered two new BVMOs to complement the available collection of known BVMOs. The extreme environment of the Atacama Desert may develop as an interesting source for new robust enzymes. Using metagenomic approaches it should become feasible to tap novel biocatalysts from this rich source of actinobacterial "biosynthetic dark matter", which is unique due to the special soil subsurface geochemistry, ecological diversity and environmental conditions of this hyper-arid desert[22,23].
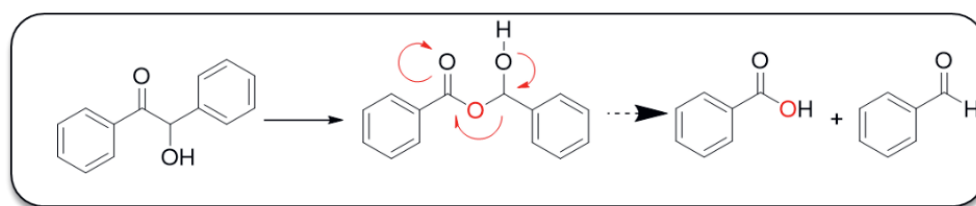
**Figure 3. Proposed reaction of benzoin oxidation by Sle_62070.** Mechanism of the hydrolysis of oxidation product to benzaldehyde and benzoic acid.

## ACKNOWLEDGMENTS

**6**

# SUPPORTING INFORMATION CHAPTER VI

## Tables

**Table S1.** Gas chromatography parameters

| Analysis | Column | GC program | | | Split ratio | Injection [μL] |
|---|---|---|---|---|---|---|
| | | Rate [°C min⁻¹] | Temperature [°C] | Hold time [min] | | |
| Racemic bicyclo[3.2.0] hept-2-en-6-one | | - | 40 | - | 50.0 | 3 |
| | | 10 | 130 | 15 | | |
| | | 10 | 40 | - | | |
| 4-Phenylcyclohexanone | CP Chiralsil Dex CB Agilent, 25 m x 0.25 mm x 0.25 μm | - | 80 | - | 50.0 | 3 |
| | | 10 | 110 | - | | |
| | | 1 | 200 | 20.00 | | |
| | | 10 | 80 | - | | |
| 2-Phenylcyclohexanone | | - | 40 | - | 20.0 | 1 |
| | | 10 | 180 | 0 | | |
| | | 1 | 200 | 15 | | |
| | | 10 | 40 | 0 | | |
| Phenylacetone | | - | 30.0 | 5.00 | 5.0 | 2 |
| | | 5.00 | 70.0 | 5.00 | | |
| | | 5.00 | 130.0 | 10.00 | | |
| 2-Phenylcyclohexanone 4-Phenylcyclohexanone | | - | 30.0 | 5.00 | 10.0 | 2 |
| | | 10.00 | 160.0 | 5.00 | | |
| | | 5.00 | 250.0 | 5.00 | | |
| Benzoin | HP-1 Agilent, 30 m x 0.25 mm x 0.25 μm | - | 55.0 | 5.00 | 5.0 | 2 |
| | | 5.00 | 78.0 | 3.00 | | |
| | | 10.00 | 170.0 | 1.00 | | |
| | | 5.00 | 180.0 | 3.00 | | |
| | | 5.00 | 215.0 | 5.00 | | |
| | | 10.00 | 250.0 | 3.00 | | |
| Substrate scope analysis | | - | 30 | 5 | 5.0 | 2 |
| | | 5 | 70 | 5 | | |
| | | 5 | 130 | 5 | | |
| | | 10 | 325 | 5 | | |

**Table S2. Identification of putative Type I BVMO in *S. leeuwenhoekii* proteome.** The genome analyses are shown in the table with the predicted proteins, their respective Uniprot accession number, the predicted amino acid sequence of the conserved motif (two Rossman and the Type I BVMO fingerprints) and the names of the cluster judged by percentage of identity.

| Predicted protein | Uniprot Accesion number | Rossman motif G-x-G-x-x-[G/A] | Type I BVMO fingerprints [A/G]-G-x-W-x-x-x-[F/Y]-P-[G/M]-x-x-x-D | F-x-G-x-x-H-x-x-x-W-[P/D] | Rossman motif G-x-G-x-x-[G/A] | 90% identity Cluster name |
|---|---|---|---|---|---|---|
| Sle_13190 | A0A0F7VUW7 | GAGIGG | GGTWYWNRFPGVRCD | FAGHSFHTSRWD | GTGSTT | Pentalenolactone D synthase |
| Sle_62070 | A0A0F7W6X7 | GGGFGG | GGTWYWNRYPGIHCD | FEGHTFHTSRWD | GTGATG | Phenylacetone monooxygenase |
| Sle_41160 | A0A0F7VV32 | GSGFGG | GGTWRDNSYPGCACD | FPGKVFHSARWD | GTGASA | Uncharacterized monooxygenase |

6
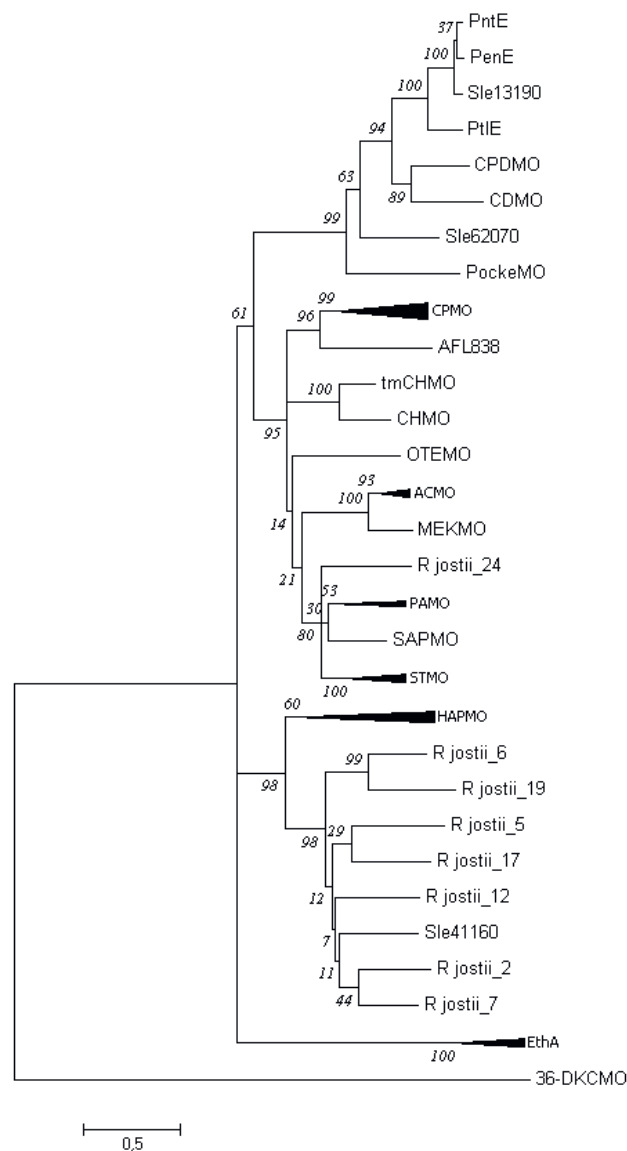
# Figures



**Figure S1. Molecular phylogenetic analysis.** The evolutionary history was inferred by using the Maximum Likelihood method; the tree with the highest log likelihood (-14431,27) is shown. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 45 amino acid sequences including CDMO (cyclododecanone monooxygenase), CPDMO (cyclopentanone monooxygenase), PockeMO (polycyclic ketone monooxygenase), CPMO (cyclopentanone monooxygenase), MEKMO (methylethylketone monooxygenase), ACMO (acetone monooxygenase), PAMO (phenylacetone monooxygenase), SAPMO (4-sulfoacetophenone monooxygenase) STMO (steroid monooxygenase), OTEMO (2-oxo-$\Delta^3$-4,5,5- trimethylcyclopentenylacetyl-CoA monooxygenase), CHMO (cyclohexanone monooxygenase), HAPMO (4-hydroxyacetophenone monooxygenase), EthA (Ethionamide monooxygenase), pentalenolactone biosynthetic genes PenE, PntE and PtlE, 24 BVMOs from *Rhodococcus jostii* RHA1 and the three putative proteins from *S. leeuwenhoekii* C34.
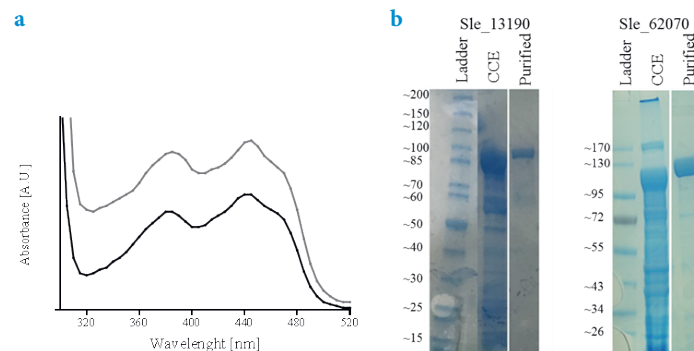
**Figure S2. Purification of Sle_13190 and Sle_62070.** (a) UV-Vis absorbance spectra of purified Type I BVMOs; Sle_13190 in black and Sle_62070 in gray. (b) SDS-PAGE of both purified flavoproteins and their respective clarified crude extracts.

**6**
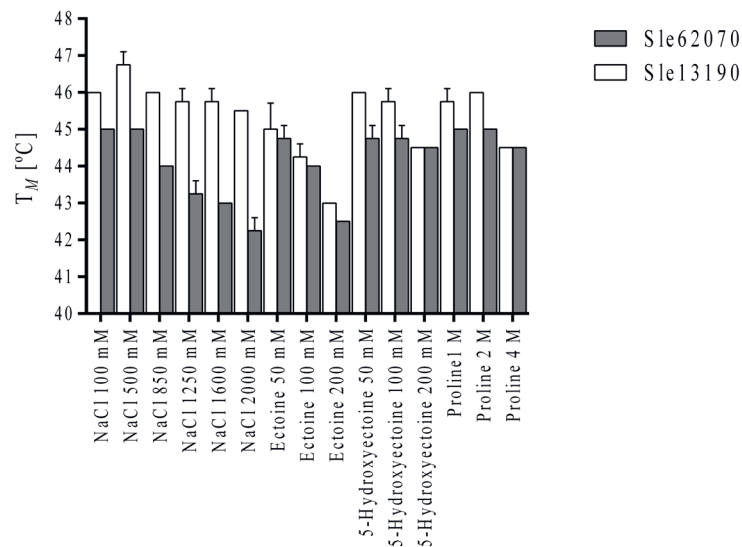


**Figure S3. Determination of apparent melting temperatures ($T_M$') of *S. leeuwenhoekii* Type I BVMOs with additives.** The $T_M$' of Sle_62070 (gray column) and Sle_13190 (white column) were determined using increasing concentrations of NaCl, ectoine, 5-hydroxyectoine and proline as additives, n=3.
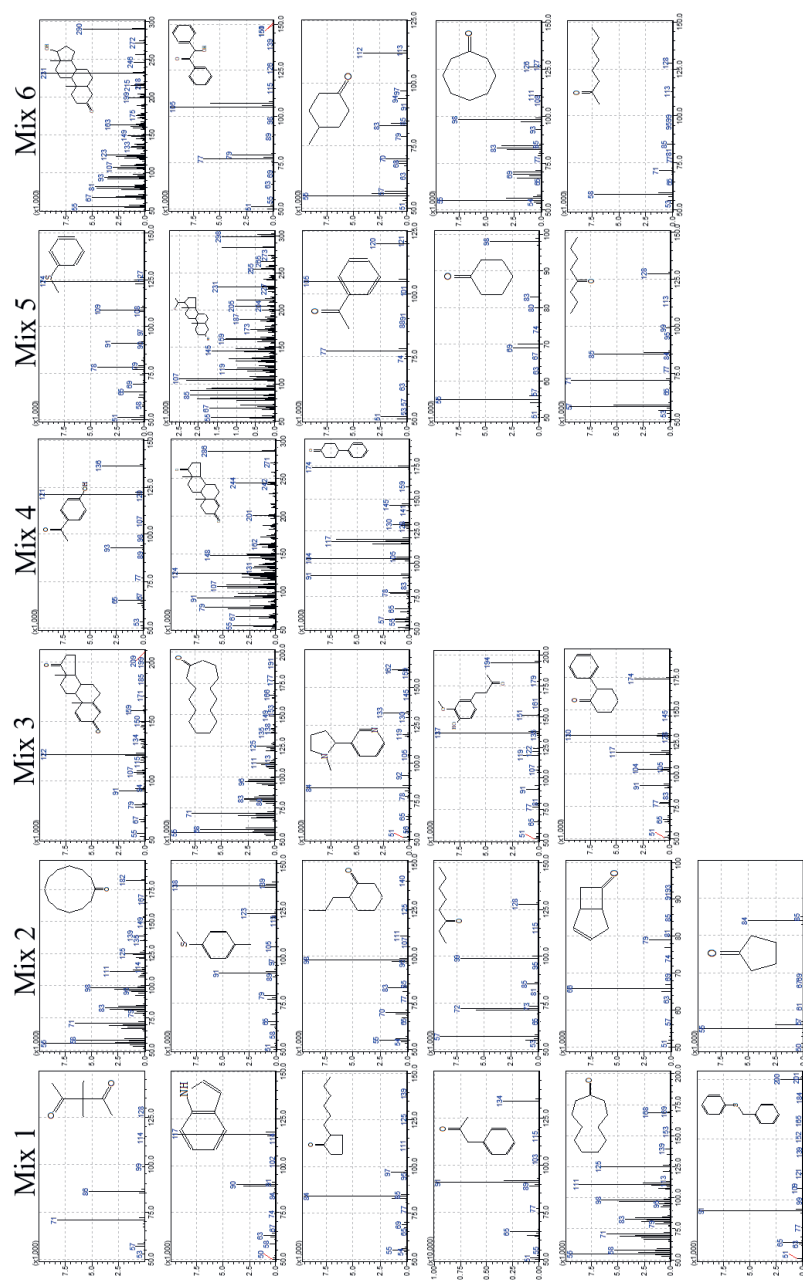
**Figure S4. List of compounds used for substrate scope analysis.** 30 Compounds were used in the analysis of substrate acceptance using 6 different compound mixtures. The composition of the mixtures are shown in columns with the structure and the respective ms spectrum obtained from GCMSsolution Postrun Analysis 4.11 (Shimadzu). Mix 1, 3-methyl-2,4-pentanedione, indole, 2-hexylcyclopentanone, phenylacetone, cycloundecanone and benzylphenyl sulfide; mix 2, cyclododecanone, methyl-p-tolyl sulfide, 2-propylcyclohexanone, 3-octanone, biclye[3.2.0]hept-2-en-6-one and cyclopentanone; mix 3, androst-1,4-diene-3,17-dione, cyclopentadecanone, nicotine, vanillylacetone and 2-phenylcyclohexanone; mix 4, 4-hydroxyacetophenone, androst-4-ene-3,17-dione, 4-phenylcyclohexanone; mix 5, thioanisole, pregnalone, acetophenone, cyclohexanone and 4-octanone; mix 6, stanolone, benzoin, 4-methylcyclohexanone, cyclooctanone and 2-octanone.
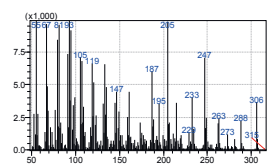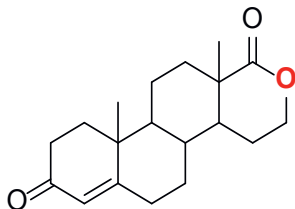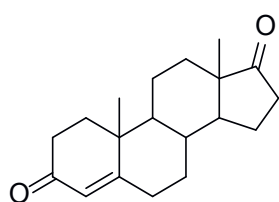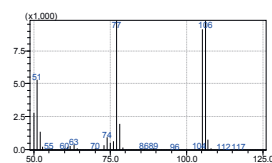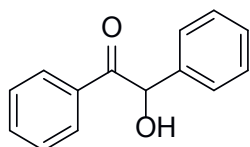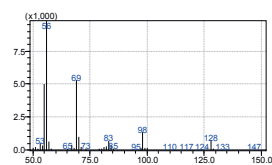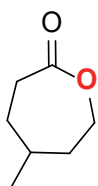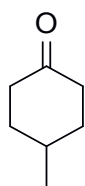
**Figure S5. Substrate and product MS identification.** *The products of 2-phenylcyclohexanone and benzoin were analyzed by ^{1}H-NMR. While the product for 4-phenylcyclohexanone was not in the MS spectrum library: the molecular structure depicts the expected BV product.

| Substrate | Product | MS spectra |
|---|---|---|
| | | |
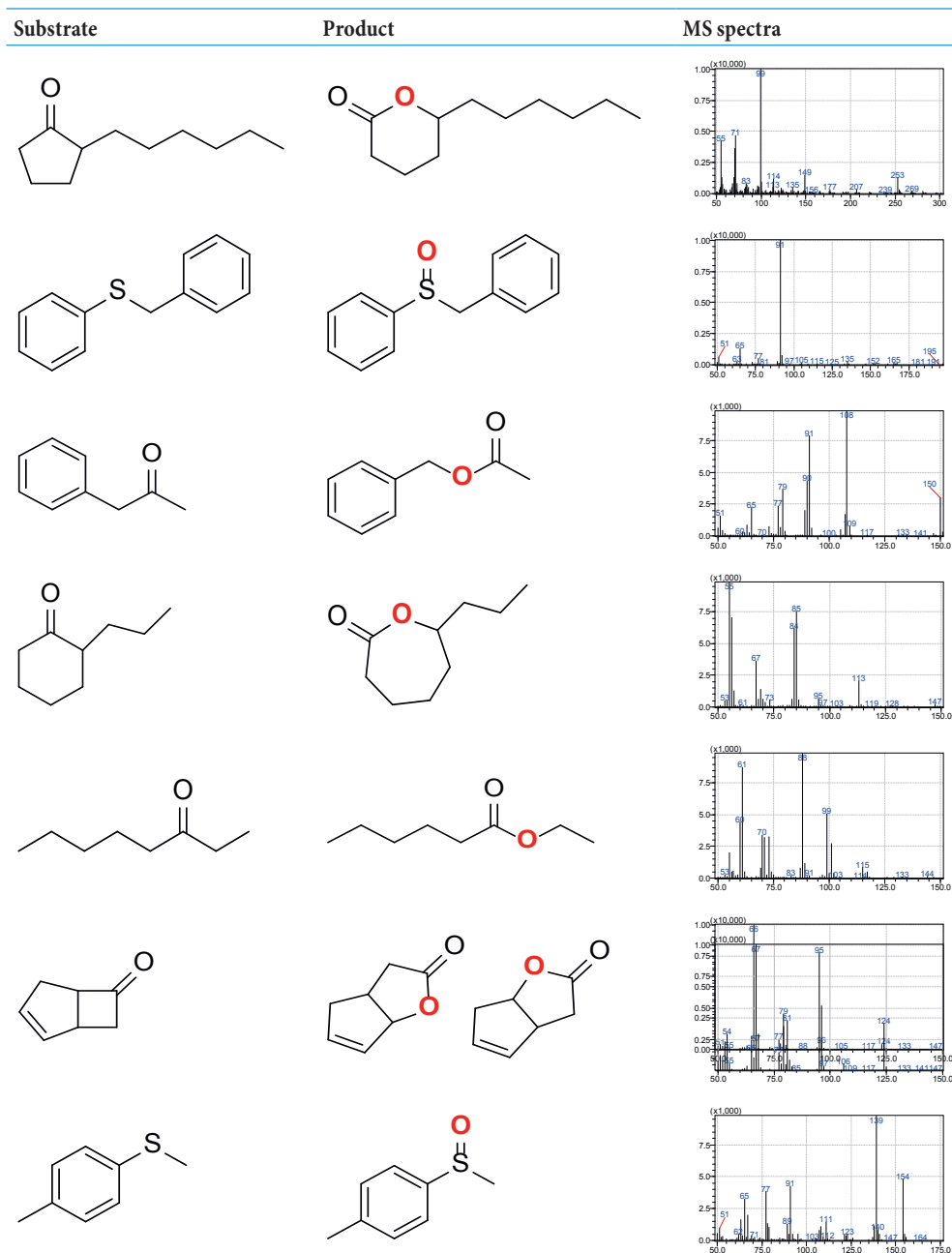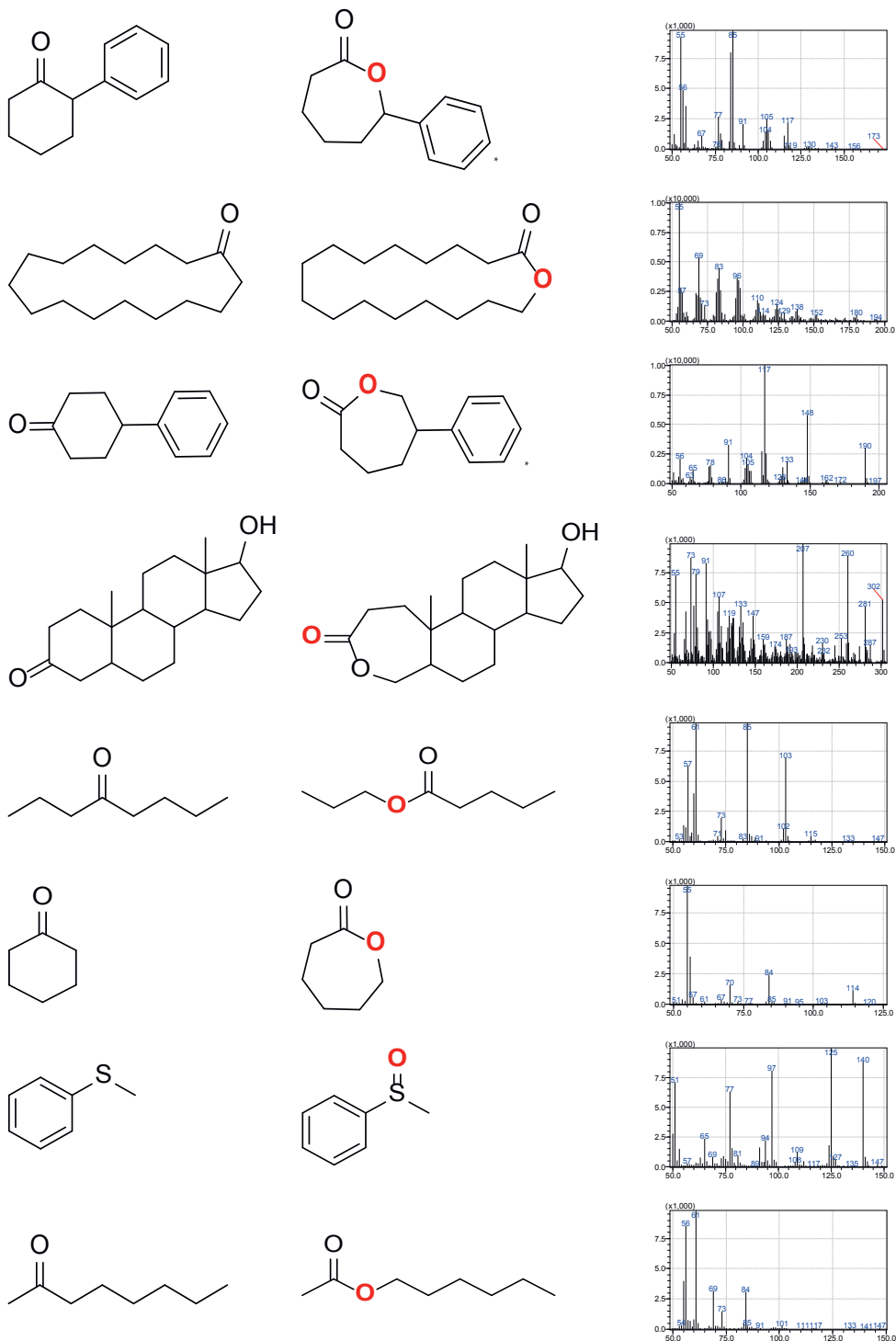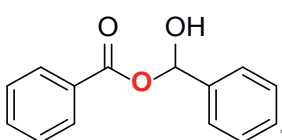
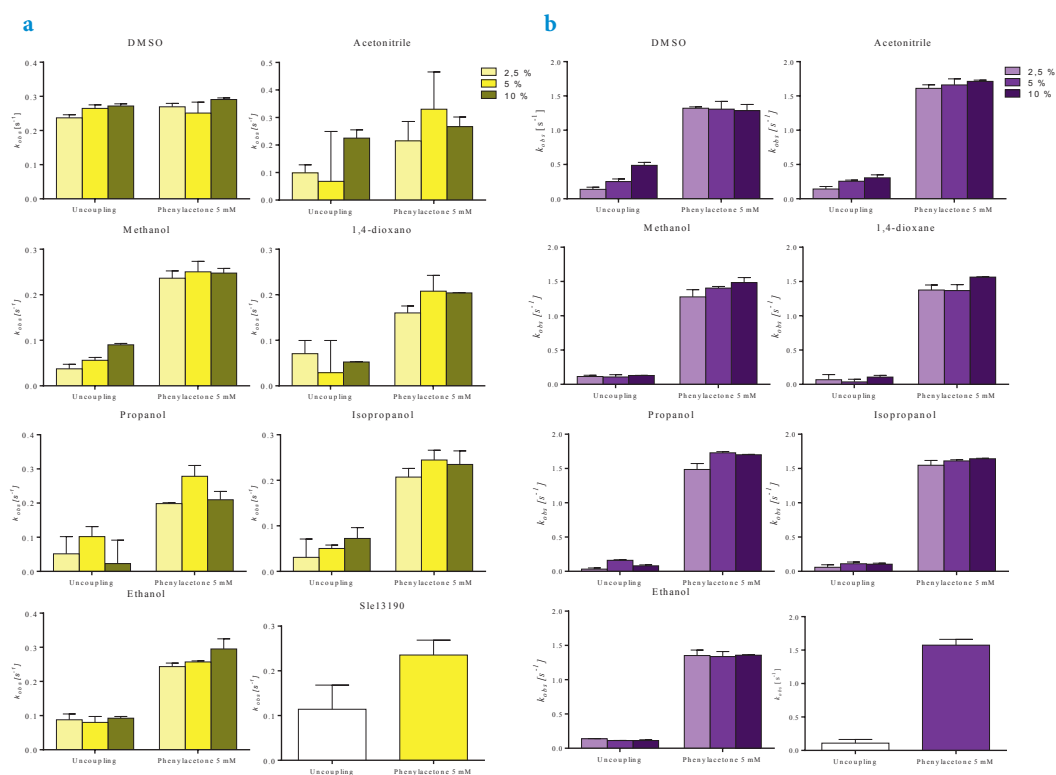**Figure S6. Effects of water-miscible cosolvents in BVMO and NADPH oxidase activity.** Activity ($k_{obs}$) was measured using 5.0 mM phenylacetone or in the absence of substrate [(a) Sle_13190 and (b) Sle_62070]. The effect of cosolvent was tested using 2.5 (light column), 5 (dark column) and 10 % v/v (black column) of the cosolvent (DMSO, acetonitrile, methanol, 1,4-dioxane, propanol, isopropanol or ethanol), n=3
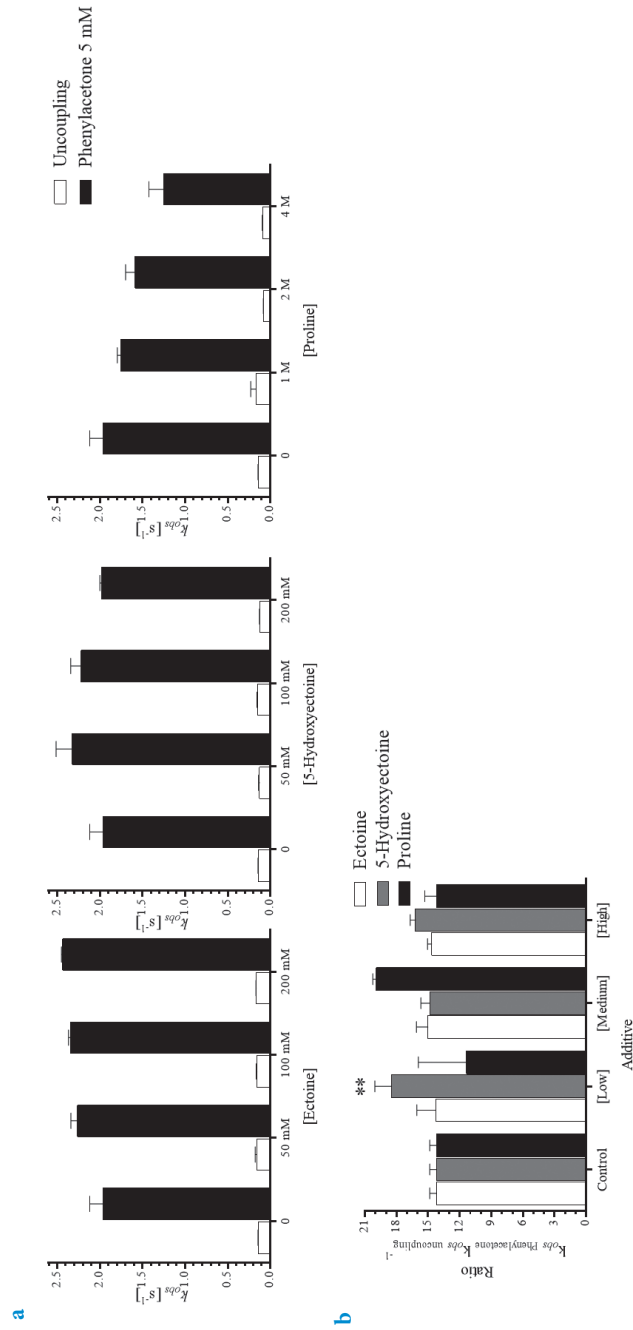
**Figure S7. Effect of additives in BVMO and NADPH oxidase activity of Sle_62070.** (a) The $k_{obs}$ of Sle_62070 in presence of 5.0 mM phenylacetone (black column) and in absence of substrate (white column) was measured at increasing concentrations of ectoine, 5-hydroxyectoine and proline. (b) The ratio $k_{obs}$ with phenylacetone and in absence of substrate was calculated for the additive conditions using ectoine (white column), 5-hydroxyectoine (gray column) and proline (black column), n=3, **=statistical significance with $p < 0.05$.

**6**
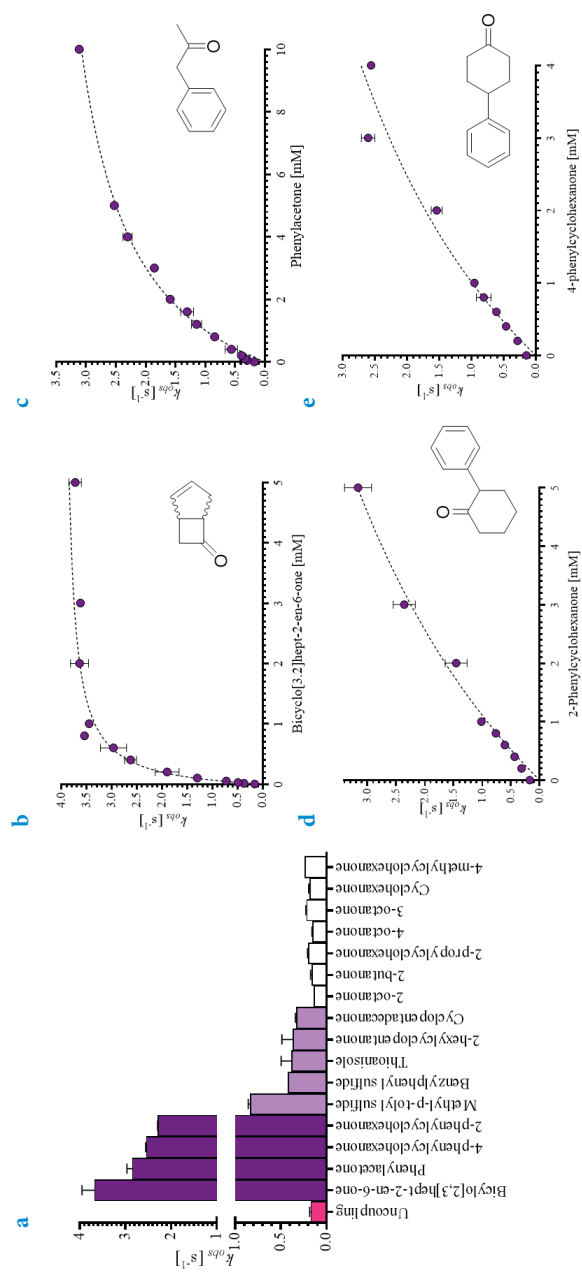
**Figure S8. NADPH oxidation activity analysis of Sle_62070.** (a) NADPH oxidation was spectrophotometrically followed at 340 nm using Sle_62070 as catalyst with different ketones and sulfides. As a control the uncoupling rate was measured (red column). The substrates that showed highest activity (dark purple) were selected for a more detailed kinetic analysis: (b) bicyclo[3.2.0]hept-2-en-6-one, (c) phenylacetone, (d) 2-phenylcyclohexanone and (e) 4-phenylcyclohexanone, n=3.
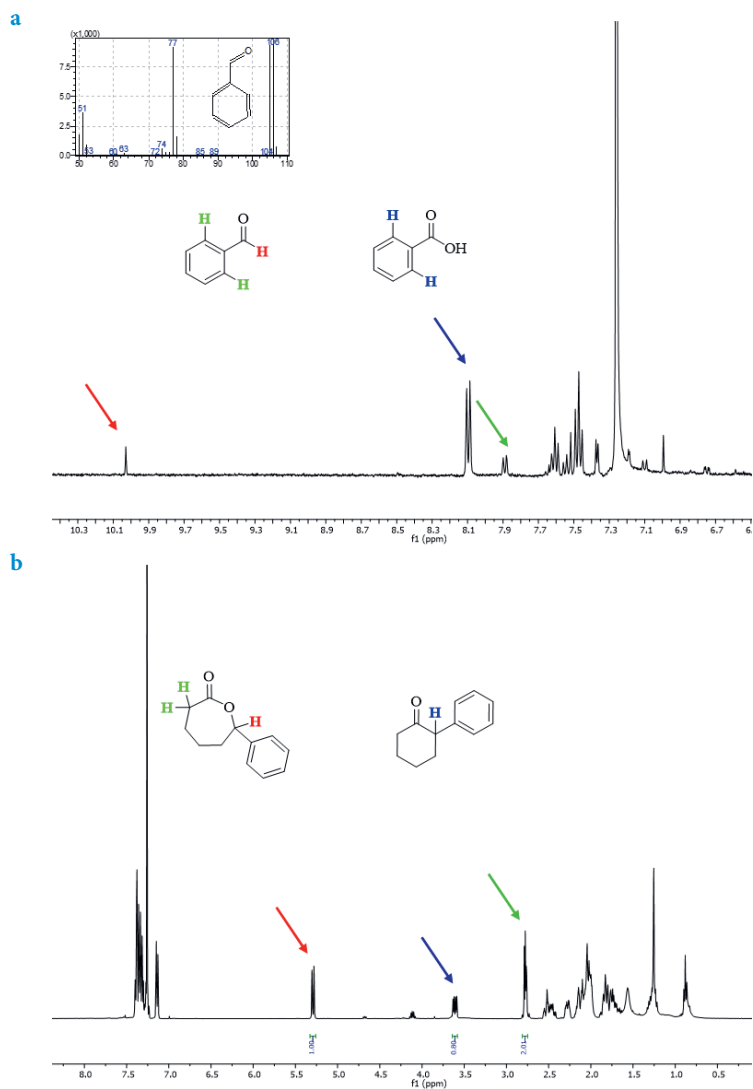
**Figure S9. Identification of reactions products using Sle_62070 as biocatalyst.** (a) ^{1}H-NMR analysis in CDCl$_3$ at 400 MHz of reaction with benzoin as substrate and MS spectrum of benzaldehyde obtained in GC-MS. (b) ^{1}H-NMR analysis of 2-phenylcyclohexanone reaction in CDCl$_3$ at 400 MHz.
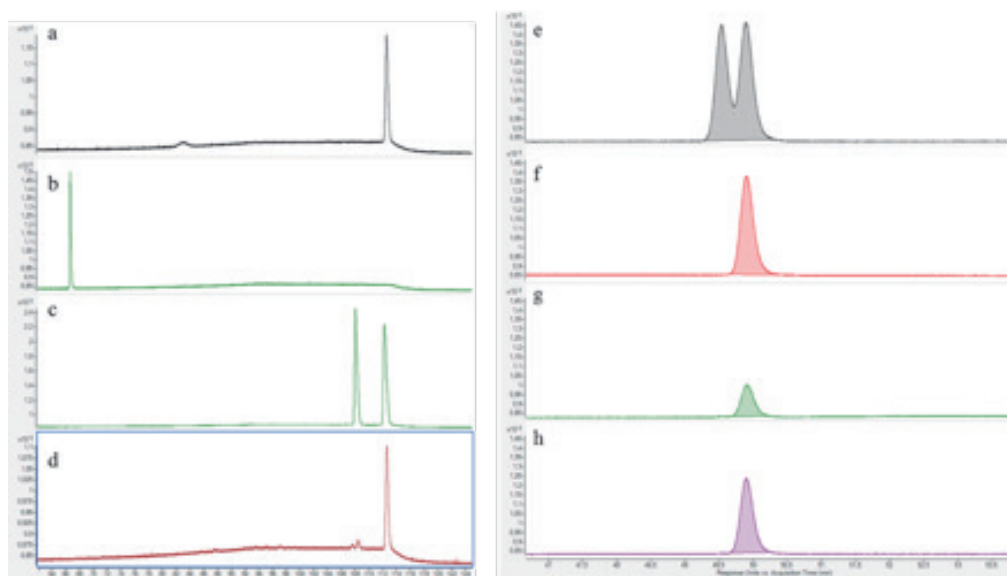
**Figure S10. Biotransformation of Sle_62070 using 2- and 4-phenylcyclohexanone.** Reactions contained purified enzyme (2.7 µM), 2-phenylcyclohexanone or 4-phenylcyclohexanone (5.0 mM), NADPH (150 µM), $Na_2PO_3 \cdot 5H_2O$ (10 mM), FAD (30 µM), NaCl (100 mM), glycerol (10 % w/v) in 50 mM Tris-HCl at pH 8.0 and 1,4-dioxane as cosolvent (2.5 % v/v). For (a) 4-phenylcyclohexanone reactions were analyzed by chiral GC after 2 h at 24 °C. (b) Substrate and (c) synthetized racemic products were also analyzed. Chromatograms were compared with (d) TmCHMO-PTDH described to produce preferably the *S* lactone. For 2-phenylcyclohexanone reactions were analyzed after (e) 0, (f) 2 and (g) 24 h at 24 °C. Chromatograms were compared with (h) TmCHMO-PTDH described to display *R* selectivity.
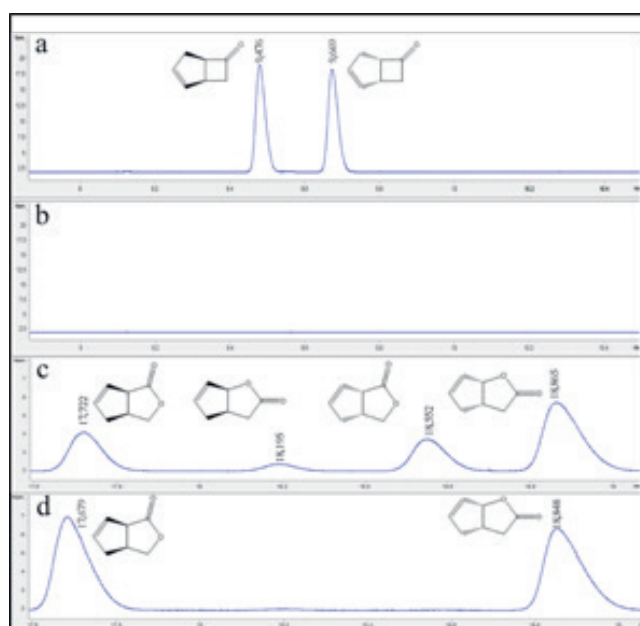
**Figure S11. Chromatograms of the biotransformation using Sle_62070**. Reactions contained purified enzyme (2.7 µM), *rac*-bicyclo[3.2.0]hept-2-en-6-one (5.0 mM), NADPH (150 µM), Na$_2$PO$_3$•5H$_2$O (10 mM), FAD (30 µM), NaCl (100 mM), glycerol (10 % w/v) in 50 mM Tris-HCl at pH 8.0 and 1,4-dioxane as cosolvent (2.5 % v/v). Reactions containing rac-bicyclo[3.2.0]hept-2-en-6-one were analyzed by chiral GC after (a) 0 and (b & d) 2h at 24 °C. Sle_62070 fully converted *rac*-bicyclo[3.2.0]hept-2-en-6-one and produced almost exclusively one regioisomer from each enantiomer, namely (1S,5R)-2-oxabicyclo[3.3.0]oct-6-en-3-one (normal product) and (1R,5S)-3-oxabicyclo[3.3.0]oct-6-en-2-one (abnormal product). Retention times were compared with (c) where PAMO-PTDH was used for converting *rac*-bicyclo[3.2.0]hept-2-en-6-one and which is known to yield all four possible lactones.

**6**
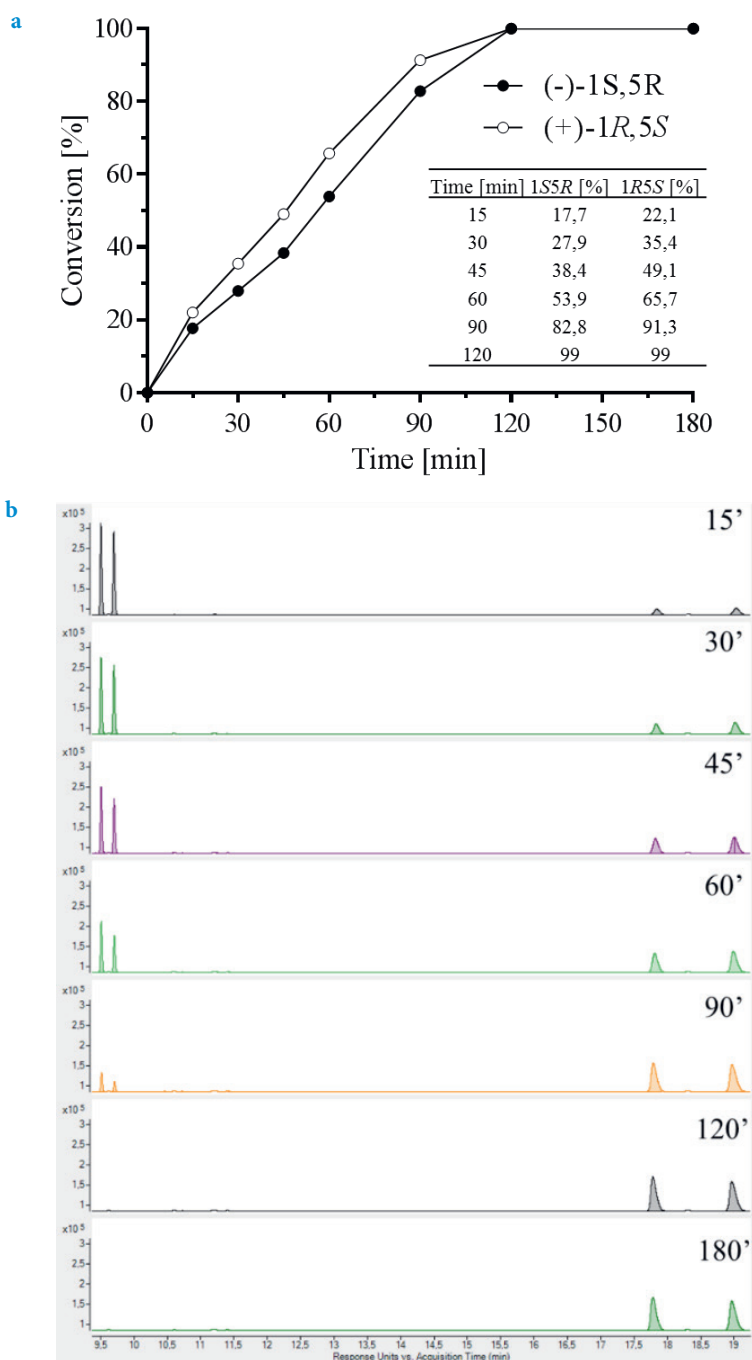
**Figure 12. Biotransformation of *rac*-bicyclo[3.2.0]hept-2-en-6-one in time.** Biotransformations using Sle_62070 and *rac*-bicyclo[3.2.0]hept-2-en-6-one were followed in time. (a) Conversion of (-)-1S,5R (black) or (+)-1R,5S (white) substrate are shown in time. (b) Chromatograms of time point reaction are shown (15, 30, 45, 60, 90, 120 and 180 min). Retention time of substrates (-)-1S,5R, (+)-1R,5S and the products (abnormal and normal) were 9.47, 9.67, 17.68 and 18.85, respectively.

# REFERENCES

1. Santiago, M., Ramírez-Sarmiento, C. A., Zamora, R. A. & Parra, L. P. Discovery, molecular mechanisms, and industrial applications of cold-active enzymes. *Front. Microbiol.* 2016, 7.

2. Dijkman, W. P., de Gonzalo, G., Mattevi, A. & Fraaije, M. W. Flavoprotein oxidases: classification and applications. *Appl. Microbiol. Biotechnol.* 2013, 97, 5177–5188.

3. Pellis, A., Cantone, S., Ebert, C. & Gardossi, L. Evolving biocatalysis to meet bioeconomy challenges and opportunities. *N. Biotechnol.* 2018, 40, 154–169.

4. Reetz, M. T. Biocatalysis in Organic Chemistry and Biotechnology: Past, present, and future. *J. Am. Chem. Soc* 2013, 135, 12480–12496.

5. Woodley, J. M. Integrating protein engineering with process design for biocatalysis. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* 2018, 376, 20170062.

6. Dewan, S. S. Global markets for enzymes in industrial applications. *BCC Res. Wellesley* 2014.

7. Leisch, H., Morley, K. & Lau, P. C. K. Baeyer– Villiger monooxygenases: more than just green chemistry. *Chem. Rev.* 2011, 111, 4165–4222.

8. Bordewick, S., Beier, A., Balke, K. & Bornscheuer, U. T. Baeyer-Villiger monooxygenases from *Yarrowia lipolytica* catalyze preferentially sulfoxidations. *Enzyme Microb. Technol.* 2018, 109, 31–42.

9. Mthethwa, K. S. *et al.* Fungal BVMOs as alternatives to cyclohexanone monooxygenase. *Enzyme Microb. Technol.* 2017, 106, 11–17.

10. Ceccoli, R. D., Bianchi, D. A. & Rial, D. V. Flavoprotein monooxygenases for oxidative biocatalysis: recombinant expression in microbial hosts and applications. *Front. Microbiol.* 2014 5, 25.

11. Fürst, M.J., Savino, S., Dudek, H.M., Gómez Castellanos, J.R., Gutiérrez de Souza, C., Rovida, S., Fraaije, M.W. & Mattevi, A. Polycyclic ketone monooxygenase from the thermophilic fungus *Thermothelomyces thermophila*: a structurally distinct biocatalyst for bulky substrates. *J. Am. Chem. Soc* 2016, 139, 627–630.

12. Torres Pazmiño, D. E., Riebel, A., de Lange, J., Rudroff, F., Mihovilovic, M. D., & Fraaije, M. W. Efficient biooxidations catalyzed by a new generation of self-sufficient Baeyer–Villiger monooxygenases. *ChemBioChem* 2009, 10, 2595–2598.

13. Johannes, T. W., Woodyer, R. D. & Zhao, H. Directed evolution of a thermostable phosphite dehydrogenase for NAD (P) H regeneration. *Appl. Environ. Microbiol.* 2005, 71, 5728–5734.

14. Karan, R., Capes, M. D. & DasSarma, S. Function and biotechnology of extremophilic enzymes in low water activity. *Aquat. Biosyst.* 2012 8, 4.

15. Fraaije, M. W., Wu, J., Heuts, D. P., Van Hellemond, E. W., Spelberg, J. H. L., & Janssen, D. B. Discovery of a thermostable Baeyer-Villiger monooxygenase by genome mining. *Appl. Microbiol. Biotechnol.* 2005, 66, 393–400.

16. de Gonzalo, G., Ottolina, G., Zambianchi, F., Fraaije, M. W. & Carrea, G. Biocatalytic properties of Baeyer–Villiger monooxygenases in aqueous–organic media. *J. Mol. Catal. B Enzym.* 2006, 39, 91–97.

17. Rodríguez, C., de Gonzalo, G., Torres Pazmiño, D. E., Fraaije, M. W. & Gotora, V. Selective Baeyer–Villiger oxidation of racemic ketones in aqueous–organic media catalyzed by phenylacetone monooxygenase. *Tetrahedron: Asymmetry* 2008, 19, 197–203.

18. Romero, E., Castellanos, J. R. G., Mattevi, A. & Fraaije, M. W. Characterization and crystal structure of a robust cyclohexanone monooxygenase. *Angew. Chem. Int. Ed. Engl.* 2016, 55, 15852–15855.

19. Pulschen, A.A., Rodrigues, F., Duarte, R.T., Araujo, G.G., Santiago, I.F., Paulino-Lima, I.G., Rosa, C.A., Kato, M.J., Pellizari, V.H. & Galante, D. UV-resistant yeasts isolated from a high-altitude volcanic area on the Atacama Desert as eukaryotic models for astrobiology. *Microbiologyopen* 2015, 4, 574–588.

20. Azua-Bustos, A., Urrejola, C. & Vicuña, R. Life at the dry edge: microorganisms of the Atacama Desert. *FEBS Lett.* 2012, 586, 2939–2945.

21. Paulino-Lima, I.G., Azua-Bustos, A., Vicuña, R., González-Silva, C., Salas, L., Teixeira, L., Rosado, A., da Costa Leitao, A.A. & Lage, C Isolation of UVC-tolerant bacteria from thehyperarid Atacama Desert, Chile. *Microb. Ecol.* 2013, 65, 325–335.

22. Idris, H., Goodfellow, M., Sanderson, R., Asenjo, J. A. & Bull, A. T. Actinobacterial rare biospheres and dark matter revealed in habitats of the Chilean Atacama Desert. *Sci. Rep.* 2017, 7, 1–11.

23. Bull, A. T., Asenjo, J. A., Goodfellow, M. & Gómez-Silva, B. The Atacama desert: technical resources and the growing importance of novel microbial diversity. *Annu. Rev. Microbiol.* 2016, 70, 215–234.

24. Bull, A. T. & Asenjo, J. A. Microbiology of hyper-arid environments: recent insights from the Atacama Desert, Chile. *Antonie Van Leeuwenhoek* 2013, 103, 1173–1179.

25. Okoro, C.K., Brown, R., Jones, A.L., Andrews, B.A., Asenjo, J.A., Goodfellow, M. & Bull, A.T. Diversity of culturable actinomycetes in hyper-arid soils of the Atacama Desert, Chile. *Antonie Van Leeuwenhoek* 2009, 95, 121–133.

26. Busarakam, K., Bull, A.T., Girard, G., Labeda, D.P., van Wezel, G.P. & Goodfellow, M. *Streptomyces leeuwenhoekii* sp. nov., the producer of chaxalactins and chaxamycins, forms a distinct branch in *Streptomyces* gene trees. *Antonie Van Leeuwenhoek* 2014, 105, 849–861.

27. Gomez-Escribano, J.P., Castro, J.F., Razmilic, V., Chandra, G., Andrews, B., Asenjo, J.A. and Bibb, M.J. The *Streptomyces leeuwenhoekii* genome: de novo sequencing and assembly in single contigs of the chromosome, circular plasmid pSLE1 and linear plasmid pSLE2. *BMC Genomics* 2015, 16, 485.

28. Rateb, M.E., Houssen, W.E., Harrison, W.T., Deng, H., Okoro, C.K., Asenjo, J.A., Andrews, B.A., Bull, A.T., Goodfellow, M., Ebel, R. & Jaspars, M. Diverse metabolic profiles of a *Streptomyces* strain isolated from a hyper-arid environment. *J. Nat. Prod.* 2011, 74, 1965–1971.

29. Gibson, M., Nurealam, M., Lipata, F., Oliveira, M. A. & Rohr, J.Characterization of kinetics and products of the baeyer–villiger oxygenase MtmOIV, the key enzyme of the biosynthetic pathway toward the natural product anticancer drug mithramycin from *Streptomyces argillaceus*. *J. Am. Chem. Soc.* 2005, 127, 17594–17595.

30. Seo, M. J., Zhu, D., Endo, S., Ikeda, H. & Cane, D. E. Genome mining in *streptomyces*. elucidation of the role of Baeyer-Villiger monooxygenases and non-heme iron-dependent dehydrogenase/oxygenases in the final steps of the biosynthesis of pentalenolactone and neopentalenolactone. *Biochemistry* 2011, 50, 1739–1754.

31. Park, J., Kim, D., Kim, S., Kim, J., Bae, K., & Lee, C. The analysis and application of a recombinant monooxygenase library as a biocatalyst for the Baeyer-Villiger reaction. *J. Microbiol. Biotechnol.* 2007, 17, 1083–1089.

32. Jiang, J., Tetzlaff, C. N., Takamatsu, S., Iwatsuki, M., Komatsu, M., Ikeda, H., & Cane, D. E. Genome mining in *Streptomyces avermitilis*: a biochemical baeyer-villiger reaction and discovery of a new branch of the pentalenolactone family tree. *Biochemistry* 2009, 48, 6431–6440.

33. Fraaije, M. W., Kamerbeek, N. M., van Berkel, W. J. H. & Janssen, D. B. Identification of a Baeyer–Villiger monooxygenase sequence motif. *FEBS Lett.* 2002, 518, 43–47.

34. Riebel, A., Dudek, H. M., De Gonzalo, G., Stepniak, P., Rychlewski, L., & Fraaije, M. W. Expanding the set of rhodococcal Baeyer-Villiger monooxygenases by high-throughput cloning, expression and substrate screening. *Appl. Microbiol. Biotechnol.* 2012, 95, 1479–1489.

35. Whelan, S. & Goldman, N. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol.* 2001, 18, 691–699.

36. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 2016, 33, 1870–1874.

37. Hobbs, G. Dispersed growth of *Streptomyces* in liquid culture. *Appl Microbiol Biotechnol* 1989, 31, 272–277.

38. Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., & Smith, H. O. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 2009, 6, 343–345.

39. Forneris, F., Orru, R., Bonivento, D., Chiarelli, L. R. & Mattevi, A. ThermoFAD, a Thermofluor®-adapted flavin ad hoc detection system for protein folding and ligand binding. *FEBS J.* 2009, 276, 2833–2840.

40. Vogel, M. & Schwarz-Linek, U. Bioorganic Chemistry: Highlights and New Aspects, *ed. U. Diederichsen*. 1999.

41. Bocola, M., Schulz, F., Leca, F., Vogel, A., Fraaije, M. W., & Reetz, M. T. Converting phenylacetone monooxygenase into phenylcyclohexanone monooxygenase by rational design: towards practical Baeyer–Villiger monooxygenases. *Adv. Synth. Catal.* 2005, 347, 979–986.

42. Alphand, V., Furstoss, R., Pedragosa-Moreau, S., Roberts, S. M. & Willetts, A. J. Comparison of microbiologically and enzymatically mediated Baeyer–Villiger oxidations: synthesis of optically active caprolactones. *J. Chem. Soc. Perkin Trans.* 1996, 1, 1867–1872.

43. Rial, D. V *et al.* Stereoselective desymmetrizations by recombinant whole cells expressing the Baeyer–Villiger monooxygenase from *Xanthobacter* sp. ZL5: a new biocatalyst accepting structurally demanding substrates. *European J. Org. Chem.* 2008, 1203–1213.

**6**

# VII

# Summary and conclusions

Alejandro Gran-Scheuch

## SUMMARY AND CONCLUSIONS

Nowadays, biocatalysis has a crucial role in society. Enzymes have been developed for a plethora of biotechnological applications. New biocatalytic advances aim to reduce the environmental impact of chemical processes. In this context, many years have passed since the field of biocatalysis abandoned the merely curiosity-driven research in academia and became a more mature field for e.g. the chemical and pharmaceutical industries. This occurred because biocatalysts allow a tight control over the selectivity of reactions, making enzymes exquisite alternatives for chemical catalysts. They can also contribute to a greener industry and untapped new chemistry.

Flavoenzymes are a remarkable example of biotechnologically interesting biocatalysts. These enzymes are highly versatile biomolecular machines that perform a wide range of reactions, and represent a vast number of biocatalysts in the toolbox that enables redox chemistry. Flavoenzymology is a multidisciplinary field, which aims to increase the overall understanding of flavin-containing enzyme, by gaining a deeper insight into their catalytic mechanisms, kinetic properties and sequence-structure-function relationships. An improved understanding of these enzymes permits the further development of enhanced biocatalysts. For example, knowledge-based enzyme engineering can lead to improved chemo-, regio- and enantioselectivities required for the synthesis of valuable compounds.

The advantages of biocatalysis are well-exemplified for the Baeyer-Villiger oxidation reactions reported in **Chapter 1**. Even though the chemical method for such oxidations offers a simple synthetic route to oxidize ketones into esters/lactones, it requires the usage of hazardous reagents and is typically not chemo-, regio- or enantioselective. As an appealing alternative, the flavin-containing Baeyer-Villiger monooxygenases (BVMOs) have been investigated for a long time for their potential as biocatalysts[1–5]. In the last few decades various BVMOs have been discovered or engineered that exhibit activities suitable for the synthesis of valuable compounds. In **chapter 1**, we summarized the state of the art for BVMOs as biocatalysts, thereby focusing on their biochemical, mechanistic and structural properties. Although it appears to be a relatively mature field, there are still intricacies for flavoenzymologists to unravel. Limitations, such as moderated stability, undesired specificities and cofactor dependency often prohibit industrial applications. Ongoing research is making progress in overcoming these issues.

Many (flavo)enzymes depend on a dissociable cofactor, such as NAD(P)H, for catalysis. The costs related to such cofactor dependency can be partly overcome by using self-sufficient bifunctional enzymes that are efficient in cofactor regeneration. In **chapter 2** we designed an experimental protocol to evaluate the effect of the length of a flexible

glycine-rich linker on the properties of the self-sufficient TbADH-TmCHMO fusions. This experimental protocol enables the generation of optimized biocatalysts. Moreover, the procedure can be easily modified to generate a similar library of other fusion proteins. As demonstration reaction, the generated ADH-BVMO fusion enzymes were tested for the synthesis of ε-caprolactone, a valuable polymer precursor, starting from cyclohexanol. Such reaction takes advantage of the fact that both enzyme activities satisfy their cofactor dependency: ADH activity converts $NADP^+$ into NADPH, while, in turn, BVMO activity oxidizes NADPH into $NADP^+$. Furthermore, fused bifunctional biocatalysts may enhance the protein stability and promote product intermediate channeling by proximity effects. There are no uniform design rules for an optimal linker to fuse two enzymes. However, a 'wrong' linker can lead to detrimental effects on the biocatalytic performance[6]. The presented work concerned the evaluation of the effect of fifteen linker variants, differing in size, on: expression, thermostability, activity and conversion levels. All the obtained variants exhibited high expression levels (250-360 mg $L^{-1}$) and were obtained mainly as holo proteins. For both TmCHMO and TbADH, it was observed that the length of the linker did not show a significant deleterious effect on their thermostability. Concerning activity, the alcohol oxidation (ADH) and sulfoxidation (BVMO) activities were similar for the 9 shortest variants, while the fusions with the linker length of 10, 12 and 15 amino acids showed a slightly increased activity. Small-scale bioconversions resulted in highest turnover numbers (TON) for the fusions with 2, 3, 6, 7, 13 and 14 amino acid linkers (TONs of 20,000-25,000). The TON of the reaction catalyzed by the non-fused enzymes was around 12,000.

In **chapter 3**, we studied the reactivity of flavoenzymes with dioxygen, a captivating topic for flavoenzymologists. Dioxygen is a four-electron oxidant that can be enzymatically activated and reduced to hydrogen peroxide or water through consecutive one-electron transfers. In some cases, other reactive oxygen species (ROS) can be formed, such as superoxide. In **chapter 3**, we investigated ROS formation —or uncoupling— during the flavin-mediated reduction of dioxygen by flavoprotein oxidases and monooxygenases, using $PAMO_{WT}$, $PAMO_{C65D}$, EUGO and HMFO as test enzymes. The analysis of the uncoupling in flavoproteins is of great relevance because ROS have an important role in biology and may complicate the use of flavoenzymes as biocatalysts. Moreover, the mechanism of flavoenzyme-mediated ROS formation is not fully understood.

Accurate profiles of hydrogen peroxide and superoxide production under different operational conditions were determined for all studied flavoenzymes. Remarkably, all proteins were found to produce significant amounts of superoxide. Furthermore, increased superoxide levels were detected at higher pH, which could be indicative of a pH-sensitive caged radical pair dissociation. Despite the accumulation of superoxide, no detrimental effect on biocatalysis was demonstrated. Interestingly, for $PAMO_{WT}$ and EUGO, the

addition of catalase significantly increased the catalytic performance. The results provide a better view on the conditions that promote ROS formation in flavoenzymes which is industrially relevant and could help to reduce formation of hazardous ROS and avoid the waste of valuable reducing equivalents.

The second part of this thesis deals with the identification and characterization of several newly discovered flavin-containing monooxygenases. Such studies on new flavoenzymes provide more insights into the chemistry feasible with natural enzymes and may also lead to new enzyme-based applications. In **chapter 4**, in collaboration with the Institute of Microbiology of the ETH (Zürich, Switzerland), the involvement of BVMOs in the production of some specific polyketides was studied. It was shown that oxygen insertion via a Baeyer-Villiger oxidation into a nascent polyketide backbone is achieved by bacterial enzymes. The biochemical properties of two of such BVMOs, Oock and LmbC-Ox, were established. These FAD-containing enzymes were shown to be involved in an oxygen-insertion reaction that is part of the biosynthesis of the secondary metabolites oocydin and lobatamide.

**Chapter 5** describes efforts in finding promising type I or type II flavin-containing monooxygenase (FMOs). Through genome mining we identified two proteins from *Chloroflexi* bacterium and the tardigrade *Hypsibius dujardini:* CbFMO (type II) and HdFMO (type I), respectively. Both enzymes displayed distinctive biochemical features. HdFMO was found to be only active with sulfides, did only accept NADPH as hydride donor, and presented a moderated thermostability ($T_M^{app}$ of 45°C). In contrast, CbFMO converted preferentially ketones into the respective esters or lactones, and displayed a poor thermostability ($T_M^{app}$ of 34°C). The motivation for studying CbFMO, a type II FMO, was partly because of the relaxed nicotinamide cofactor specificity of previously reported type II FMOs[7]. Yet, CbFMO was found to have a strong preference for NADPH. Therefore, comparing with the already described collection of flavoprotein monooxygenases, both FMOs do not seem very appealing for biocatalysis.

Finally, in **chapter 6** two type I BVMOs from *Streptomyces leeuwenhoekii* C34 were identified, namely: Sle_13190 and Sle_62070. Both enzymes were successfully expressed with phosphite dehydrogenase as fusion partner. Similar to other type I BVMOs, both proteins showed NADPH-dependent Baeyer-Villiger oxidation. The sequences of Sle_13190 and Sle_62070 clustered, based on sequence homology, with other BVMOs that are known to act on bulky compounds. In this context, it was not surprising that they accepted rather complex compounds as substrate, including biphenyls and a steroid. Moreover, both enzymes were found to be moderately robust, exhibiting a $T_M^{app}$ of 45 °C and tolerating water-miscible cosolvents. In particular, Sle_62070 was found to be highly active with cyclic ketones and displayed a high regioselectivity producing only one

lactone from 2-phenylcyclohexanone, and high enantioselectivity producing only normal (-)-1*S*,5*R* and abnormal (-)-1*R*,5*S* lactones (*e.e.* >99 %) from bicyclo[3.2.0]hept-2-en-6-one. These two newly discovered BVMOs may develop as valuable additions to the known collection of BVMOs.

The work described in this thesis delivered several new enzymes, that can be employed as biocatalysts, and new insights into their catalytic properties. The work confirms that extreme environments can be a great source of untapped robust enzymes, as demonstrated with the two BVMOs identified in a bacterium isolated from Atacama Desert (**chapter 6**). Also, an flavoenzyme from a tardigrade, a microscopic animal that can be cryopreserved and had not been considered before as source for biocatalysts, was obtained and studied. Unfortunately, this did not yield a very promising biocatalyst. It shows that organisms that can survive extreme conditions do not always (only) harbor robust biocatalysts. Besides of exploring sequence genomes of (thermophilic) microorganisms, employing a metagenomic approach may also be useful if one is looking for a robust biocatalyst. Current metagenomic approaches allow the handling of rich sources of biosynthetic dark matter, which can lead to new (bio)chemistry[8,9]. In this sense, the combined effort in the discovery of novel enzymes, more mechanistic insights and the engineering of enzymes will make biocatalytic reactions even more efficient, reliable and environmentally friendly. This will allow biocatalysis to gain ground to compete in a sustainable way with classical chemical routes, or allow totally new enzyme-based applications.

**7**

# REFERENCES

1. Mihovilovic, M. D. Enzyme mediated Baeyer-Villiger oxidations. *Curr. Org. Chem.* 10, 1265–1287 (2006).

2. Kirschner, A. & Bornscheuer, U. T. Directed evolution of a Baeyer-Villiger monooxygenase to enhance enantioselectivity. *Appl. Microbiol. Biotechnol.* 81, 465–472 (2008).

3. Leisch, H., Morley, K. & Lau, P. C. K. Baeyer-Villiger monooxygenases: More than just green chemistry. *Chem. Rev.* 111, 4165–4222 (2011).

4. Alphand, V., Carrea, G., Wohlgemuth, R., Furstoss, R. & Woodley, J. M. Towards large-scale synthetic applications of Baeyer-Villiger monooxygenases. *TRENDS Biotechnol.* 21, 318–323 (2003).

5. Romero, E., Gómez Castellanos, J. R., Mattevi, A. & Fraaije, M. W. Characterization and crystal structure of a robust cyclohexanone monooxygenase. *Angew.Chem.Int.* 55, 15852–15855 (2016).

6. Aalbers, F. S. & Fraaije, M. W. Enzyme fusions in biocatalysis: coupling reactions by pairing enzymes. *ChemBioChem* 20, 20–28 (2019).

7. Riebel, A., Fink, M. J., Mihovilovic, M. D. & Fraaije, M. W. Type II flavin-containing monooxygenases: A new class of biocatalysts that harbors Baeyer–Villiger monooxygenases with a relaxed coenzyme specificity. *ChemCatChem* 6, 1112–1117 (2014).

8. Zhang, J. J. & Moore, B. S. Natural products: Digging for biosynthetic dark matter. *Elife* 4, e06453 (2015).

9. Idris, H., Goodfellow, M., Sanderson, R., Asenjo, J. A. & Bull, A. T. Actinobacterial rare biospheres and dark matter revealed in habitats of the chilean Atacama Desert. *Sci. Rep.* 7, (2017).

**7**

# Appendices

**Nederlandse samenvatting**

**Resumen en español**

**List of publications**

**Acknowledgements**

# NEDERLANDSE SAMENVATTING

**Vertaald door Yannick Kok en Titia R. Oppewal**

Biokatalyse vervult tegenwoordig een cruciale rol in de samenleving. Zo zijn er enzymen ontwikkeld voor verscheidene biotechnologische applicaties. Door innovatie in biokatalyse wordt ook getracht de milieulasten van chemische processen te verlagen. Het is daarom niet verrassend dat biokatalytisch onderzoek al lange tijd niet meer alleen gemotiveerd wordt vanuit pure academische nieuwsgierigheid, maar ook vanuit interesse voor chemische- en pharmaceutische industriële toepassingen. Biokatalysatoren (enzymen) zijn in staat om reacties te laten verlopen met ongekende selectiviteit, waardoor ze een excellent alternatief vormen voor chemische katalysatoren. Ze kunnen bovendien bijdragen aan een groenere industrie en nieuwe typen reacties aanboren die eerder niet mogelijke waren.

Flavine-bevattende enzymen zijn een goed voorbeeld van biotechnologisch relevante biokatalysatoren. Het zijn veelzijdige biomoleculaire machines, in staat om een grote verschijdenheid aan reacties uit te voeren. Ze zijn in grote aantallen vertegenwoordigd in de klasse van enzymen die redox chemie (oxidaties en reducties) mogelijk maken. Flavoenzymologie is een multidisciplinair onderzoeksveld dat erop gericht is om flavine-bevattende enzymen beter te begrijpen. Onder andere door beter begrip van hun katalytische mechanismen, kinetische eigenschappen, en de relatie tussen aminozuurketen, structuur en functie, wordt getracht deze enzymen te ontcijferen. Door al deze kennis te combineren kunnen enzymen worden ontworpen met betere chemische- en regio- en enantioselectiviteit voor de synthese van waardevolle chemische verbindingen.

De voordelen van biokatalyse zijn evident voor de Baeyer-Villiger oxidatiereacties die besproken worden in **Hoofdstuk 1** van dit proefschrift. Voor de oxidatie van ketonen naar esters/ lactonen, bijvoorbeeld, bestaat reeds een simpele chemische synthetische route. Echter wordt hier gebruik gemaakt van gevaarlijke stoffen en zijn de reacties bovendien vaak niet chemo-, regio- of enantioselectief. Als potentieel alternatief zijn de flavine-bevattende Baeyer-Villiger monooxygenasen (BVMO'en) uitgebreid bestudeerd. In de afgelopen decennia zijn er dan ook verscheidene BVMO'en ontdekt en ontworpen die geschikt zouden zijn om als biokatalysatoren waardevolle stoffen te produceren. In **Hoofdstuk 1** wordt de *state-of-the-art* beschreven voor BVMO'en als biokatalysatoren en gaan we met name in op hun biochemische, mechanistische en structurele eigenschappen. Hoewel het op het eerste gezicht een volwassen veld lijkt, zijn er nog altijd relevante aspecten waar enzymologen een waardevolle bijdrage kunnen leveren. Middelmatige stabiliteit, ongewenste specifiteit, en de afhankelijkheid van een cofactor beperken vooralsnog de industriële applicatie van deze enzymen. Lopend onderzoek probeert hier verandering in te brengen.

Vele (flavo)enzymen zijn afhankelijk van een dissocieerbare cofactor zoals NAD(P)H voor hun katalytische werking. De kosten van dergelijke afhankelijkheid kunnen deels worden overkomen door het gebruik van zelfvoorzienende bifunctionele enzymen die de co-factor efficiënt kunnen regenereren. In **Hoofdstuk 2** hebben we de bijdrage van een flexibele glycine-linker voor de effectiviteit van zelfvoorzienende TbADH-TmCHMO fusies onderzocht. We hebben hiertoe een protocol ontworpen om dit zo goed mogelijk uit te kunnen zoeken en zo geoptimaliseerde biokatalysatoren te kunnen ontwikkelen. De procedure kan bovendien gemakkelijk aangepast worden om vergelijkbare bibliotheken van andere eiwitfusies te verkrijgen. Ter demonstratie werden de ADH-BVMO fusies gebruikt in de synthese van ε-caprolactone, een waardevol startmateriaal voor de productie van polymeren. Deze omzetting van cyclohexanol tot ε-caprolactone maakt gebruik van het feit dat beide enzymen worden voorzien van hun cofactor door elkaar. ADH zet $NADP^+$ om in NADPH, dat BVMO op zijn beurt weer oxideert tot $NADP^+$. Het fuseren van biokatalysatoren komt bovendien de eiwitstabiliteit ten goede en bevordert het doorgeven van substraten door de fysieke nabijheid. Er bestaan geen uniforme regels voor het ontwerp van de flexibele linker tussen de te fuseren eiwitten. Echter kan een 'verkeerd' ontworpen linker wel degelijk schadelijk zijn voor de biocatalytische prestaties van de enzymen. In dit proefschrift zijn de effecten van vijftien verschillende linkers onderzocht op zowel de katalytische prestaties (activiteit en conversie), als ook de algemene thermostabiliteit en expressie van de enzymen. Alle verkregen varianten vertoonden hoge expressieniveaus (250-360 mg $L^{-1}$) en werden voornamelijk verkregen als holo-eiwitten. Voor zowel TmCHMO als TbADH had de lengte van de linker geen significant nadelig effect op hun thermostabiliteit. Wat betreft activiteit waren de alcoholoxidatie- (ADH) en sulfoxidatie- (BVMO) activiteiten vergelijkbaar voor de 9 kortste varianten, terwijl de fusies met de linkerlengte van 10, 12 en 15 aminozuren een licht verhoogde activiteit vertoonden. Bioconversies op kleine schaal resulteerden in de hoogste omzetcijfers voor de fusies met 2, 3, 6, 7, 13 en 14 aminozuurlinkers (omzetcijfers van 20.000-25.000). De omzetcijfers van de reactie gekatalyseerd door de niet-gefuseerde enzymen was ongeveer 12.000.

In **Hoofdstuk 3** hebben we de reactiviteit van flavoenzymen met zuurstof ($O_2$) bestudeerd, een boeiend onderwerp voor flavoenzymologen. Zuurstof is een oxidatiemiddel dat enzymatisch kan worden geactiveerd en gereduceerd tot waterstofperoxide of water door opeenvolgende overdrachten elektronen (resp. 2 and 4 electronen). In sommige gevallen kunnen andere reactieve zuurstofsoorten (ROS) worden gevormd, zoals superoxide. In **Hoofdstuk 3** onderzochten we ROS-vorming —of ontkoppeling— tijdens de flavine-gebaseerde reductie van zuurstof door flavoproteïne oxidasen en monooxygenasen, met $PAMO_{WT}$, $PAMO_{C65D}$, EUGO en HMFO als testenzymen. De analyse van de ontkoppeling in flavoenzymen is van groot belang omdat ROS een belangrijke rol spelen in de biologie en het gebruik van flavoenzymen als biokatalysatoren in de weg kunnen staan. Bovendien wordt het mechanisme van de ROS formatie door flavoenzymen nog niet volledig begrepen.

Voor alle bestudeerde flavoenzymen werden daarom nauwkeurige profielen bepaald van de productie van waterstofperoxide en superoxide onder verschillende operationele omstandigheden. Het is opmerkelijk dat alle eiwitten aanzienlijke hoeveelheden superoxide bleken te produceren. Bovendien werden er bij een hogere pH, verhoogde superoxideniveaus gedetecteerd. Dit zou een indicatie kunnen zijn voor pH-gevoelige dissociatie van gekooide radicaalparen. De ophoping van superoxide had echter geen nadelig effect op de biokatalyse. Interessant is dat voor PAMO$_{WT}$ en EUGO de toevoeging van katalase de katalytische prestaties significant verhoogde. Al deze resultaten samen geven een beter zicht op de omstandigheden die de vorming van ROS in flavoenzymen bevorderen. Dit is relevant voor de industrie en zou kunnen helpen de vorming van gevaarlijke ROS te verminderen met als gevolg minder van de aanwezige reducerende cofactoren te verspillen.

Het tweede deel van dit proefschrift behandelt de identificatie en karakterisatie van een aantal recent ontdekte monooxygenases die flavine bevatten. Dergelijk onderzoek naar nieuwe flavoenzymen geeft steeds meer inzicht in de chemie die mogelijk is met natuurlijke enzymen en kan leiden tot nieuwe toepassingen. In **Hoofdstuk 4** is in samenwerking met het Microbiologisch Instituut van de ETH (Zürich, Zwitserland) de rol van BVMO'en in de productie van enkele specifieke polyketiden bestudeerd. We ontdekten dat zuurstofinbouw in een nieuwe polyketide-keten wordt bereikt door bacteriële enzymen via een Baeyer-Villiger oxidatie. De biochemische eigenschappen van twee van zulke BVMO'en, Oock en LmbC-Ox, werden vastgesteld. Deze FAD-bevattende enzymen bleken betrokken te zijn bij een zuurstof-insertiereactie die deel uitmaakt van de biosynthese van de secundaire metabolieten oocydine en lobatamide.

**Hoofdstuk 5** beschrijft de inspanningen om veelbelovende type I of type II flavine-bevattende monooxygenasen (FMO'en) te vinden. Door middel van genoom "mining" hebben we twee eiwitten geïdentificeerd van respectievelijk de *Chloroflexi* bacterie en het beerdiertje *Hypsibius dujardini*: CbFMO (type II) en HdFMO (type I). Beide enzymen bleken onderscheidende biochemische kenmerken te hebben. HdFMO bleek alleen actief te zijn met sulfiden, accepteerde alleen NADPH als hydridedonor en vertoonde een gematigde thermostabiliteit ($T_M^{app}$ van 45 °C). Daarentegen zette CbFMO bij voorkeur ketonen om in de respectievelijke esters of lactonen, en vertoonde het een slechte thermostabiliteit ($T_M^{app}$ van 34 °C). De motivatie voor het bestuderen van CbFMO, een type II FMO, was deels vanwege eerder onderzoek naar de binding van nicotinamide door type II FMO'en. Er werd eerder namelijk voor de FMO'en een lage coenzym bindingsspecificiteit gerapporteerd[7]. Toch bleek CbFMO een sterke voorkeur te hebben voor NADPH. Vergeleken met de reeds beschreven flavoproteïne monooxygenasen, lijken beide FMO'en daarom niet erg aantrekkelijk voor biokatalyse.

Ten slotte werden in **Hoofdstuk 6** twee type I BVMO'en van *Streptomyces leeuwenhoekii* C34 geïdentificeerd, namelijk: Sle_13190 en Sle_62070. Beide enzymen werden succesvol tot expressie gebracht met fosfietdehydrogenase als fusiepartner. Net als bij andere type I BVMO'en vertoonden beide eiwitten NADPH-afhankelijke Baeyer-Villiger oxidatie. De sequenties van Sle_13190 en Sle_62070 clusteren, gebaseerd op sequentiehomologie, met andere BVMO'en waarvan reeds bekend is dat ze actief zijn op vrij grote substraten. In deze context was het niet verrassend dat ze tamelijk complexe verbindingen als substraat accepteerden, waaronder bifenylen en een steroïde. Daarbij bleken beide enzymen matig robuust te zijn, met een $T_M^{app}$ van 45 °C en tolereerden ze aanwezigheid van met water mengbare oplosmiddelen. In het bijzonder bleek Sle_62070 zeer actief te zijn op cyclische ketonen en vertoonde het een hoge regioselectiviteit die slechts één lacton produceerde uit 2-fenylcyclohexanon. Ook vertoonde het een hoge enantioselectiviteit en produceerde het alleen normale (-)-1*S*,5*R* en abnormale (-)-1*R*,5*S* lactonen (e.e. > 99%) uit bicyclo [3.2.0] hept-2-en-6-one. Deze twee nieuw-ontdekte BVMO'en kunnen uitgroeien tot waardevolle toevoegingen aan de bekende verzameling BVMO'en.

Het werk dat in dit proefschrift wordt beschreven, heeft verscheidene nieuwe enzymen opgeleverd die kunnen worden gebruikt als biokatalysatoren. Dit proefschrift geeft daarnaast nieuwe inzichten in de katalytische eigenschappen van deze nieuwe enzymen. Het werk bevestigt dat in natuurlijke extreme omstandigheden vele onaangeboorde robuuste enzymen kunnen worden gevonden, zoals aangetoond met de twee BVMO'en die zijn geïdentificeerd in een bacterie die is geïsoleerd uit de Atacama-woestijn (**Hoofdstuk 6**). Ook werd een flavoenzym verkregen en bestudeerd uit een beerdiertje, een microscopisch klein dier dat kan worden gecryopreserveerd en nog niet eerder werd beschouwd als bron van biokatalysatoren. Helaas leverde dit geen veelbelovende biokatalysator op. Het laat zien dat organismen die extreme omstandigheden kunnen overleven niet altijd robuuste biokatalysatoren herbergen. Naast het onderzoeken van het genoom van (thermofiele) micro-organismen, kan het toepassen van een metagenomische benadering ook nuttig zijn als men op zoek is naar een robuuste biokatalysator. De huidige metagenomische benaderingen maken het mogelijk om met verschillende typen biosynthetische donkere materie om te gaan, waar nieuwe bio(chemie) uit kan ontstaan [8,9]. Op die manier zal het bundelen van krachten in de zoektocht naar- en in het ontwerpen van nieuwe enzymen, meer mechanistische inzichten opleveren en de biokatalytische reacties nog efficiënter, betrouwbaarder en milieuvriendelijker maken. Hierdoor kan biokatalyse terrein winnen om op een duurzame manier te concurreren met klassieke chemische routes, of totaal nieuwe enzymgebaseerde toepassingen mogelijk maken.

A

## LITERATUUR

1. Mihovilovic, M. D. Enzyme mediated Baeyer-Villiger oxidations. *Curr. Org. Chem.* 10, 1265–1287 (2006).

2. Kirschner, A. & Bornscheuer, U. T. Directed evolution of a Baeyer-Villiger monooxygenase to enhance enantioselectivity. *Appl. Microbiol. Biotechnol.* 81, 465–472 (2008).

3. Leisch, H., Morley, K. & Lau, P. C. K. Baeyer-Villiger monooxygenases: More than just green chemistry. *Chem. Rev.* 111, 4165–4222 (2011).

4. Alphand, V., Carrea, G., Wohlgemuth, R., Furstoss, R. & Woodley, J. M. Towards large-scale synthetic applications of Baeyer-Villiger monooxygenases. *TRENDS Biotechnol.* 21, 318–323 (2003).

5. Romero, E., Gómez Castellanos, J. R., Mattevi, A. & Fraaije, M. W. Characterization and crystal structure of a robust cyclohexanone monooxygenase. *Angew.Chem.Int.* 55, 15852–15855 (2016).

6. Aalbers, F. S. & Fraaije, M. W. Enzyme fusions in biocatalysis: coupling reactions by pairing enzymes. *ChemBioChem* 20, 20–28 (2019).

7. Riebel, A., Fink, M. J., Mihovilovic, M. D. & Fraaije, M. W. Type II flavin-containing monooxygenases: A new class of biocatalysts that harbors Baeyer–Villiger monooxygenases with a relaxed coenzyme specificity. *ChemCatChem* 6, 1112–1117 (2014).

8. Zhang, J. J. & Moore, B. S. Natural products: Digging for biosynthetic dark matter. *Elife* 4, e06453 (2015).

9. Idris, H., Goodfellow, M., Sanderson, R., Asenjo, J. A. & Bull, A. T. Actinobacterial rare biospheres and dark matter revealed in habitats of the chilean Atacama Desert. *Sci. Rep.* 7, (2017).

## RESUMEN EN ESPAÑOL

En los últimos años se han desarrollado una variedad de enzimas para una gran cantidad de aplicaciones biotecnológicas, haciendo que hoy en día, la biocatálisis tenga un papel fundamental en la sociedad. Los nuevos avances en biocatálisis tienen como objetivo reducir el impacto ambiental de los procesos químicos. En este contexto, ya hace mucho tiempo que la biocatálisis abandonó la mera curiosidad academica; hoy en día se ha convertido en un campo maduro y ampliamente explotado, como por ejemplo, por las industrias química y farmacéutica. Esto ocurrió porque los biocatalizadores (enzimas) permiten un control estricto sobre la selectividad de sus reacciones, esto hace a las enzimas alternativas exquisitas para catálisis química. Por otro lado, también contribuyen al desarrollo de una industria más ecológica y a destapar *nueva* química que está aún sin explotar.

Las flavoenzimas son un excelente ejemplo de biocatalizadores biotecnológicamente interesantes. Estas enzimas son máquinas biomoleculares altamente versátiles que realizan una amplia gama de reacciones y representan una gran cantidad de los biocatalizadores que actualmente se encuentran disponibles para la química RedOx. La flavoenzimología es un campo multidisciplinario, cuyo objetivo es aumentar la comprensión general de estas enzimas que contienen flavina como cofactor. Esto último, bajo un profundo enfoque en el estudio de los mecanismos catalíticos, propiedades cinéticas y relaciones de secuencia-estructura-función. Una mejor comprensión de estas enzimas ha permitido el desarrollo de biocatalizadores mejorados.

En el **Capítulo 1**, las ventajas de la biocatálisis para las reacciones de oxidación de tipo Baeyer-Villiger son bien ejemplificadas. Aunque el método químico para tales oxidaciones ofrece una ruta sintética simple para oxidar cetonas en ésteres o lactonas, requiere el uso de reactivos peligrosos. Además, son típicamente no quimio-, regio- o enantioselectivos. Como alternativa atractiva, las monooxigenasas de tipo Baeyer-Villiger (BVMOs) que contienen flavina como grupo prostético, se han investigado durante mucho tiempo por su alto potencial biocatalítico[1-5]. En las últimas décadas se han descubierto o diseñado varias BVMOs que exhiben actividades adecuadas para la síntesis de compuestos valiosos. En el **capítulo 1**, resumimos el estado del arte del uso de estas enzimas como biocatalizadores, nos centramos así en sus propiedades bioquímicas, mecanísticas y estructurales. Aunque pareciera ser un campo relativamente maduro, todavía existen desafíos que los flavoenzimólogos deben desentrañar. Limitaciones como estabilidad moderada, baja especificidad y la dependencia de cofactores, a menudo disminuyen las opciones de ser usadas a nivel industrial. De todos modos, actualmente, se está avanzando en la superación de todas estas limitaciones.

**A**

Muchas (flavo)enzimas dependen de cofactores disociables para la catálisis, como el NAD(P)H. Los costos relacionados con tal dependencia de cofactores pueden superarse en parte, mediante el uso de enzimas bifuncionales autosuficientes que regeneran el cofactor de manera eficiente. En el **capítulo 2**, diseñamos un protocolo experimental para evaluar el efecto de la longitud de un *linker* flexible rico en glicinas sobre las propiedades de fusiones de TbADH y TmCHMO. Este protocolo experimental permite la generación de biocatalizadores optimizados. Además, el procedimiento se puede modificar fácilmente para generar una biblioteca similar de otras proteínas de fusión. Como reacción de demostración, las enzimas de fusión ADH-BVMO generadas, se analizaron para la síntesis de ε-caprolactona, un precursor de polímero valioso, usando ciclohexanol como sustrato inicial de la cascada. Tal reacción aprovecha el hecho de que ambas actividades enzimáticas satisfacen su dependencia de cofactor: la actividad ADH convierte NADP$^+$ en NADPH, mientras que, a su vez, la actividad BVMO oxida NADPH en NADP$^+$. Además, los biocatalizadores bifuncionales fusionados pueden mejorar la estabilidad de la proteína y promover la canalización de intermediarios del producto por efectos de proximidad. Vale la pena mencionar, que no existen reglas uniformes para diseñar *linkers* óptimos para proteínas de fusión. Sin embargo, un enlazador "incorrecto" puede provocar efectos perjudiciales sobre el rendimiento biocatalítico[6]. El trabajo presentado se enfocó en la evaluación del efecto de quince variantes de *linker* de diferente largo sobre la: expresión, termoestabilidad, actividad y niveles de conversión. Todas las variantes obtenidas exhibieron altos niveles de expresión (250-360 mg L$^{-1}$) y se obtuvieron principalmente como holoproteínas (según la razón $A_{280}$:$A_{440}$). Tanto para TmCHMO como para TbADH, se observó que la longitud del *linker* no mostró un efecto perjudicial significativo sobre su termoestabilidad. En cuanto a la actividad, las actividades de oxidación del alcohol (ADH) y sulfoxidación (BVMO) fueron similares para las 9 variantes más cortas, mientras que las fusiones con la longitud del conector de 10, 12 y 15 aminoácidos mostraron una actividad ligeramente aumentada. Las bioconversiones a pequeña escala dieron como resultado *turnover numbers* (TON) más altos para las fusiones con *linker* de 2, 3, 6, 7, 13 y 14 aminoácidos (TON de 20.000-25.000). El TON de la reacción control catalizada por las enzimas no fusionadas fue de alrededor de 12.000.

En el **capítulo 3**, estudiamos la reactividad de flavoenzimas con dioxígeno, un tema fascinante para la flavoenzimólogia. El dioxígeno es un oxidante de cuatro electrones que puede activarse enzimáticamente y reducirse a peróxido de hidrógeno o agua mediante transferencias consecutivas de un electrón. En algunos casos, se pueden formar otras especies reactivas de oxígeno (ROS), como el superóxido. En el **capítulo 3**, investigamos la formación de ROS —o desacoplamiento— durante la reducción de dioxígeno mediada por el cofactor de flavoproteínas (oxidasas y monooxigenasas), utilizando PAMO$_{WT}$, PAMO$_{C65D}$, EUGO y HMFO como enzimas a analizar. El análisis del desacoplamiento en flavoproteínas es de gran relevancia porque el ROS tienen un papel relevante en biología (transduccion de señales), por otro lado, puede complicar el uso de flavoenzimas como

biocatalizadores. Además, el mecanismo de formación de ROS mediada por flavoenzimas aún no se comprende completamente.

Para este capítulo, se determinaron perfiles precisos de producción de peróxido de hidrógeno y superóxido en diferentes condiciones operativas para todas las flavoenzimas estudiadas. Sorprendentemente, se descubrió que todas las proteínas producen cantidades significativas de superóxido. Además, se detectaron mayores de esta molécula a pH más altos, esto sugiere una disociación de los pares de radicales sensible al pH. A pesar de la acumulación de superóxido, no se demostró ningún efecto perjudicial de este sobre la biocatálisis. Curiosamente, para PAMO$_{WT}$ y EUGO, la adición de catalasa aumentó significativamente el rendimiento catalítico. Los resultados proporcionan una mejor visión de las condiciones que promueven la formación de ROS en flavoenzimas y podría ayudar a reducir la formación de estas especies a nivel industrial, evitando el desperdicio de valiosos equivalentes reductores.

La segunda parte de esta tesis trata de la identificación y caracterización de varias monooxigenasas que contienen flavina. Estudios sobre nuevas flavoenzimas proporcionan más información sobre la química con enzimas naturales y también pueden conducir a nuevas aplicaciones biocatalíticas. En el **capítulo 4**, en colaboración con el Instituto de Microbiología del ETH (Zürich, Suiza), se estudió la participación de BVMO en la producción de algunos policétidos específicos. Se demostró que BVMO de origen bacteriano eran capaz de insertar un atomo de oxígeno a través de una oxidación de Baeyer-Villiger en esqueletos de policétido naciente. Se establecieron las propiedades bioquímicas de dos BVMOs: Oock y LmbC-Ox. Finalmente, se demostró que estas flavoenzimas están involucradas en la biosíntesis de los metabolitos secundarios oocidina y lobatamida.

El **Capítulo 5** describe los esfuerzos para encontrar prometedoras *flavin-containing monooxigenases* (FMOs) de tipo I o tipo II. Mediante minería genómica, identificamos dos proteínas: una de origen bacteriano (*Chloroflexi)* y otra proveniente del tardígrado *Hypsibius dujardini*: CbFMO (FMO de tipo II) y HdFMO (FMO de tipo I), respectivamente. Ambas enzimas mostraron características bioquímicas distintas. HdFMO mostró solo actividad con sulfuros, y solo aceptó NADPH como donante de hidruro, además de presentar una termoestabilidad moderada (T$_M$$^{app}$ de 45 ° C). Por el contrario, CbFMO convirtió preferentemente cetonas en los respectivos ésteres o lactonas y mostró una baja termoestabilidad (T$_M$$^{app}$ de 34 ° C). La motivación para estudiar CbFMO, un FMO de tipo II, se debió en parte, a que este grupo de enzimas han sido descritas con una promiscua especificidad por el cofactor de nicotinamida[7]. Sin embargo, se encontró que CbFMO tiene una fuerte preferencia por NADPH. Por lo tanto, en comparación con la colección ya descrita de monooxigenasas de flavoproteínas, ambas FMO no parecieron muy atractivas para procesos biocatalíticos.

Finalmente, en el **capítulo 6,** se identificaron dos BVMO de tipo I de *Streptomyces leeuwenhoekii* C34: Sle_13190 y Sle_62070. Ambas enzimas se expresaron con éxito, fusionadas con un  regenerador de cofactor (fosfito deshidrogenasa). Al igual que otros BVMO de tipo I, ambas proteínas mostraron oxidación de Baeyer-Villiger dependiente de NADPH. Las secuencias de Sle_13190 y Sle_62070 se asociaron basándose en la homología de secuencia, con otras BVMOs descritas que actúan sobre compuestos voluminosos. En este contexto, no fue sorprendente que aceptaran como sustrato compuestos bastante complejos, incluidos bifenilos y un esteroide. Además, se encontró que ambas enzimas eran moderadamente robustas, exhibiendo una $T_M^{app}$ de 45 ° C y tolerancia a cosolventes miscibles en agua. En particular, se encontró que Sle_62070 es altamente activo con cetonas cíclicas y mostró una alta regioselectividad produciendo solo la  lactona de 2-fenil-ciclohexanona, y una alta enantioselectividad para la conversion de  biciclo[3.2.0]hept-2-en-6-ona produciendo solo las lactonas (-)-1*S*, 5R normal y (-)-1*R*, 5*S* abnormal (*e.e.* > 99 %). Estas dos BVMOs recién descubiertas pueden convertirse en valiosas adiciones a la colección de BVMOs.

El trabajo descrito en esta tesis proporcionó varias enzimas nuevas, que pueden emplearse como biocatalizadores, además de nuevos conocimientos sobre sus propiedades catalíticas. El trabajo recalca que los ambientes extremos pueden ser una gran fuente de enzimas robustas sin explotar, como se demostró con las dos BVMOs identificadas en una bacteria aislada del desierto de Atacama (**capítulo 6**). Además, se obtuvo y estudió una flavoenzima de un tardígrado, un animal microscópico que se puede criopreservarse, el cual no se había considerado antes como fuente de biocatalizadores. Desafortunadamente, esta busqueda no condujo a la obtencion de un catalizador muy prometedor. Muestra de que los organismos que pueden sobrevivir en condiciones extremas no siempre (solo) albergan biocatalizadores robustos. Además de explorar las secuencias de genomas de microorganismos (termófilos), el empleo de un enfoque metagenómico también puede ser de alta útilidad para la busqueda de enzimas con enfoque industrial. Los enfoques metagenómicos actuales permiten el manejo de fuentes ricas en *materia biosintetica oscura*, lo que puede conducir a descubir nueva (bio)química[8,9]. En este sentido, el esfuerzo combinado en el descubrimiento de nuevas enzimas, estudios de mecanismos y la ingeniería de enzimas hará que las reacciones biocatalíticas sean aún más eficientes, fiables y amigables con el medio ambiente. Esto permitirá que la biocatálisis sea un competidor sustentable a las rutas químicas clásicas, permitiendo nuevas aplicaciones basadas en enzimas a nivel industrial.

# REFERENCIAS

1. Mihovilovic, M. D. Enzyme mediated Baeyer-Villiger oxidations. *Curr. Org. Chem.* 10, 1265–1287 (2006).

2. Kirschner, A. & Bornscheuer, U. T. Directed evolution of a Baeyer-Villiger monooxygenase to enhance enantioselectivity. *Appl. Microbiol. Biotechnol.* 81, 465–472 (2008).

3. Leisch, H., Morley, K. & Lau, P. C. K. Baeyer-Villiger monooxygenases: More than just green chemistry. *Chem. Rev.* 111, 4165–4222 (2011).

4. Alphand, V., Carrea, G., Wohlgemuth, R., Furstoss, R. & Woodley, J. M. Towards large-scale synthetic applications of Baeyer-Villiger monooxygenases. *TRENDS Biotechnol.* 21, 318–323 (2003).

5. Romero, E., Gómez Castellanos, J. R., Mattevi, A. & Fraaije, M. W. Characterization and crystal structure of a robust cyclohexanone monooxygenase. *Angew.Chem.Int.* 55, 15852–15855 (2016).

6. Aalbers, F. S. & Fraaije, M. W. Enzyme fusions in biocatalysis: coupling reactions by pairing enzymes. *ChemBioChem* 20, 20–28 (2019).

7. Riebel, A., Fink, M. J., Mihovilovic, M. D. & Fraaije, M. W. Type II flavin-containing monooxygenases: A new class of biocatalysts that harbors Baeyer–Villiger monooxygenases with a relaxed coenzyme specificity. *ChemCatChem* 6, 1112–1117 (2014).

8. Zhang, J. J. & Moore, B. S. Natural products: Digging for biosynthetic dark matter. *Elife* 4, e06453 (2015).

9. Idris, H., Goodfellow, M., Sanderson, R., Asenjo, J. A. & Bull, A. T. Actinobacterial rare biospheres and dark matter revealed in habitats of the chilean Atacama Desert. *Sci. Rep.* 7, (2017).

**A**

## LIST OF PUBLICATIONS

Durán-Toro, V., **Gran-Scheuch, A.**, Órdenes-Aenishanslins, N., Monrás, J.P., Saona, L.A., Venegas, F.A., Chasteen, T.G., Bravo, D. & Pérez-Donoso, J.M. Quantum dot-based assay for $Cu^{2+}$ quantification in bacterial cell culture. *Analytical biochemistry* 450 (**2014**): 30-36.

**Gran-Scheuch, A.**, Fuentes, E., Bravo, D., Jiménez, J.C. & Pérez-Donoso, J.M. Isolation and characterization of phenanthrene degrading bacteria from diesel fuel-contaminated Antarctic soils. *Frontiers in microbiology* 8 (**2017**): 1634.

**Gran-Scheuch, A.**, Trajkovic, M., Parra, L. & Fraaije, M.W. Mining the genome of *Streptomyces leeuwenhoekii*: Two new type I Baeyer–Villiger monooxygenases from Atacama Desert. *Frontiers in microbiology* 9 (**2018**): 1609.

Fürst, M.J.L.J., **Gran-Scheuch, A.**, Aalbers, F.S., & Fraaije, M.W. Baeyer–Villiger monooxygenases: Tunable oxidative biocatalysts. *ACS Catalysis* 9 (**2019**): 11207-11241.

Reiko, U., Meoded, R.A., **Gran-Scheuch, A.**, Bhushan, A., Fraaije, M.W. & Piel, J. Genome mining of oxidation modules in trans-acyltransferase polyketide synthases reveals a culturable source for lobatamides. *Angewandte Chemie International Edition* 59 (**2020**): 7761-7765.

**Gran-Scheuch, A.**, Aalbers, F.S, Woudstra Y., Parra, L. & Fraaije, M.W. Optimizing the linker length for fusing an alcohol dehydrogenase with a cyclohexanone monooxygenase. *Methods in Enzymology* 647(**2020**). Academic Press.

**Gran-Scheuch, A.**, Ramos-Zuñiga, J., Fuentes, E., Bravo, D., & Pérez-Donoso, J.M. Effect of Co-contamination by PAHs and Heavy Metals on Bacterial Communities of Diesel Contaminated Soils of South Shetland Islands, Antarctica. *Microorganisms* 8 (**2020**), 1749.

## ACKNOWLEDGMENTS/ AGRADECIMIENTOS

I would like to thank the many people who have helped me over the past few years to obtain my Ph.D. and for making this experience not only academically better, but also, enjoyable and easier through scientific discussions and biertjes.

Primero que todo, quiero agradecer a la Agencia Nacional de Investigación y Desarrollo (ANID, ex-CONICYT), en particular, por la beca otorgada a través del concurso de becas de doctorado nacional, año académico 2015 (beca #21151392) y al programa de beneficios adicionales de estadía de cotutela en el extranjero.

Dear **Marco**, first of all, thank you so much for giving me the opportunity to be part of your research group. This has been an unique experience and I'm extremely grateful for that. I must say that by far you've been the best boss I've ever had, I greatly appreciate your advices and guidance that helped me to finish this work (it goes without saying, that I'm infinitely grateful for your lightning-fast corrections of manuscripts /presentations /posters). It was a pleasure to have a supervisor who combined patience, humor and passion for science with an excellent balance. I would also like to thank you for all the times we shared during the lab retreat (and the times that I won playing grote dalmuti that made me realize that I couldn't handle absolute power). Thanks also to the amazing hosts **Marleen**, **Laura** and **Niek** for being part of the times you invited us for dinner (or for sinterklaas or beer tasting or board games) at your place.

**Loreto**, muchísimas gracias por haberme aceptado como tu primer estudiante de doctorado. El comienzo fue un tanto diferente, como estabas comenzando con tu grupo de investigación, me tocó no solo jugar de alumno de doctorado, sino que también de técnico de laboratorio, lab manager, encargado del wifi, mover cajas, barrer y estandarizar los equipos. Esto sin duda fue una experiencia más que interesante, y creo que, al fin y al cabo, amplió mi paquete de habilidades blandas. Además, te quiero agradecer lo apañadora que fuiste desde un principio, ya sea, por cuando fuimos a Massachusetts por el Global UC o gestionando la cotutela en Groningen (en este punto en particular, también quiero agradecer a **Pedro Bouchon** por apurar el acuerdo RuG-UC). Loreto, te deseo mucho éxito en tus próximas aventuras tanto académicas como personales.

Dear **Dick**, I would like to thank you for your useful suggestions and ideas during the work discussions. Although we did not share research projects together, your critical way of thinking partly shaped my scientific development. I really appreciate that you have given me an exciting new challenge, which I hope will be very fruitful.

**A**

Quisiera agradecer a los integrantes de mi comité de candidatura (UC), Prof. **Mónica Vásquez**, Prof. **Juan Asenjo**, Prof. **Cesar Ramírez**, Prof. **Jorge Vásquez** y Prof. **Juan de Dios**. Muchas gracias por su tiempo y disposición al coordinar las fechas de candidatura y avance. Por sus sugerencias, que me ayudaron a mejorar este trabajo de manera significativa. En particular a la Prof. **Mónica,** por preocuparse por mi bienestar tanto académico como personal. Y a **César**, muchas gracias por tus sugerencias y por haberme resuelto muchas dudas durante la primera etapa de mi doctorado (fuiste de muchísima ayuda). I would like to express my deepest gratitude to the reading committee (RuG), Prof. **van Berkel** and Prof. **Driessen,** thank you for your time and the valuable comments on my thesis. Dear collaborators, **Roy** and Prof. **Piel**, thank you very much for inviting me to be part of our joint-project, which happily ended in a very interesting paper (chapter 5).

Dear officemates (104s), **Milos**, **Hemant**, **Qinglong, Mohammed, Friso** and **Yiming. Milos,** my squash partner, thank you for your friendship (moving the wardrobe was not easy) and for your helpful character, you are a great scientist, but above all, a terrific person. I'm very happy to have met you and your family. I wish you, **Jasna** and little **Sofi** all the best in The Netherlands! **Hemant,** you are an incredible humble and chill guy, I really appreciated our conversations in the office and on coffee breaks (by the way, you should improve your reflexes when people faint in front of you). **Qinglong,** I'm infinitely grateful that you were my dealer of competent cells (the best competent cells on the block!). Also thank you very much for teaching me a bit of chinese, although I failed in the attempt (xiè xiè!). I wish all the best to you and your family! **Mohamed**, I really appreciate our conversations about our different cultures, and from time to time about football, you are a very calm and nice person, stay that way! I wish you all the best in your new projects and challenges! **Friso**. I want to thank you for our projects together (chapter 1 and 2) that turned out to be very good at the end. Also, thanks for teaching me a bit about the complex Dutch culture (even though you are more Frisian than Dutch). I really appreciated that you wanted to learn from my sparse Spanish, I'm also very grateful for your openness with me. **Nina,** thanks for all the good times! **Yiming,** you are a very cheerful and smiley guy with a contagious laugh, I hope you continue like this (thank you for continuing expanding my chinese).

**Nikola,** amigo, thank you for organizing so many bbq and/or borrels, it was great that you were always making an effort to gather people (even if it was to watch an Iranian movie), I hope you keep doing it (someone has to …). Thank you for the advices and the good vibe. I'm sorry I didn't want to join the cooking club, but it seemed like a lot of effort (jajaja). I wish you all the best with GECCO and with your next challenges. **Maxi,** I would first like to thank you for your professional maturity and great organization skills during the writing of the review, I think thanks to you it came out fast and neat (although sometimes it hurt to see the corrections). It was always fun to party and share with you; the most memorable time for me was when you brought meat to a NON-barbecue party. I wish you all the best in UK

(and please forget about the damn plants of your old place!). I miei **Principessos,** you two were a great editing team during my Kubrick airs. **Simone,** you will always be my favorite drag partner, we may have a future in that field if nothing else goes well. I really appreciate your humor and thanks for helping me whenever I needed it with the AKTA. Also, I really enjoyed our conversations in and outside of the lab and…. just writing to you these words made me thirsty, when are we going to The Crown? **Ivana**, you two make a perfect match, sorry if I spoke to you in Spanish when I was at parties (but it seemed to me like you could speak Spanish fluently). Thanks for adopting me the summer that I was alone! **Nikk,** il princippeso di Sardegna, bonjour! You are a great friend, I really like your humor and your cooking skills (I enjoyed them both during the lab retreats or personalized at my place), I hope I could go to Sardegna someday (Sardegna, not Italy, is different, of course) and try some of your seafood (thank you also for making it clear to me that I cannot distinguish between *B* and *V*). **Lilas**, thanks for the good times, it was nice meeting you! **Cate,** you have an underrated patience…well, not really, but you make the best pizzas in town! Thank you for being a great friend and a nice colleague, I admire your organization skills and efficiency, I would have liked to receive a bit of them by osmosis, but it did not happen. I wish you all the best working at GECCO! **Fermín,** tío, eres muy majo y siempre dispuesto a celebrar, gracias por eso y por los buenos momentos, ¡les deseo lo mejor!

**Elvira**, eres monísima, muy esforzada y humilde, no hace falta decir que te encuentro una excelente científica, te quiero agradecer por haberme ayudado y enseñado un kilo sobre el mundo de las flavoproteinas cuando llegué al lab; tus consejos (tanto científicos como personales) me ayudaron muchísimo. Además, me encanta como dejas todo en la cancha cuando es hora de carretear (rumbear)! **Cora**, sos regrosa!, aunque al principio traté de hablar contigo en un español más menos neutral (¿durante las primeras semanas quizás?), después fue simplemente en chileno, y nos (me) ayudó a tener una comunicación mucho más fluida, ¡te veo como una latina más! Encuentro que eres una persona la zorra, fue la raja compartir y carretear contigo. ¡Les deseo lo mejor en UK! Kaixo **Lur!**, encuentro que eres muy motivada y trabajadora, muchas gracias por tu buena onda y alegría, ¡espero que pronto podamos estar celebrando en San Sebastián! **Estela,** eres una persona muy simpática y entretenida, fue siempre un agrado conversar contigo, te deseo todo lo mejor con **Jan,** ¿ya sea aquí o quizás en México?

**Elvira**, thanks for all the good times! It was a lot of fun sharing office with you (at least until I was sadly kicked out), non era silenziosa ma almeno era divertente. I wish you all the best! **Jeroen,** it was great having you around and hanging out with you in the lab, mostly during this last period. All the best for you and Nate (I hope he never makes you a tuna casserole again). **Edwin,** I would like to thank you for your help correcting a few of my chapters during the end of my PhD, they were very quick and useful! **Titia,** you are a very happy and cheerful person, thanks again for your help with the samenvattig! **Rudi**, thank you for your

happiness and kindness. You are a very enthusiastic person, I know you will do very well. **Hugo**, thank you for your advice and help, you often came up with insightful suggestions that were handy for my labwork. **Yapapei**, you are a very nice person, thank you for your willingness in the lab, I'm very curious to know how you will end up bossing everyone in the lab, I wish you all the best! **Hein,** I find you incredibly smart, I really appreciated your suggestions and comments thinking outside the box. I hope we have great results in our next project together. I would like to thank **Sandra** and **Chienes** for helping me with many administrative stuffs (you made my life much easier). Additionally, I would also like to thank **Pim**, **Gabriela**, **Paulina and Pranoti** for letting me be part of your academic training, but also, because being your supervisor was also very educational for me, good luck in the future! Thanks to RuG for giving us free printing and coffee, which I estimate I drank at least around 2,700 so far, it was the driving force of every day.

I would like to thank all current and former members of the groups of Marco, Dick and people from upstairs that I have not mentioned yet: **Clemens**, **Ilias, Elisa**, **Peter**, **Ivana**, **Vakil**, **Dana**, **Andy, Eduardo, Carlos**, **Ulises**, **Tamara**, **Engel, Christiaan**, **Henriette**, **Antonija**, **Brenda**, **Daniël**, **Guang**, **Bart, Roxana, Claudia, Reuben, Alessandro, Xiaoyu, Andrea** & **Nyoman** for contributing to my work through discussions, chit-chats and administration issues. Thank you for making my PhD life not only better but also fun.

I would also like to thank my vecinos **Marcel** and **Luca**. **Marcelzinho**, eres una excelente persona, muchas gracias por las veces que nos invitaste a tu casa y por tu increíble disposición, y mil gracias por cuidar de nuestras plantas. **Luca**, thanks for all the good times! I'm glad that you accomplished your mermaid dream. También me gustaría agradecer a la gente de la UC: **Dani, Andrea, Blasquez, Vicente, Luis, Sergio, Tamara, Isaac, Ariel, Aníbal** y **Luis.** Muchas gracias por la buena onda, y a la ayuda en general. En especial a la **Dani**, muchísimas gracias por tu buena onda y por hacerme todos los favores administrativos que te pedí durante estos últimos años. Te deseo mucho éxito en tus próximos proyectos.

Quiero agradecerle a **Rodri** y **Feli**, los quiero un kilo, Uds. han sido en cierto grado parte de nuestra familia en Groningen, me encantó haberlos conocido y que pudiésemos hacer juntos una absurda variedad de distintas cosas, desde tomar oncecita los domingos, ¡a viajar a Egipto!, o terminar en el McDonalds a las 1 de la mañana, eso que teníamos planeado ver una película tranqui en la tarde. Les deseo lo mejor en sus próximos desafíos, y espero que algún día Feli active su filtro social, me parece que vino así de fábrica, pero puede ser reprogramado. **Patri,** querida! ¡Te encuentro una persona genial! Muchísimas gracias por organizar tantas tardes en tu casa, ¡han sido siempre muy entretenidas! Me encanta que seas tan preocupada por los tuyos y que compartamos nuestro lado geek (nunca he podido entender qué clase de café toman con Mel, espero algún día me lo cuentes) **Alfre,** eres muy buena onda, gracias por habernos preparado esa carne al trapo (ojalá podamos

repetirla) y por invitarnos a la cena de las velitas. Aunque muchas veces te pierdes en tus propios pensamientos (Grolo? Quien Grolo?), se agradece que seas una muy buena persona. **Yannick**, you are my favorite Dutch person! thank you for teaching me such useful and wise stuffs like "*Hola supermercado de la bancos por aquí*" and "*donde esta Genietos*"! I hope you can finish your PhD as soon as possible so we can visit you at your restaurant. Thank you for putting so much effort every time you cooked for us (at your house is where I have eaten the best in The Netherlands). **Christoff**, you are my favorite South African person, I guess? You are a super chill guy with a great sense of humor, and I enjoyed that a lot. I'm so sorry I made you fly home with a horrible hangover (not sorry). Yannick / Christoff, I hope that one day we can meet our spirit guide: Mr. Ferrel (aka volcano man).

**Vichito**, fue la raja haberte tenido cerca y que nos hayamos pinponeado un par de veces entre Groningen y Bremen (siento que hayas perdido tus zapatos, pero porque se te ocurrió venir a escalar?!) ¡Te quiero un montón! **Nadine**, thanks for all the good times! I loved it when you secretly told us to go to Vicho's birthday, I wish you both all the best! **Pancha** y **Tavo,** fue entretenidísimo haberlos tenido tan cerca y que hayamos podido salir juntos de vacaciones, aunque al final todas las catedrales se veían igual (#RoteandoEuropa). Muchas gracias Tavo por ser nuestro fotógrafo personal y darnos mucho material de Instagram (y lamento mucho haberlos aburrido a ti y a la Mel con el infinito y asombroso mundo de los tubos y microplacas). ¡Ojalá podamos organizar otro viaje pronto! También quiero agradecer a mis amigos en Chile, **Waldo (**que en tus propias palabras eres la luz en mi camino**), Diego** (que acompañó cada memiercoles**), Lucho** (tu marca sigue en mi piso, ¿gracias por eso?**), Nico** (que constantemente me haces énfasis que debo aprovechar a concho la oportunidad de estar acá**), Foncho (**que siempre ha sido súper apañador**), Noni y Chaplin**. Creo que ser capaz de soportar los reiterados fracasos durante la tesis fue en parte gracias a Uds. Gene Gunners nos hizo inmune a la frustración. Quisiera agradecer especialmente LMG, muchísimas gracias por su buena onda y sentido del humor, y que cada día se mandaran un *buenos días!* o inclusive solo *BD*, espero poder ir pronto para que tomemos un desayuno ejecutivo (y rápido).

I also want to thank my paranymphs: **Gucci** and **Masun**. **Gautier,** one of our first interactions was at Flavins & Flavoproteins, where we sat together and during one of the speakers' presentation you started to laugh at me (loudly) because I woke up to my own snoring. I think that the second interaction was on the stairs of Dog Bollocks. So we had a kind of belated friendship (mainly because you didn't live in Groningen at first), but then we found out that we had a similar sense of humor and the friendship blossomed. So, thank you so much for being very supportive, and such a good friend, but also, for the morning memes and for sharing many good times! I wish you and **Bernice** all the best in Canada**!**

**Misun**, you are a great friend! Well now, because at first you deliberately ignored me at Vismarkt. But after that, we became very good friends (thank you for choosing me as your paranymph, it was very fun). I really enjoy our conversations in the lab, but also, that we can share the good and bad moments. I want to thank you for your good vibes and that you never hesitated to help me when I had a problem or doubt, even if I had a silly question; or for helping me figuring out about whom people were talking when my memory failed. **Martin**, thanks for all the good times, it's great that you don't have to drive so much anymore! I wish you both all the best in your upcoming challenges & projects.

A mi familia en Chile, a **Fabián** y **Gema**, que durante estos años me han considerado como uno más (a esta altura me imagino que ya me pasaron por la libreta de familia). Gracias por haberme acompañado y porque siempre se han sentido orgullosos de mí, los quiero mucho! Al **Sapolio** y **Nachita**, por ser excelentes cuñis, por su infinito cariño, las tardes de película en Chile o de vez en cuando las llamadas de teléfono cuando iba en la bici. Las quiero mucho, gracias por considerarme como su hermano...aunque no sea su hermano.

A mis papás (**Fernando** y **Cecilia**), gracias por el inmenso apoyo y por estar constantemente preocupados por su concho (ya sea por cómo iba mi tesis, por saber que había hecho el fin de semana o si tenía que abrigarme porque estaba muy helado). A mis hermanes: **Pelao, Bubu, Lolo, Gera, Paula, Fran y Pía,** por estar siempre cuando los he necesitado, y por compartir la buena onda ya sea cuando tomamos once o por WhatsApp, (Gera/Pia/Lolo/Fran, fue genial haberlos tenido en nuestra casa y poder compartir un poco de Groningencito con Uds. Gera/Pía, ojalá la próxima vez vayamos a acantilados de verdad y no ir a aplaudirles a un par de focas). A mis sobrines (**Trini, Pollo** y **Gaspacho**), por ser unas ternuras máximas. Muchísimas gracias por apoyarme desde la distancia, los amo un montón (llamen más seguido si).

**Mel** (mi corazón, slippers), haber viajado a Groningen ha sido infinitamente enriquecedor (tanto académica como personalmente), pero también fue a veces un poco triste estar lejos de la familia, sobre todo en fechas especiales. Sin embargo, durante todos estos años, cuando más lo he necesitado, siempre has sido un soporte incondicional. Me has acompañado y apañado en todas, te agradezco por estar conmigo en las buenas, y también por ser mi pilar en momentos más complicados. Gracias por alentarme a seguir adelante y aconsejarme en las situaciones difíciles. Pero también, gracias por compartir conmigo risas y fiestas. Me encanta que seas una locura mañosa (y un poco ruidosa), que nos tengamos los dos para cuando necesitemos conversar, reír o regalonear. Que hayamos armado nuestra primera casita juntos, que me ames infinitamente de vuelta (como cuando no vuelves a la casa sin un regalo para tu cora) y que hayamos recolectado un montón de recuerdos juntos. Sinceramente, fuiste una ayuda inmensa para lograr terminar esta tesis, siempre has sido mi motor. Gracias, te amo infinito.