Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Leann Quertinmont Teadt Date

Engineering the Stereospecificity and Regioselectivity of Flavoenzymes

By

Leann Quertinmont Teadt
Doctor of Philosophy

Chemistry

_____

Dr. Stefan Lutz
Advisor

_____

Dr. Dennis Liotta
Committee Member

_____

Dr. Emily Weinert
Committee Member

Accepted:

_____

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

_____

Date

Engineering the Stereospecificity and Regioselectivity of Flavoenzymes

By

Leann Quertinmont Teadt
B.A., Carthage College, 2012

Advisor: Dr. Stefan Lutz, PhD

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Chemistry
2018

Abstract

Engineering the Stereospecificity and Regioselectivity of Flavoenzymes

By Leann Quertinmont Teadt

Enzymes are biocatalysts which accelerate biological reactions while typically maintaining high regio-, stereo-, and chemo- selectivity. However, most enzymes have evolved over thousands of years for specificity towards their native substrates. As a result, their usefulness as tools for biocatalysis is limited beyond these native substrates. Therefore, researchers have turned to protein engineering to adapt enzymes to work more efficiently towards desired compounds. This dissertation focuses specifically on engineering flavoenzymes, which are enzymes that contain a flavin cofactor, to alter their stereospecificity or regioselectivity. The specific flavoenzymes this work will focus on are Old Yellow Enzyme 1 (OYE1) from *Saccharomyces pastorianus* and cyclododecanone monooxygenase (CDMO) from *Rhodococcus* sp. HI-31.

Initially, the RApid Parallel Protein EvaluatoR (RAPPER) system was developed to rapidly and semi-quantitatively assess OYE1 variants. Preliminary work with RAPPER also reports OYE1 variants with enhanced activity as a result of increased flexibility in the β6 loop. These studies were followed up by the expansion of the RAPPER system to assess a variety of OYE1 variants. These variant libraries included single amino acid substitutions, deletions, and alanine scanning studies. The importance of substrate profiling was demonstrated as each variant was tested with multiple substrates to more accurately assess the impact of each substitution. As a result of this substrate profiling, several promising variants proved to be highly substrate dependent with respect to their catalytic activity.

Finally, efforts were made to engineer the regioselectivity of CDMO towards an N-protected β-amino ketone. In the course of these engineering efforts, two main regions of interest were identified as important towards determining product regioselectivity. Through a series of amino acid changes, both the chemically preferred (‘normal’) and the ‘abnormal’ products can be obtained with high selectivity. Taken together, these collective engineering efforts illustrate the capacity to specifically tailor biocatalysts for the production of industrially valuable products.

Engineering the Stereospecificity and Regioselectivity of Flavoenzymes

By

Leann Quertinmont Teadt
B.A., Carthage College, 2012

Advisor: Dr. Stefan Lutz, PhD

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Chemistry
2018

Acknowledgements

I would like to thank my advisor Dr. Stefan Lutz for giving me the opportunity to work in his lab and his mentorship throughout the years. You have encouraged me to grow and learn as a scientist throughout my graduate career and helped make my graduate experience unforgettable. I would also like to thank my committee members, Dr. Dennis Liotta and Dr. Emily Weinert for all of their input in my research over the years. Your feedback and guidance have been incredibly helpful. Additionally, I would like to thank all of the members of the Atlanta Flavin Meetings for their input and exciting intellectual discussions over the past several years.

For all of their support throughout my time at Emory, I would like to thank all of the members, both past and present, of the Lutz lab. All of you made being in lab a fun and enjoyable experience, providing me with continuous support, laughter, and penguins since the start of my time in the lab. I couldn't have asked for better people to spend the last several years beside. Samantha Iamurri and Matt Jenkins thank you for all of your help and support since the beginning of my graduate school experience, this journey wouldn't have been the same without either of you. I would like to thank Huanyu Zhao as well for all of the help and materials you've provided to me over the past few years. I also thank Elsie Williams for being not only a wonderful office mate and coffee buddy, but for being a great friend and support system the past few years. Every member of the lab has made my life so much richer, and I have enjoyed working with and getting to know every one of you.

I have developed great friendships outside the Lutz lab throughout my time at Emory, and I thank all of you for your guidance and support over the past several years. All of you, especially Jessica Hurtak, Michael Sullivan, and Joshua and Stephanie Bartlett, made Atlanta feel like home much faster.

Finally, I'd like to thank my family, especially my parents and siblings, who have always believed in me and pushed me to reach for my goals. Your love and encouragement has helped get me where I am today. And last but certainly not least, I would like to thank my husband, Thomas. You have kept me smiling even throughout the most stressful times. Thank you for your endless support and encouragement, your love and support means more to me than I can ever say.

List of frequently used abbreviations

Abbreviation	Full name
BVMO	Baeyer-Villiger Monooxygenase
CASTing	Combinatorial Active-Site Saturation Testing
CDMO	Cyclododecanone Monooxygenase
CHMO	Cyclohexanone Monooxygenase
FAD	Flavin Adenine Dinucleotide
FMN	Flavin Mononucleotide
ISM	Iterative Saturation Mutagenesis
OYE1	Old Yellow Enzyme 1
PAMO	Phenylacetone Monooxygenase
PockeMO	Polycyclic Ketone Monooxygenase
PURE	Protein Synthesis Using Recombinant Elements
RAPPER	RApid Parallel Protein EvaluatoR

Table of Contents

Table of Figures and Schemes

Table of Tables

Chapter 1: General Introduction

General Introduction

Enzymes are catalysts found in biological systems necessary for carrying out the reactions essential to life at a useful timescale, typically maintaining strict regio-, stereo-, and chemo- selectivity[1]. Biocatalysis is the use of enzymes or microbes to perform chemical transformations[2-5]. Since Eduard Buchner's work with yeast extracts over a century ago, scientists have been exploiting enzymes for their natural properties in the lab[6-12]. In the middle of the twentieth century the focus was primarily on optimizing the isolation and purification of enzymes for use in stereoselective transformations[4, 13]. However, due to the frequently high specificity and selectivity of native enzymes, their usefulness as tools for biocatalysis is limited beyond their native substrate scope[14]. In the past two decades, adapting enzymes for biocatalysis has been a challenge that has promoted a great deal of protein engineering work, especially towards enzymes involved in industry or industrially relevant processes. One major reason biocatalysts have become so widely applied is due to their environmentally friendly nature. Enzymes are biodegradable, biocompatible, non-toxic and they avoid the use of precious metals; additionally, they react under mild, aqueous conditions and can provide high selectivities[15]. A review by Bornscheuer et al. classified the progression of protein engineering in the recent past into waves[2]. The first wave has already been discussed as it consisted primarily of isolating enzymes in their native form and utilizing them in accordance with their natural functions. The second wave presented the use of two engineering ideologies, rational design and directed evolution, to effect desirable biocatalytic properties. The third wave added structural data into the engineering strategies of the second wave to create semi-rationally designed libraries.

Enzyme Engineering

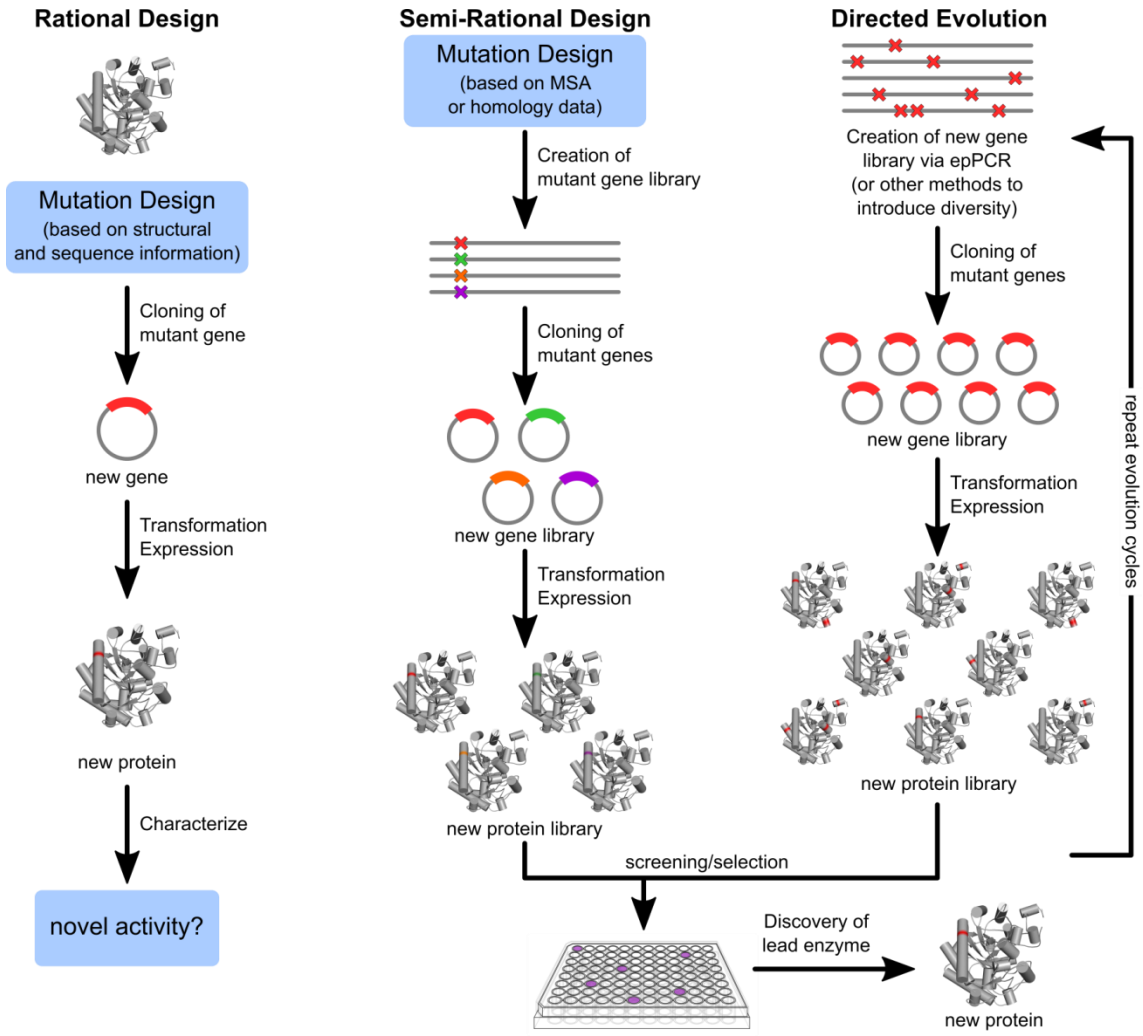


Figure 1.1. Overview of rational design, directed evolution, and semi-rational design. Figure adapted from Porter, J. L.; Rusli, R. A.; Ollis, D. L., *ChemBioChem* 2016, 17 (3), 197-203.

Since the early 1990s, the biocatalysis field has been continuously growing with the advent of new technologies[2]. Significant strides have been made in several areas that have permitted the field to grow at such a rapid pace. Advances in DNA sequencing have substantially expanded the gene pool available and lowered the cost of synthetic genes[16]. Protein immobilization has increased recovery and recycling of biocatalysts[16]. Protein engineering methodologies have significantly improved in the last 25 years, with the emergence of rational design, directed evolution, and semi-rational design techniques (Figure 1.1)[13, 15]. The first successful efforts at engineering enzymes by site-directed mutagenesis in a controlled and reproducible manner were reported in the mid 1980s[10].

These efforts were made possible by advances in molecular biology at the time. The development of recombinant DNA methodology[17], along with the development of PCR[25] proved to be invaluable foundational tools for the nascent protein engineering field.

Rational Design

Rational protein engineering efforts have allowed scientists to answer fundamental questions about the role individual amino acids play in protein structure and function. However, a general condition for rational design efforts is some ab initio structural knowledge of the protein (such as a crystal or cryogenic electron microscopy structure) prior to targeting residues for mutagenesis[18]. The success stories of rational design highlight the potential of this method for protein engineering[19-20]. However, the effects of rational amino acid substitutions are not always easy to predict, and as a result, engineering to improve catalytic or physical properties is a non-trivial task that generally requires knowledge of the sequence and structure, close ancestral and evolutionary relations, and the active site and reaction mechanism[16]. The long and distinguished list of failures of rational design is also a reflection of our current limited understanding of the functional and structural complexity of enzymes[14].

Directed evolution

When nature encounters the same issue of functional complexity within enzymes, it uses Darwinian evolution to overcome it. Laboratory directed evolution mimics natural evolution and involves repeating cycles of sequence diversification, protein expression, and some selection or screening method to isolate enzymes exhibiting the preferred properties[16]. Under selection conditions, some pressure is applied to 'select for' the desired phenotype, allowing the exclusive survival of certain organisms[21]. A classic example of this

is antibiotic-based selections in which exogenous DNA is transformed into a cell that contains both a gene of interest and an additional gene that confers resistance against a specific antibiotic. Cells that successfully take up the DNA are able to grow on agar medium that contains the corresponding antibiotic, while any cells that did not successfully take up the DNA cannot survive with the antibiotic present[21]. Unlike selections, screens allow the degree to which any organism displays a specific phenotype to be observed, but doing so requires an active search through a variant library[21]. Screens that sort based on the fluorescence of the compound of interest are some of the most common examples, since they have both a higher sensitivity and can allow for a higher throughput than many other types of screens[22]. Another example of a screen is determining the enantioselectivity preference of enzymes within a variant library using a chiral GC or HPLC, although this tends to be lower throughput than many fluorescence based screens[23]. Whether to choose a selection or a screen depends upon the library size, which is determined by the method of diversification chosen to create the variants.

The earliest methods of directed evolution in the laboratory relied mostly on random mutagenesis through error prone PCR (epPCR) for diversification of gene sequences[24]. One of the advantages of epPCR is that it requires no structural information and is a 'shotgun method'[25]. However, this method creates very large libraries, and a meaningful assessment of these large libraries demands a high throughput assay[13, 25-27]. Another disadvantage to epPCR is that it does not have any method to share beneficial gene mutations between different variants, since each library member represents an asexual lineage[14]. The Stemmer and Arnold groups independently offered an effective solution to this problem through the introduction of DNA shuffling[28-29]. Recombination of highly homologous DNA fragments allows for crossover of beneficial mutations among individual library members[29].

Directed evolution libraries have the potential to be very large. Libraries with up to 10^{15} sequence variants have been reported, but even libraries that large are not big enough to cover all possible variations if there are more than two or three corresponding amino acid changes on average in a 1000-bp gene[14]. Library size is of critical importance since the ability to assess millions of variants is limited to fluorescence activated cell sorting (FACS), agar plate screens or selections, or microfluidic screening methods. Since screening is currently the bottleneck point for directed evolution strategies[2, 26, 30-41], an appropriate screen (or selection) must be developed for each library. The highest throughput techniques are selection techniques, but when implemented, they only indicate whether or not cells possess a specific property. Although the extent of a selection pressure can be varied, in general, selection techniques do not indicate the degree to which that property has been altered relative to the parent gene. High throughput screening techniques can involve the use of microfluidic devices, which utilize picoliter–volume drops of aqueous reaction mixture dispersed in oil. Such devices have been used to dramatically increase the rate of analysis primarily for fluorescently labeled substrates[42-43] and, more recently, non-labeled substrates using mass spectrometry as a detection method[44-45]. While microfluidic devices have proven to be a useful high throughput screening method, limitations involved with the technology are currently being addressed, including minimizing leakage between droplets and expanding the detection methods available for these devices[46]. Another option is microtiter plate screenings, which are much more laborious than the previous methods mentioned, but yield more useful information. These rely on GC, HPLC, or a colorimetric assay, depending upon the library size. Since they are more laborious, library sizes for microtiter plate screens are usually limited to no larger than 10^4, even with the help of high end automated systems. However, they can screen for a wider variety of substrates and products[16]. Due to the challenges of library size that directed evolution presents, more

recent efforts have focused on 'semi-rational' design strategies that provide small, focused libraries.

Library Selection and Screening

Challenges remain in engineering biocatalysts, especially when large libraries are created, related to screening the libraries that result from engineering efforts. Two fundamental aspects are still critical to any method chosen for library analysis. The first is the rule, '*you get what you screen for*'[47-48], and the second is that a library's genotype must be linked in some fashion with its phenotype. In regards to '*you get what you screen for*,' this refers primarily to the fact that in many cases experimentalists rely too heavily on model conditions, or do not take all of the parameters into account that will eventually factor into their respective assays. These factors can include: pH, temperature, sample preparation conditions or buffer composition. As a result, they can engineer an enzyme that has wonderful activity, but that activity is only observable for an analogous compound or in a non-useful physical environment[48]. A clear example of this phenomenon is the work performed in Steve Withers' laboratory in engineering glycosyltransferases using FACS. with 400-fold activity only displayed that enhanced activity towards the fluorescently labeled lactose rather than the lactose itself[49]. This oversight was later corrected using a method which ensured they were selecting for the native substrate[22].

For the second aspect of library analysis, the genotype of the expressed phenotype must be linked in order to trace back the mutation that caused the beneficial phenotype change. Currently, protein sequencing requires more material and is much more expensive than DNA sequencing; therefore, keeping the genetic elements encoding the phenotypic output spatially linked is significantly faster and more cost-effective for library analyses. The most common way this link is established is through transformation of the DNA library

into an expression host, such as *Escherichia coli*. However, functional assays can be problematic within host organisms for a variety of reasons. These can include difficulty transporting reagents across host cell membranes, background interference of native proteins within the host, and the expense of the throughput of the assays[14]. Another method of linking genotype and phenotype is using some sort of surface display system. These systems can vary in complexity and in the choice of host organisms, ranging from viral capsids[50], to prokaryotes[51], to eukaryotes[52], to bacterial spores[53]. A core feature that all surface display systems have in common is the use of a surface protein that is amenable to recombination which anchors the target protein to the surface of the display system, using an encoded signal peptide to direct the target protein for surface presentation[54]. Surface display systems have some distinct advantages over assays within a host, such being outside the membrane, thereby circumventing some of the issues associated with transporting reagents into host cells[14]. Additionally, some eukaryotic systems are capable of sophisticated post-translational modifications, allowing for more complex enzymes to be displayed[54]. Although display systems are usually very effective for identifying high affinity interactions, they present more of a challenge when isolating enzyme variants that provide multiple turnovers as the products of their reactions diffuse, making it difficult to isolate a single mutant[14, 54]. Another alternative to both these methods was developed by Tawfik and Griffiths which allowed for *in vitro* compartmentalization in water-in-oil emulsions, creating an artificial linkage of genotype and phenotype[55]. This was achieved through the use of a cell-free *in vitro* transcription/translation system.

Semi-rational design

Semi-rational design strategies utilize information from the rapidly growing bioinformatics fields[56]. Specifically, information from multiple sequence alignments of

proteins and structure databases[57], as well as computational and machine-learning algorithms, have all been used to help design smaller libraries of focused protein variants[14]. One example of semi-rational design is Reetz and coworkers development of a process called combinatorial active-site saturation testing (CASTing), or targeting only residues in or near the active site in combination for mutagenesis in order to limit library size[58-59]. They chose to focus on such residues based on recent work indicating substitutions near the active site were more effective in altering catalytic activity than distant substitutions[60]. The idea behind this methodology is to limit library size by applying saturation mutagenesis to two residues that are spatially close to one another near the active site, increasing the chance for synergistic effects to occur when variants are combined[59]. These positions are then simultaneously randomized, creating a library of double and single variants which, due to its smaller size, is easier to screen than libraries that contain variants with three or four mutations[59]. CASTing can be applied at several different positions in the active site, creating multiple small libraries to screen (around 10^3 different variants), rather than creating one very large library with a wide variety of enzyme variants (over 10^6 different variants)[59]. The Reetz group then followed this work up with demonstrations of iterative saturation mutagenesis (ISM) as a method to help with 'accelerated directed evolution' efforts such as CASTing[61]. At its core, ISM involves the creation of a series of saturation mutagenesis libraries with mutations located at predicted 'hot sites' of the enzyme, based on structural information[61]. These initial libraries are screened, then the best variants are used as the template for the next round of saturation mutagenesis at different residues[61]. ISM allows combinations of beneficial mutations to arise quickly and in a focused manner, which are unlikely to appear through epPCR or DNA shuffling[61]. Although the library sizes for CASTing and ISM are significantly smaller than those from standard directed evolution studies, they still require screening thousands of variants. Additionally, they assume that

the most beneficial mutations will arise in or around the active site, which limits identification of beneficial mutations at distant sites that may have been discovered through more traditional methods.

Other semi-rational design strategies attempt to take advantage of the idea that new enzymatic functions can be developed through a promiscuous enzyme intermediate, similar to an ancestral enzyme[62]. Work performed in the Tawfik lab supports this strategy: early efforts highlighted the possibilities for creating promiscuous enzymes, and more recent work has taken advantage of the ideas of using promiscuous, or ancestral, enzymes as a starting point for evolution to minimize the amount of directed evolution necessary[63-65]. The Tawfik lab was able to evolve three different enzymes, a phosphotriesterase, serum paraoxonase, and carbonic anhydrase II, to have promiscuous activity while still retaining their native functions[64]. More recent work using different model enzymes (serum paroxanose and cytosolic sulfotransferase) demonstrated that by preparing an ancestral library, these libraries could be subjected to directed evolution, allowing for a relatively small library that was comprised of more evolvable and active variants, on average, than traditional random mutagenesis methods provide[65]. By starting with a promiscuous enzyme, the necessary step of evolving away from the enzyme's native activity was already completed, and therefore fewer steps were necessary to obtain a useful enzyme for the desired catalysis. Using an ancestral, or promiscuous, enzyme as an intermediate or starting point can be a highly useful way to minimize library size and alter an enzymes catalytic activity. However, if an enzyme already has some activity towards a desired substrate, this type of engineering work will be of less use since the enzyme does not need to be evolved away from this native activity. Additionally, creating the ancestral enzymes requires significant reconstruction through phylogenetic analyses[66].

Another emerging semi-rational design based protein engineering strategy is to use computational or machine learning based techniques. These approaches include ProteinGPS, ProSAR, and Rosetta. ProteinGPS is a technology created by Atum (formerly DNA 2.0) that identifies beneficial substitutions based on machine learning algorithms[67-69]. ProSAR is a similar technology to ProteinGPS, as it also involves incorporating machine learning algorithms towards directed evolution[70]. Specifically, these techniques combine activity and sequence information with statistical analysis to determine the effect that each individual substitution has on a protein's activity for every round of evolution. These methods allow for improvements through combinations of amino acid changes that are not substantially beneficial individually in the corresponding variants[71]. Both of these technologies need a sequence-activity dataset in order to apply the machine learning algorithms to continue forward with directed evolution strategies, which generally requires a large initial screen, especially when applying ProSAR[71]. Finally, as computational methods develop further, the impact of amino acid substitutions on protein structure and function can also be presampled *in silico* using programs such as Rosetta. Work performed by Baker, Hilvert and coworkers has demonstrated the current computational power in enzyme engineering as they have created catalysts entirely *in silico* with some success[72-77]. A classic example of this work is the *de novo* computational design and creation of an enzyme for a stereoselective Diels-Alder reaction using Rosetta[75]. They succeeded in creating an active, stereoselective enzyme, however only 2 out of 50 designed enzymes that were tested had measurable activities, demonstrating the current limitations of this technology[75]. One way to improve these computationally designed enzymes is through directed evolution after *in silico* design, as work done with Kemp elimination catalysts demonstrates[76]. The work performed by Röthlisberger, et al. showed that by screening a relatively small number of variants (800-1600 per round), they were able to produce an enzyme with a >200-fold

increase in k_{cat}/K_M[76]. By combining computational design with focused directed evolution libraries, they were able to create a new Kemp elimination biocatalyst. One drawback worth noting is that the computational power required to design these enzymes is significant and may not be readily available to all reasearchers[78]. Additionally, designing proteins *de novo* currently faces challenges with robustness, aggregation, and lack of functionalities (such as allostery and signaling) present in most natural proteins[78]. While these methods are very attractive, especially for the creation of enzymes to perform non-natural activity, computational methods are more useful currently as a way to assist semi-rationally designed evolution methods rather than as stand-alone creators of biocatalysts.

Cell-free translation systems

Cell-free translation systems have been around since the mid 1970s[79]. For protein engineering, the major advantages of these systems include avoiding the need for transformation into a host, as well as eliminating the time typically required for expression and purification of proteins. There are two approaches to these systems: the first is based on crude cell extracts, typically from *Escherichia coli*, wheat germ, or rabbit reticulocytes[80]; the second approach is to purify and reconstitute the components of transcription and translation[81]. While crude cell extracts have been optimized for high yield, they have limited applications due to complications that typically arise from materials contained within the lysate. One of these issues is product degradation from proteases or nucleases inherent within the cell[82-83]. Another issue is that lysates contain all of the pathways found natively in cell metabolism, which can be beneficial if the activity from those pathways is desired[84]. However, if these activities are not desired, the background metabolism can interfere with characterization of the desired activity of the evolved enzyme[85].

The drawbacks of the cell lysate based systems led the Ueda group to develop the Protein synthesis Using Recombinant Elements (PURE) system in 2001[81]. The PURE system uses purified enzymes for protein synthesis as well as the 70S ribosome[86]. As a result of the clearly defined nature of the system, experimenters have more control over the compounds that are present within the reaction mixture, thus minimizing any background interference. Additionally, due to the purified nature of all of the enzymes and the ribosome in the PURE system, there is minimal protease and nuclease activity[87]. The PURE system has been used to investigate the individual components of protein transcription/translation[88], including how the ribosome interacts with specific molecules such as macrolide antibiotics[89]. The PURE system has also been applied to various protein engineering efforts, such as the incorporation of non-canonical amino acids in a site specific manner[90-91]. Other protein engineering efforts that use the PURE system have included directed evolution studies[92-93], as well as previous work in our lab to assess a library of Old Yellow Enzyme 1 variants[94]. The most attractive feature in regard to protein engineering efforts is that the protein produced by the PURE system can be used directly for functional analysis, without further purification. This can significantly reduce the time needed for screening efforts.

While the PURE system has its clear advantages, it is not without its drawbacks. In a direct comparison, *E. coli* cell lysate formed about 1.5 times more functional protein compared to the PURE system, using firefly luciferase as a reporter[87]. The lower yields of the PURE system are most likely due to either problems with efficient energy recycling, or issues with the translation efficiency of the system[87]. The problems arising from translation efficiency have been addressed and can be assisted with the addition of chaperones such as GroEL/GroES[88, 95-96]. One of the biggest drawbacks currently of the PURE system is the cost of the system, which limits its utility to most researchers[97]. An early criticism of the PURE system was that the lack of any membrane elements made membrane proteins difficult to

synthesize within the system: however, a new method with the PURE system has recently been devised to allow for the synthesis of membrane proteins[98]. Researchers are now heading in the direction of a minimal, artificial cell within synthetic membranes, and the PURE system has been proposed as the core technology in these minimal cells[99]. In spite of its costs and limitations, the applicability of the PURE system to a wide variety of enzyme classes makes it a useful tool.

Flavoenzymes

The class of enzymes that this work will focus on is flavoenzymes, which are enzymes that require a flavin cofactor and catalyze a wide variety of reactions[100]. Due to the wide range of chemistry that can be performed and the selectivity of these enzymes, they are industrially attractive for a variety of purposes[100-104]. Flavoenzymes have important roles in biosynthetic pathways for a variety of natural products including antibiotics and antitumor agents[105-107]. They are also used in the food[108] and fine chemical[100, 109] industries as well as being useful in a variety of synthetic routes due to their typically high enantio- and regio- selectivity[106]. Generally, flavoproteins contain either flavin mononucleotide (FMN) or flavin adenine dinucleotide (FAD), shown in Figure 1.2, bound noncovalently to the enzyme[110]. As a result of the wide variety of reactions flavins can catalyze, flavoenzymes are found in a wide range of biological processes, from the classical oxidation/reduction reactions they are typically associated with to energy production to DNA repair to chromatin remodeling[100].

The reactive part of the flavin is the isoalloxazine ring system (Figure 1.2A). The redox potential of flavin is around -200 mV but varies from -400 mV to +60 mV depending upon the local enzymatic environment in which the flavin is contained[111-112]. Flavins can exist in oxidized, semiquinone, or reduced form as shown in Figure 1.2B. The 'classic' redox

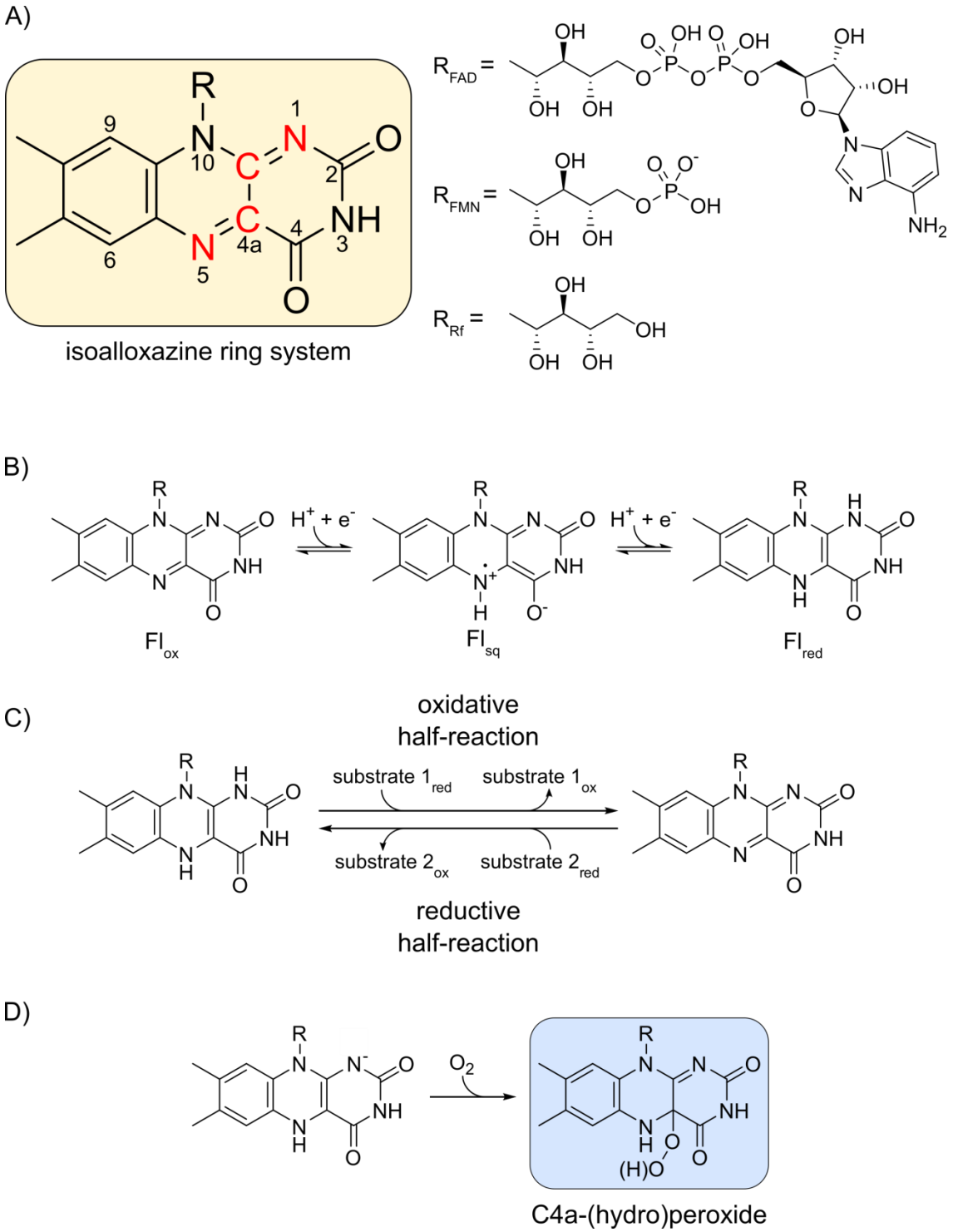


Figure 1.2. The structures and classical reactions of flavin cofactors. A) The isoalloxazine ring structure of flavins with atomic numbering and the most relevant reactive atoms are highlighted (red). The chemical structures for the FAD, FMN and riboflavin tails are shown as well. B) The classic redox reactions of flavins are shown as well as the oxidized, semi-quinine, and reduced states accessed via single electron transfer. C) The hydride transfer reaction typically divided into the oxidative and reductive half reaction. D) The creation of the C4a-(hydro)peroxide flavin via reaction with molecular oxygen. Figure adapted from Piano, V.; Palfey, B. A.; Mattevi, A., *Trends Biochem. Sci.* 2017, *42* (6), 457-469.

chemistry of flavins hinges upon their ability to perform either a single or double electron transfer[113]. Flavoenzymes typically have two half-reactions: an oxidative half-reaction and a reductive half-reaction (Figure 1.2C). In the oxidative half reaction, the flavin is oxidized by an electron acceptor, while in the reductive half reaction, the flavin is reduced by the substrate with the hydride transfer originating from N5[106, 111-112]. Within the classical flavin reactions, the versatility arises from the ability of the flavin to react differently with molecular oxygen depending upon its protein environment[114]. Oxidases generally use molecular oxygen as an electron acceptor to produce hydrogen peroxide[114]. Monooxygenases are oxidized by molecular oxygen to form a C4a-hydroperoxide intermediate (Figure 1.2D) which can then perform a hydroxylation or insert an oxygen atom in a substrate[106]. Dehydrogenases and electron transferases, however, typically do not react or react very slowly with oxygen and instead use other electron acceptors[114]. These redox reactions with molecular oxygen represent the classical chemistry for which flavoenzymes have been known.

Monooxygenases are a common and well-studied part of the flavoenzyme family, and until recently, they have only been known to use the C4a-hydroperoxide for oxygen transfer[115-116]. In recent years, more unprecedented flavin mechanisms are emerging, widening the amount of known chemistry that flavoenzymes can perform[115]. Generally, these new types of flavoenzyme mechanisms involve an N5-covalent adduct formation. One example of this is the formation of an N5-oxygen adduct[115]. One of these novel enzymes, flavoenzyme EncM, has been found within the biosynthetic pathway of enterocin and was isolated with its FAD in the N5-oxide form (Figure 1.3)[105, 117]. This suggests that the reduced flavin reacts with oxygen to form a transient N5 peroxide, leading to the N5-oxide which can perform the oxidation of the substrate[117]. Interestingly, EncM does not need any external

reducing agent to complete its catalytic cycle, allowing it to behave as an 'internal monooxygenase' which is unlike most other flavin oxygenases[115].

Another novel mechanism involves the formation of an imine adduct on the N5 of the flavin. A classic example of this is UDP galactopyranose mutase (UGM), which catalyses a nonredox reaction with the N5 atom of the reduced flavin reacting with the substrate, leading to an iminium adduct[118]. This covalent adduct was recently crystallized, conclusively demonstrating the formation of the iminium intermediate (Figure 1.3)[119]. A different version of these imine-N5 adducts is represented by alkyl-dihydroxyacetone phosphate synthase (ADPS). Although ADPS also forms a flavin iminium adduct in a nonredox reaction, the intermediate in this case requires an oxidized flavin instead of a reduced flavin[120]. With both UGM and ADPS, a covalent trap is created by the flavin that protects a reactive intermediate throughout the reaction[115]. Another flavoenzyme proposed to use a flavin-iminium intermediate is flavin-dependent thymidylate synthase (FDTS). Specifically, FDTS has been proposed to perform a methylene transfer using a flavin-iminium intermediate[121]. The reduced flavin forms the N5-methylene intermediate and the FAD acts as the methylene carrier between the donor and acceptor molecules[115, 122].

Figure 1.3. Structures of unusual N5-covalent adducts on flavins.

There has also been the recent discovery of a flavin analog with an isoalloxazine ring modification at the N5 position, allowing for novel reactions. The modified flavin is a prenylated flavin and is the result of two enzymes: UbiX, a prenyltransferase, and ferulic acid decarboxylase (Fdc1). UbiX derivatizes the FMN as shown in Figure 1.3[123]. The prenyl-FMN is then oxidized by Fdc1[124]. This allows the iminium form of the prenyl-FMN to react

with the substrate (cinnamic acid) in a 1,3-dipolar cycloaddition, leading to the decarboxylation of the substrate and formation of a styrene adduct on the prenyl-FMN, which is released upon completion of the catalytic cycle[125]. This N5-alkylated derivative is a robust prosthetic group in contrast to the iminium-flavin systems described earlier which are relatively unstable, transiently formed reactive intermediates[126]. This novel prenylated flavin demonstrates the wide versatility of flavin catalysts, some of which have yet to be discovered.

While there are new, intriguing flavin mechanisms emerging, there is still a wide variety of interesting chemistry that can be performed with 'classical' flavin chemistry. The rest of this work will focus on an ene-reductase that performs a hydride transfer and on a traditional monooxygenase that uses the C4a-peroxyflavin. The specific flavoenzymes discussed in more detail are Old Yellow Enzyme 1 and Baeyer-Villiger Monooxygenases.

Old Yellow Enzyme 1

Old Yellow Enzyme 1 (OYE1) was first isolated in 1932 from *Sacchromyces pastorianus* and described by its distinct yellow color, named 'yellow ferment', and noted to be involved in the oxidation of NADPH by molecular oxygen[127]. A few years later a 'new' yellow enzyme was discovered[128], which led to the original enzyme being renamed the 'old yellow enzyme', which has persisted to this day for the entire family of OYE-like enzymes. Since the discovery of the family, OYE-like enzymes have been found in bacteria, plants, other yeasts, and eukaryotes[129]. New OYE homologs have been discovered in cyanobacteria recently by screening environments for ene-reductase activity[130]. Additionally, with the ability to genome mine databases now, previously unidentified OYE homologs are being identified, such as one from *Lactobacillus casei*[131]. An in-depth identification of OYEs from fungal species was recently performed to gain more insight about fungal OYEs[132]. A few of

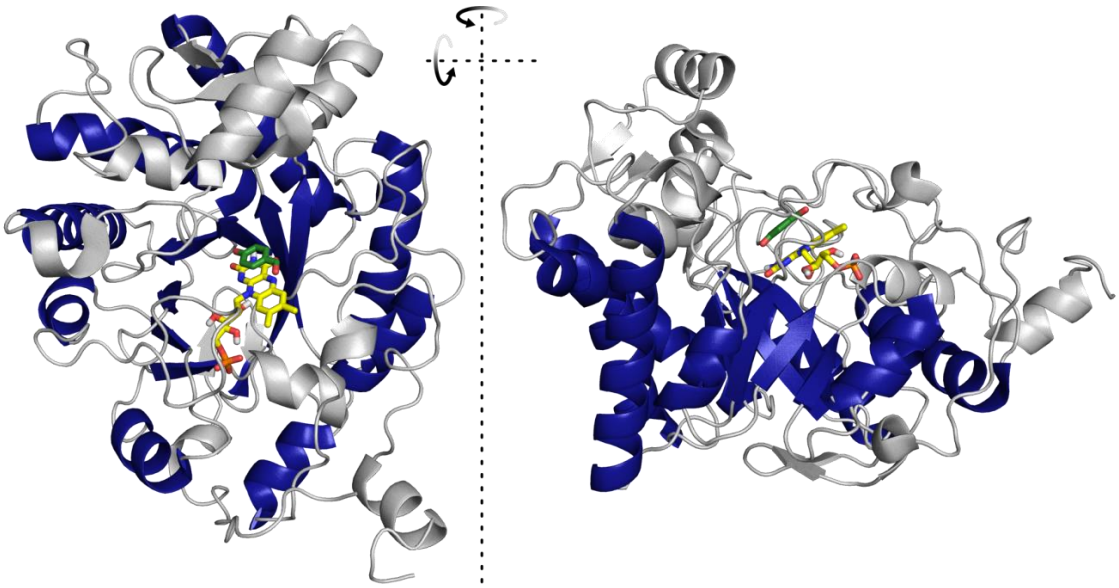


Figure 1.4 Structure of OYE1 with two different views (PDBID: 1OYB[139]). The FMN (yellow) and *p*-hydroxybenzaldehyde (green) are shown bound in the active site. The $(\beta/\alpha)_8$ barrel fold is highlighted in blue.

the OYE homologs that have been well characterized and will be discussed further are pentaerythritol tetranitrate (PETN) reductase, from *Enterobacter cloacae* PB2[133], 12-oxophytodienoate reductase 1 (OPR1) and 12-oxophytodienoate reductase 3 (OPR3) from tomato[134], NCR from *Z. mobilis*[135], and YqjM from *Bacillus subtilis*[136]. YqjM, when crystallized, revealed a second class of OYEs based upon several structural features[136]. As a result, classification of OYE homologs is now separated into OYE-like and YqjM-like proteins[137]. These YqjM-like proteins also contain the thermophilic OYEs[138].

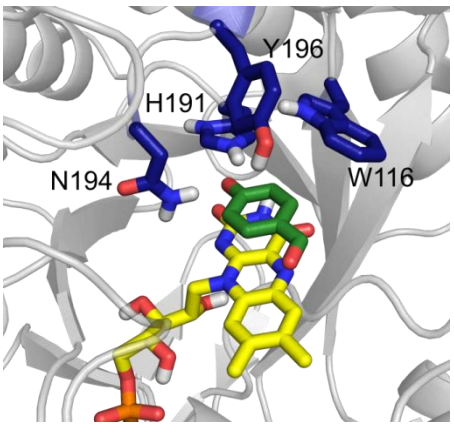


Figure 1.5. Close up of OYE1 active site with relevant residues labeled. (PDBID: 1OYB[139])

The OYE family shares a few structural features. Among these is the single domain $(\beta/\alpha)_8$ barrel fold as the core structure, visible in Figure 1.4[112]. Another key component found in the OYE family is the presence of a noncovalently bound FMN molecule in the C-terminal end of the $(\beta/\alpha)_8$ barrel[139]. As shown in Figure 1.5, the substrate binds over the

si-face of the isoalloxazine ring[136, 139-140]. Finally, there are several conserved residues within the family. Key among these residues is Y196 (Figure 1.5), which is believed to be a proton donor during catalysis[141], but some studies indicate that it is not strictly essential for catalysis[138, 142-143]. Two other conserved residues, H191 and N194 (Figure 1.5), are believed to play roles in substrate binding and orientation for catalysis[139, 144].

The OYE family is composed of ene-reductases, which catalyze the hydride transfer reaction shown in Figure 1.6. Substrates of OYE1 are typically α,β-unsaturated aldehydes, ketones, imides, nitroalkenes, carboxylic acids and esters. OYE1 has also been shown to be active towards compounds such as nitroaromatics and both cyclic and acyclic enones[138]. The reaction for OYE1 proceeds via a hydride transfer from the N5 of the reduced flavin onto the beta carbon (Figure 1.6) of the substrate with a proton transfer in a *trans*-fashion to the alpha carbon.[141, 145-146] This proton transfer generally leads to high stereoselectivities that are conserved within the entire OYE family[138].

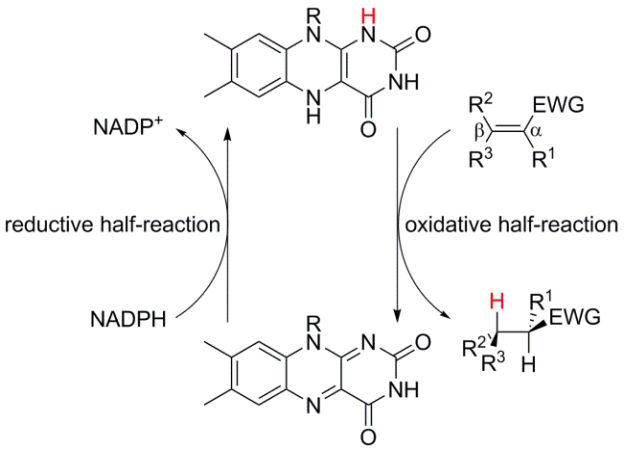


Figure 1.6. Catalytic cycle for OYEs.

This conservation of stereoselectivity can be observed across a variety of different OYE family members, with a wide range of substrates. PETN reductase is one of the members of the OYE family that is able to reduce a variety of explosives, such as nitrate esters or nitroaromatics[133], and has generally high stereoselectivity, depending upon the reaction conditions it is subject to[147]. OPR1 and OPR3 display a range of stereoselectivities, depending upon the substrates, generally reacting with similar stereoselectivities[134]. However, they did demonstrate stereocomplementarity towards a nitroalkene, 2-phenyl-1-

nitropropene[134]. When OPR1 and OPR3 were compared to YqjM as well, all three enzymes demonstrated excellent stereoselectivity toward the reduction of nitroolefins, enals, malemides, and enones[148]. Although with the exception of the aforementioned nitroalkene, the stereoselectivity did not differ between the enzymes[148]. This is unsurprising, since most of the ene-reductases have the same stereoselectivities, with a few noted exceptions[149]. Part of the reason for this retention of stereoselectivity between homologs is the high conservation of the active site residues discussed previously[149].

This retention of stereochemistry has led to a wide variety of protein engineering work on the OYE family[138, 149-154]. In attempts to improve the thermal stability of NCR, the Hauer group successfully applied rational mutagenesis to shorten loops within the enzyme to make it more similar to the more thermostable YqjM-like OYEs[155]. They then built upon this initial thermostability work by grafting in loops from different OYE homologs to alter the stereospecificity of the chimeric enzymes towards certain substrates[156]. PETN reductase has also been the focus of significant protein engineering efforts, especially by the Scrutton group. They have performed directed evolution to increase the catalytic performance of PETN reductase towards a variety of substrates[151]. They engineered two variants (T26S and W102F) that switched the enantiopreference of PETN reductase as well as developing several variants with significantly faster rates[151]. The Reetz group performed ISM on YqjM, which lead to the discovery of a variant with reversed enantioselectivity[153]. Using ISM in the form of CASTing, they were able to evolve both (*R*)- and (*S*)- selective variants of YqjM towards 3-methylcyclohexenone[153].

While a great deal of research has been performed on the entire OYE family, specifically the research performed in this thesis will focus on OYE1 since it builds upon previous work with OYE1 in our lab which provided the basis for further engineering. Significant pioneering work was performed on OYE1 by the Massey group, including (but

not limited to) the elucidation of the mechanism and the role of key residues within the active site through protein engineering[141, 144, 157-161]. As a result much of the knowledge of OYE1 comes from the studies carried out in the Massey lab. Protein engineering studies on OYE1 have primarily focused on expanding its substrate range or changing its substrate specificity since OYE1 has a modest active site, restricting it to smaller substrates[149, 162]. Efforts to change the stereoselectivity have primarily come from the Stewart lab and focused on the W116 position (Figure 1.5). Substitutions at W116 were observed to be critical towards changing the stereoselectivity in favor of the reduction of (*S*)- or (*R*)-carvone. In particular, the alanine and valine variants at W116 showed a reversal in stereoselectivity with both substrates, while other substitutions such as tyrosine or tryptophan showed no change in stereoselectivity[163-165]. Previous work aimed at changing the substrate specificity has involved circular permutation studies of OYE1 from our lab[166]. These studies have improved the catalytic efficiency of OYE1 towards three different substrates, by altering the structural dynamics and changing the active site volume within the enzyme[167]. These two studies are the primary engineering efforts on OYE1, prior to the work discussed largely in the third chapter of this work. The third chapter of my dissertation will focus on further engineering OYE1 by evaluating the impact of additional structural changes on stereoselectivity and substrate specificity using a cell-free transcription translation system.

Baeyer-Villiger Monooxygenases

Adolf von Baeyer and Victor Villiger first reported the transformation of ketones to esters or cyclic ketones to lactones using Caro's acid in 1899[168]. Since its discovery, the eponymous Baeyer-Villiger reaction has become a standard in the field of organic chemistry and has continued to develop over the last century; it has a wide variety of applications in

areas including the synthesis of steroids, antibiotics and monomers for polymerization[169]. A Baeyer-Villiger reaction proceeds through the nucleophilic attack or addition of the peroxy acid to the carbonyl to form the Criegee intermediate[170] (indicated in Figure 1.7). Next, a migration of one of the adjacent carbons to oxygen leads to the release of the carboxylate anion, or carboxylic acid. This migration can result in one of two regioisomers, depending upon the migratory aptitude of the R groups on the ketone. The migratory aptitude of the substituents is predictable and follows the following order: tertiary alkyl > cyclohexyl > secondary alkyl > benzyl > phenyl > primary alkyl > CH_3[171]. The primary stereoelectronic effect that controls the migration ability of these groups is that the group that migrates does so from an anti-periplanar position relative to the oxygen-oxygen bond of the peroxide[172-175]. The classical Baeyer-Villiger oxidation is limited in its enantio-, chemo- and regioselectivity to the chemically preferred products due to the rules that dictate migratory ability of the substituents. Also, oxidants such as trifluoroperacetic acid or 90% hydrogen peroxide required for synthesis are either unstable, costly, shock sensitive or explosive.[176]

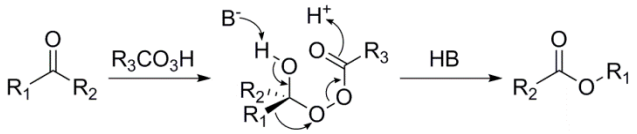


Figure 1.7. General mechanism for the Baeyer-Villiger reaction, featuring the Criegee intermediate as the second step.

Baeyer-Villiger monooxygenases (BVMOs) offer an enzymatic alternative to the traditional Baeyer-Villiger reaction. The first indication of a biological Baeyer-Villiger reaction was demonstrated in 1948 by Turfitt in a series of microbiological investigations of steroids[177]. However, Trudgill and coworkers were the first to purify a Baeyer-Villiger monooxygenase in 1976 and demonstrate it to be a flavoprotein with broad substrate specificity[178-179]. Since the discovery of the family, several more BVMOs have been discovered and purified, leading to the discovery that many BVMOs are able to oxygenate heteroatoms such as sulfur, boron, nitrogen, or selenium in addition to carbon[180-181].

Additionally, a classification system has been developed that divides BVMOs into type 1, type 2, or 'atypical' (type O) BVMOs. Type 1 BVMOs, considered the 'prototype BVMOs', have been the focus of the majority of BVMO research and consist of only one polypeptide chain[176]. They contain FAD tightly bound in the active site and are dependent solely on NADPH for reduction of the flavin[176]. The key features of type 1 BVMOs are the characteristic BVMO fingerprint motif, FXGXXXHTXXW(P/D), and two dinucleotide-binding domains, or Rossmann folds, which bind cofactors such as FAD or $NADP^+$[176, 182-183]. Type 2 BVMOs consist of two different subunits, an oxygenase, which has an (β/α) barrel structure, and a reductase, and they lack a BVMO fingerprint motif[176, 184]. Additionally, type 2 BVMOs differ from type 1 in that they use FMN rather than FAD and are able to use NADH for reduction[176]. The last class is the atypical, or type O, BVMOs which do not fit into either type 1 or type 2 classifications, although they are a single polypeptide and FAD dependent[176]. The classic representation for this class is MtmOIV, a key enzyme in the biosynthetic pathway toward mithramycin from *Streptomyces argillaceus*, which has a low sequence similarity to other known BVMOs[185]. When the crystal structure for MtmOIV was published in 2009[186], the FAD was found to be bound near the surface of the enzyme, and the structure was more similar to an FAD-dependent hydroxylase[106]. The main focus of this work will be on type 1 BVMOs.

The mechanism for several BVMOs has been previously studied and is shown in Figure1.8[187-188]. Of note in this catalytic cycle is the formation of the C4a-peroxyflavin, which is the reactive species towards the substrate to form the Criegee intermediate. Additionally, NADPH reacts with the FAD to form $NADP^+$, then remains bound in the active site throughout the entire catalytic cycle. The movement of the $NADP^+$ has been captured further in several crystal structures of cyclohexanone monooxygenase from *Rhodococcus* sp. HI-31 ($CHMO_{Rhodo}$) obtained by the Berghuis lab showing how the enzyme domains move

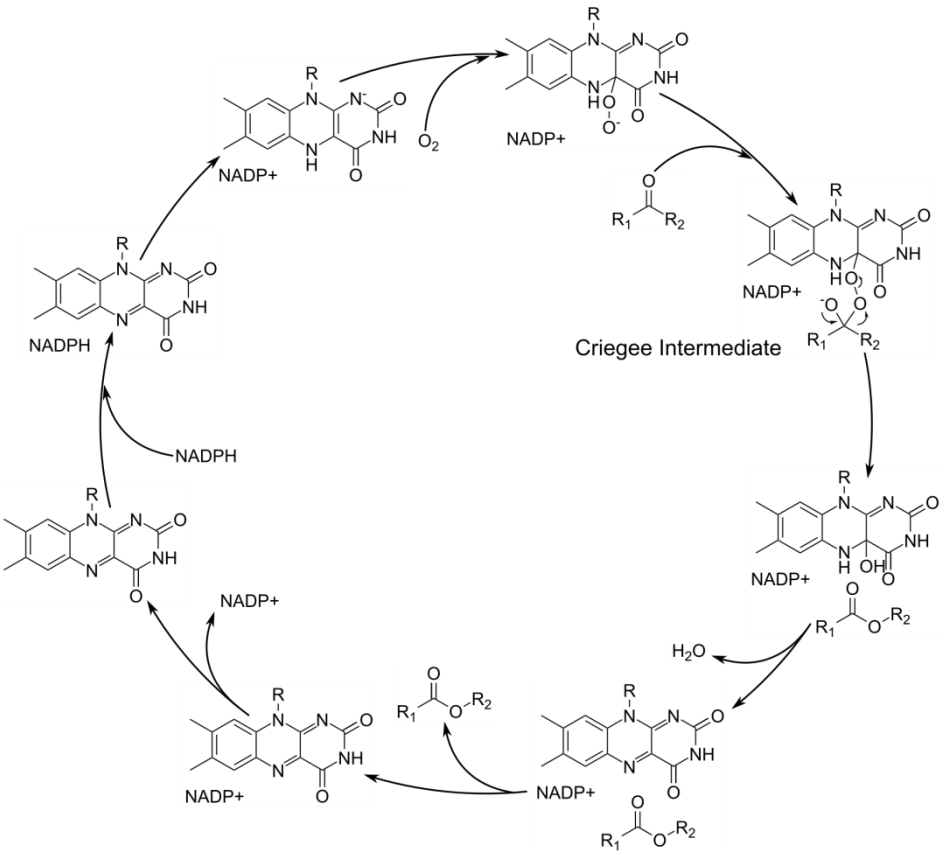


Figure 1.8 General catalytic cycle for CHMO The Criegee intermediate formed is indicated in the cycle. Figure adapted from Sheng, D.; Ballou, D. P.; Massey, V., *Biochemistry* 2001, 40 (37), 11156-11167.

during the catalytic cycle, shown in Figure 1.9A and B[189-191]. These crystal structures demonstrate the movement of the NADPH domain in relation to the rest of the protein as the catalytic cycle progresses(Figure 1.9A). They hypothesized that the movement through several states (open, closed, rotated, tight, and loose) helps position the NADPH and substrate into catalytically relevant conformations[190-191]. These structures also clearly show the movement of the nicotinamide cofactor within the active site (Figure 1.9B) [189-191].

There have been many previous engineering efforts focused on the BVMO family. CHMO from *Acinetobacter* sp. NCIMB 9871 ($CHMO_{Acinet}$) is one of the most frequently used BVMOs[192]. Like many BVMOs, it can demonstrate high enantioselectivity, especially towards some 4-substitued cyclohexanone derivatives (i.e. 4-methyl or 4-chloro), but that enantioselectivity can easily be lost as minor modifications are made at the 4 position[193].

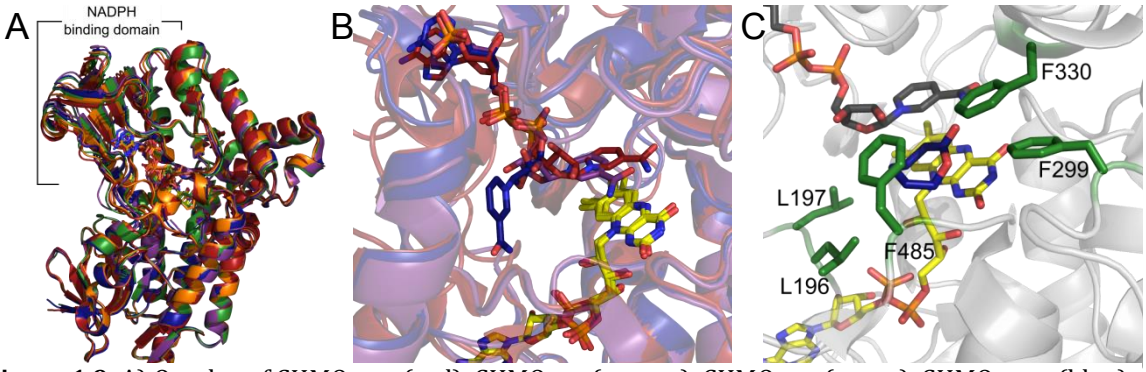


Figure 1.9. A) Overlay of $CHMO_{Closed}$ (red), $CHMO_{Tight}$ (orange), $CHMO_{Loose}$ (green), $CHMO_{Rotated}$ (blue), and $CHMO_{Open}$ (purple) at their FAD binding domains with the movement of the NADPH binding domain clearly visible. B) A close up view of the overlaid active sites of $CHMO_{Closed}$ (red), $CHMO_{Rotated}$ (blue), and $CHMO_{Open}$ (purple) showing the movement of the nicotinamide cofactor. C) CHMO with NADPH (gray) and ε-caprolactone (blue) bound. The residues identified by the Bornscheuer group are highlighted in green. PDBID: 3GWD, 3GWF, 3UCL, 4RG3, 4RG4.[189-191]

Both directed evolution efforts and rational design efforts have been implemented to increase the substrate specificity and change the regio- or enantioselectivity of $CHMO_{Acinet}$[194-196]. Additionally, an attempt was made to increase the robustness and thermal stability of $CHMO_{Acinet}$ via rational design by replacing methionine and cysteine residues with amino acids such as valine or alanine[197]. One of the most impressive engineering works on $CHMO_{Acinet}$ was demonstrated by Codexis in order to produce esomeprazole: their enzyme variant includes 41 amino acid substitutions and demonstrates 99% enantioselectivity[198-199]. Some of the substitutions discussed in the patent are located in loops near the active site that have been demonstrated to be impactful in regard to either substrate selectivity or regioselectivity in homologs, while others are located relatively far from the active site and likely play a role in stability[198].

Another BVMO that has been frequently used is phenylacetone monooxygenase (PAMO) from *Thermobifida fusca*. PAMO is of particular note because it was the first thermostable BVMO discovered[200], and its crystal structure, published in 2004, was the first of any BVMO to be solved[201]. PAMO does have a fairly large limitation though in that it has a relatively narrow substrate scope, promoting the need for engineering PAMO to make it a more useful biocatalytic enzyme[192]. Rational design, CASTing, DNA shuffling, and directed

evolution have all been applied towards PAMO with varying levels of success in altering the substrate specificity, stereoselectivity, and activity[202-209]. Initial attempts to redesign the active site of PAMO focused on the 441-444 loop region, which is next to the active site (Figure 1.10)[202]. The earliest studies used rational design to expand the substrate scope of PAMO to allow it to accept slightly larger substrates by altering this loop region through deletions[202]. Further studies used CASTing to provide amino acid substitutions in this loop region instead of deletions, which was also successful in increasing the substrate scope of PAMO[203]. Additional work used similarities between cyclopentanone monooxygenase (CPMO) and PAMO to discover new mutations around the active site to help expand the substrate specificity, and alter the enantioselectivity, of PAMO[205]. Finally, in an effort to significantly expand the substrate range of PAMO, chimeric enzymes were created, using PAMO as the scaffold, and introducing the C-terminal domain of other BVMOs to both retain the thermostability of PAMO, while exchanging a significant portion of the substrate binding site, including the 441-444 loop region[206]. These efforts were surprisingly successful, producing stable, active enzymes with novel, enantioselective behavior[206]. While the work on PAMO has been extensive and useful in regards to substrate specificity, less research has been focused on altering the regioselectivity of PAMO.

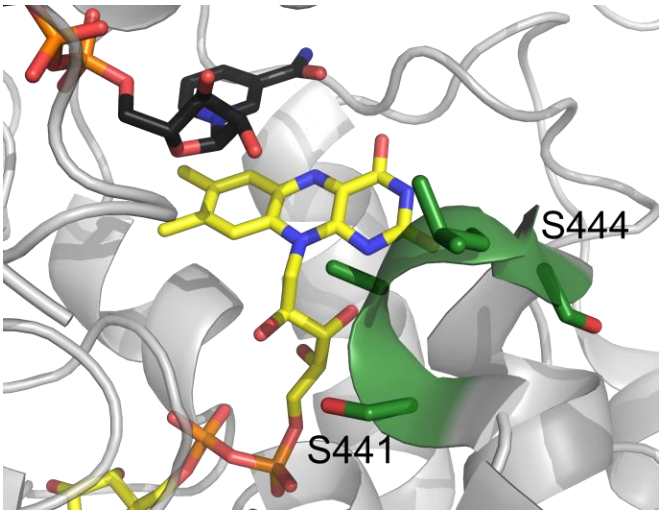


Figure 1.10. Close up of the PAMO active site with NADP (black) bound. The 441-444 loop is highlighted in green. PDBID: 2YLR.[188]

Another BVMO of note that has been engineered for regioselectivity is CHMO from *Arthrobacter* sp. BP2 ($CHMO_{Arthro}$). The Bornscheuer group was able to demonstrate a

complete switch in the regioselectivity of $CHMO_{Arthro}$ by targeting several residues in or near the active site[210]. They found three regions to be of interest for switching the regioselectivity of $CHMO_{Arthro}$ towards (+)-*trans*-dihydrocarvone, shown in Figure1.9C[210]. By mutating the phenylalanine residues shown to alanine or valine and the leucine residues to phenylalanine, they were able to relax the environment of the active site to produce the chemically preferred lactone instead of the 'abnormal' lactone formed natively by the enzyme[210]. In addition to the efforts detailed above, there has been a wide variety of other work done with BVMOs to change enantioselectivity, regioselectivity, and activity[211-216]. Two new thermostable BVMOs were also recently reported: TmCHMO from *Thermocrispum municipale*[217] and polycyclic ketone monooxygenase from *Thermothelomyces thermophile* (PockeMO)[218]. These two enzymes show promise for future engineering efforts since the wild-type enzymes display enhanced degrees of thermostability. Additionally, PockeMO demonstrates activity towards large, bulky substrates, and has a high sequence similarity to cyclododecanone monooxygenase (CDMO) from *Rhodococcus ruber* SC1[218].

The BVMO of interest for this work is CDMO. CDMO is a Type 1 BVMO, contains FAD as a cofactor and demonstrates broad substrate specificity, especially towards long-chain cyclic ketones[219]. CDMO was first cloned and characterized in 2001, and has not yet been successfully crystallized, despite attempts[219]. It is most closely related (by sequence) to PockeMO and cyclopentadecanone monooxygenase (CPDMO)[218, 220]. This sequence similarity is most likely explained by the activity all three of these enzymes have towards larger substrates, specifically compared to CHMOs. Work performed by the Mihovilovic group demonstrated the broad substrate profile of CDMO by ranking CDMO as the best scoring BVMO in three out of eight substrate classes, when compared with eight other BVMOs[221]. Their ranking was based both on stereoselectivity and activity, with more weight given to stereoselectivity where applicable[222]. Given the broad substrate specificity, CDMO

was the best enzyme for the applications needed in our lab. In particular, CDMO showed better activity towards N-protected β-amino ketones relative to other BVMOs. The fourth chapter of this work will focus on the directed evolution of CDMO for N-protected β-amino ketones, which will be further discussed within the context of the chapter.

Aim and scope of the dissertation

This dissertation focuses on engineering flavoenzymes to enhance their activity and alter their stereospecificity or regioselectivity. The initial work focuses on setting up the appropriate tools to perform these engineering efforts, followed by the development of variant enzymes with altered activity.

In chapter two, a system known as RApid Parallel Protein EvaluatoR (RAPPER) was developed. This system utilizes the PURE system to rapidly assess protein variants in a quantitative manner. The development of this protocol also reports preliminary OYE1 variants with enhanced activity due to increased flexibility in specific loop regions.

In chapter three, I report the use and expansion of RAPPER to assess a multitude of OYE1 variants. The utility of RAPPER was demonstrated with amino acid substitutions, deletions, and alanine scanning. Several substrates were assessed with each variant, demonstrating the critical importance of substrate profiling to accurately assess the value of an amino acid change, as some promising substitutions proved to be highly substrate dependent in their catalytic gains.

In chapter four, my engineering efforts toward altering the regioselectivity CDMO are discussed. In this work, two main hot spots for changing the regioselectivity were found to be of the most interest. I demonstrate the dynamic range of regioisomers that can be produced as a result of protein engineering work on CDMO.

Bibliography

1. Walsh, C., Enabling the chemistry of life. *Nature* **2001,** *409*, 226.

2. Bornscheuer, U. T.; Huisman, G. W.; Kazlauskas, R. J.; Lutz, S.; Moore, J. C.; Robins, K., Engineering the third wave of biocatalysis. *Nature* **2012,** *485* (7397), 185-94.

3. Buchholz, K.; Kasche, V.; Bornscheuer, U., *Biocatalysts and Enzyme Technology*. 2 ed.; Wiley-VCH: 2012.

4. Drauz, K.; Gröger, H.; May, O., *Enzyme Catalysis in Organic Synthesis*. 3 ed.; Wiley-VCH: 2012; p 2038.

5. Liese, A.; Seelbach, K.; Wandrey, C., *Industrial Biotransformations*. 2 ed.; Wiley-VCH: 2006; p 570.

6. Buchner, E., Alkoholische Gährung ohne Hefezellen. *Berichte der deutschen chemischen Gesellschaft* **1897,** *30* (1), 117-124.

7. Jaenicke, L., Centenary of the Award of a Nobel Prize to Eduard Buchner, the Father of Biochemistry in a Test Tube and Thus of Experimental Molecular Bioscience. *Angewandte Chemie International Edition* **2007,** *46* (36), 6776-6782.

8. Rosenthaler, L., Durch enzyme bewirkte asymmetrische synthesen. *Biochem. Z* **1908,** *14*, 238-253.

9. Sedlaczek, L.; Smith, L. L., Biotransformations of Steroids. *Crit. Rev. Biotechnol.* **1988,** *7* (3), 187-236.

10. Estell, D. A.; Graycar, T. P.; Wells, J. A., Engineering an enzyme by site-directed mutagenesis to be resistant to chemical oxidation. *J. Biol. Chem.* **1985,** *260* (11), 6518-6521.

11. Jensen, V. J.; Rugh, S., Industrial-scale production and application of immobilized glucose isomerase. In *Methods Enzymol.*, Elsevier: 1987; Vol. 136, pp 356-370.

12. Bruggink, A.; Roos, E. C.; de Vroom, E., Penicillin acylase in the industrial production of β-lactam antibiotics. *Org. Process Res. Dev.* **1998,** *2* (2), 128-133.

13. Reetz, M. T., Biocatalysis in Organic Chemistry and Biotechnology: Past, Present, and Future. *J. Am. Chem. Soc.* **2013,** *135* (34), 12480-12496.

14. Lutz, S.; Iamurri, S. M., Protein Engineering: Past, Present, and Future. In *Protein Engineering: Methods and Protocols*, Bornscheuer, U. T.; Höhne, M., Eds. Springer New York: New York, NY, 2018; pp 1-12.

15. Sheldon, R. A.; Pereira, P. C., Biocatalysis engineering: the big picture. *Chem. Soc. Rev.* **2017,** *46* (10), 2678-2691.

16. Porter, J. L.; Rusli, R. A.; Ollis, D. L., Directed Evolution of Enzymes for Industrial Biocatalysis. *ChemBioChem* **2016,** *17* (3), 197-203.

17. Glick, B. R.; Pasternak, J. J.; Patten, C. L., *Molecular Biotechnology: Principles and Applications of Recombinant DNA* ASM Press: Washington DC, 2010.

18. Brannigan, J. A.; Wilkinson, A. J., Protein engineering 20 years on. *Nat. Rev. Mol. Cell Biol.* **2002,** *3*, 964.

19. Eijsink, V. G. H.; Bjørk, A.; Gåseidnes, S.; Sirevåg, R.; Synstad, B.; Burg, B. v. d.; Vriend, G., Rational engineering of enzyme stability. *J. Biotechnol.* **2004,** *113* (1), 105-120.

20. Cedrone, F.; Ménez, A.; Quéméneur, E., Tailoring new enzyme functions by rational redesign. *Curr. Opin. Struct. Biol.* **2000,** *10* (4), 405-410.

21. Taylor, S. V.; Kast, P.; Hilvert, D., Investigating and Engineering Enzymes by Genetic Selection. *Angewandte Chemie International Edition* **2001,** *40* (18), 3310-3335.

22. Yang, G.; Rich, J. R.; Gilbert, M.; Wakarchuk, W. W.; Feng, Y.; Withers, S. G., Fluorescence Activated Cell Sorting as a General Ultra-High-Throughput Screening Method for Directed Evolution of Glycosyltransferases. *J. Am. Chem. Soc.* **2010,** *132* (30), 10570-10577.

23. Reetz, M. T., Combinatorial and Evolution-Based Methods in the Creation of Enantioselective Catalysts. *Angewandte Chemie International Edition* **2001,** *40* (2), 284-310.

24. Neylon, C., Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution. *Nuc. Acids Res.* **2004,** *32* (4), 1448-1459.

25. Economou, C.; Chen, K.; Arnold, F. H., Random mutagenesis to enhance the activity of subtilisin in organic solvents: Characterization of Q103R subtilisin E. *Biotechnol. Bioeng.* **1992,** *39* (6), 658-662.

26. Lutz, S.; Bornscheuer, U. T., *Protein engineering handbook*. John Wiley & Sons: 2012.

27. Leung, D.; Chen, E.; Goeddel, D., A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. *Technique* **1989,** *1*, 11-15.

28. Zhao, H.; Arnold, F. H., Optimization of DNA Shuffling for High Fidelity Recombination. *Nuc. Acids Res.* **1997,** *25* (6), 1307-1308.

29. Stemmer, W. P. C., Rapid evolution of a protein in vitro by DNA shuffling. *Nature* **1994,** *370*, 389.

30. Reetz, M. T., Laboratory Evolution of Stereoselective Enzymes: A Prolific Source of Catalysts for Asymmetric Reactions. *Angewandte Chemie International Edition* **2011,** *50* (1), 138-174.

31. Quin, M. B.; Schmidt-Dannert, C., Engineering of Biocatalysts: from Evolution to Creation. *ACS Catalysis* **2011,** *1* (9), 1017-1021.

32. Brustad, E. M.; Arnold, F. H., Optimizing non-natural protein function with directed evolution. *Curr. Opin. Chem. Biol.* **2011,** *15* (2), 201-210.

33. Jäckel, C.; Hilvert, D., Biocatalysts by evolution. *Curr. Opin. Biotechnol.* **2010,** *21* (6), 753-759.

34. Turner, N. J., Directed evolution drives the next generation of biocatalysts. *Nat. Chem. Biol.* **2009,** *5*, 567.

35. Bommarius, A. S.; Blum, J. K.; Abrahamson, M. J., Status of protein engineering for biocatalysts: how to design an industrially useful biocatalyst. *Curr. Opin. Chem. Biol.* **2011,** *15* (2), 194-200.

36. Shivange, A. V.; Marienhagen, J.; Mundhada, H.; Schenk, A.; Schwaneberg, U., Advances in generating functional diversity for directed protein evolution. *Curr. Opin. Chem. Biol.* **2009,** *13* (1), 19-25.

37. Bershtein, S.; Tawfik, D. S., Advances in laboratory evolution of enzymes. *Curr. Opin. Chem. Biol.* **2008,** *12* (2), 151-158.

38. Liang, J.; Lalonde, J.; Borup, B.; Mitchell, V.; Mundorff, E.; Trinh, N.; Kochrekar, D. A.; Nair Cherat, R.; Pai, G. G., Development of a Biocatalytic Process as an Alternative to the (−)-DIP-Cl-Mediated Asymmetric Reduction of a Key Intermediate of Montelukast. *Org. Process Res. Dev.* **2010,** *14* (1), 193-198.

39. Reetz, M. T., An Overview of High-Throughput Screening Systems for Enantioselective Enzymatic Transformations. In *Directed Enzyme Evolution: Screening and Selection Methods*, Arnold, F. H.; Georgiou, G., Eds. Humana Press: Totowa, NJ, 2003; pp 259-282.

40. Reymond, J.-L., *Enzyme Assays: High throughput Screening, Genetic Selection and Fingerprinting*. Wiley-VCH Verlag GmbH & Co. KGaA: 2006.

41. Reetz, M. T., Directed Evolution as a Means to Engineer Enantioselective Enzymes. In *Asymmetric Organic Synthesis with Enzymes*, Wiley-VCH Verlag GmbH & Co. KGaA: 2008; pp 21-63.

42. Agresti, J. J.; Antipov, E.; Abate, A. R.; Ahn, K.; Rowat, A. C.; Baret, J.-C.; Marquez, M.; Klibanov, A. M.; Griffiths, A. D.; Weitz, D. A., Ultrahigh-throughput screening in drop-based microfluidics for directed evolution. *Proc Natl Acad Sci USA* **2010,** *107* (9), 4004-4009.

43. Sjostrom, S. L.; Bai, Y.; Huang, M.; Liu, Z.; Nielsen, J.; Joensson, H. N.; Andersson Svahn, H., High-throughput screening for industrial enzyme production hosts by droplet microfluidics. *Lab Chip* **2014,** *14* (4), 806-813.

44. Diefenbach, X. W.; Farasat, I.; Guetschow, E. D.; Welch, C. J.; Kennedy, R. T.; Sun, S.; Moore, J. C., Enabling Biocatalysis by High-Throughput Protein Engineering Using Droplet Microfluidics Coupled to Mass Spectrometry. *ACS Omega* **2018,** *3* (2), 1498-1508.

45. Sun, S.; Slaney, T. R.; Kennedy, R. T., Label Free Screening of Enzyme Inhibitors at Femtomole Scale Using Segmented Flow Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **2012,** *84* (13), 5794-5800.

46. Mair, P.; Gielen, F.; Hollfelder, F., Exploring sequence space in search of functional enzymes using microfluidic droplets. *Curr. Opin. Chem. Biol.* **2017,** *37*, 137-144.

47. Lutz, S.; Patrick, W. M., Novel methods for directed evolution of enzymes: quality, not quantity. *Curr. Opin. Biotechnol.* **2004,** *15* (4), 291-297.

48. You, L.; Arnold, F., Directed evolution of subtilisin E in Bacillus subtilis to enhance total activity in aqueous dimethylformamide. *Protein Eng. Des. Sel.* **1996,** *9* (1), 77-83.

49. Aharoni, A.; Thieme, K.; Chiu, C. P. C.; Buchini, S.; Lairson, L. L.; Chen, H.; Strynadka, N. C. J.; Wakarchuk, W. W.; Withers, S. G., High-throughput screening methodology for the directed evolution of glycosyltransferases. *Nat. Meth.* **2006,** *3*, 609.

50. Smith, G., Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **1985,** *228* (4705), 1315-1317.

51. Freudl, R.; MacIntyre, S.; Degen, M.; Henning, U., Cell surface exposure of the outer membrane protein OmpA of Escherichia coli K-12. *J Mol Biol* **1986,** *188* (3), 491-494.

52. Boder, E. T.; Wittrup, K. D., Yeast surface display for screening combinatorial polypeptide libraries. *Nat Biotech* **1997,** *15*, 553.

53. Xu, X.; Gao, C.; Zhang, X.; Che, B.; Ma, C.; Qiu, J.; Tao, F.; Xu, P., Production of N-Acetyl-d-Neuraminic Acid by Use of an Efficient Spore Surface Display System. *Appl. Environ. Microbiol.* **2011,** *77* (10), 3197-3201.

54. Smith, M. R.; Khera, E.; Wen, F., Engineering Novel and Improved Biocatalysts by Cell Surface Display. *Ind. Eng. Chem. Res.* **2015,** *54* (16), 4021-4032.

55. Tawfik, D. S.; Griffiths, A. D., Man-made cell-like compartments for molecular evolution. *Nat Biotech* **1998,** *16*, 652.

56. Gibson, D. G.; Glass, J. I.; Lartigue, C.; Noskov, V. N.; Chuang, R.-Y.; Algire, M. A.; Benders, G. A.; Montague, M. G.; Ma, L.; Moodie, M. M.; Merryman, C.; Vashee, S.; Krishnakumar, R.; Assad-Garcia, N.; Andrews-Pfannkoch, C.; Denisova, E. A.; Young, L.; Qi, Z.-Q.; Segall-Shapiro, T. H.; Calvey, C. H.; Parmar, P. P.; Hutchison, C. A.; Smith, H. O.; Venter, J. C., Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome. *Science* **2010,** *329* (5987), 52-56.

57. Höhne, M.; Schätzle, S.; Jochens, H.; Robins, K.; Bornscheuer, U. T., Rational assignment of key motifs for function guides in silico enzyme identification. *Nat. Chem. Biol.* **2010,** *6*, 807.

58. Clouthier, C. M.; Kayser, M. M.; Reetz, M. T., Designing new Baeyer-Villiger monooxygenases using restricted CASTing. *J. Org. Chem.* **2006,** *71*, 8431-8437.

59. Reetz, M. T.; Bocola, M.; Carballeira, J. D.; Zha, D.; Vogel, A., Expanding the Range of Substrate Acceptance of Enzymes: Combinatorial Active-Site Saturation Test. *Angewandte Chemie International Edition* **2005,** *44* (27), 4192-4196.

60. Morley, K. L.; Kazlauskas, R. J., Improving enzyme properties: when are closer mutations better? *Trends Biotechnol.* **2005,** *23* (5), 231-237.

61. Reetz, M. T.; Carballeira, J. D., Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. *Nat. Protoc.* **2007,** *2*, 891.

62. Jensen, R. A., Enzyme Recruitment in Evolution of New Function. *Annu. Rev. Microbiol.* **1976,** *30* (1), 409-425.

63. Alcalde, M., When directed evolution met ancestral enzyme resurrection. *Microbial biotechnology* **2017,** *10* (1), 22-24.

64. Aharoni, A.; Gaidukov, L.; Khersonsky, O.; Gould, S. M.; Roodveldt, C.; Tawfik, D. S., The 'evolvability' of promiscuous protein functions. *Nat. Genet.* **2004,** *37*, 73.

65. Alcolombri, U.; Elias, M.; Tawfik, D. S., Directed Evolution of Sulfotransferases and Paraoxonases by Ancestral Libraries. *J Mol Biol* **2011,** *411* (4), 837-853.

66. Thornton, J. W., Resurrecting ancient genes: experimental analysis of extinct molecules. *Nat. Rev. Genet.* **2004,** *5*, 366.

67. Ehren, J.; Govindarajan, S.; Morón, B.; Minshull, J.; Khosla, C., Protein engineering of improved prolyl endopeptidases for celiac sprue therapy. *Protein Eng. Des. Sel.* **2008,** *21* (12), 699-707.

68. Minshull, J.; Govindarajan, S.; Cox, T.; Ness, J. E.; Gustafsson, C., Engineered protein function by selective amino acid diversification. *Methods* **2004,** *32* (4), 416-427.

69. Liao, J.; Warmuth, M. K.; Govindarajan, S.; Ness, J. E.; Wang, R. P.; Gustafsson, C.; Minshull, J., Engineering proteinase K using machine learning and synthetic genes. *BMC Biotechnol.* **2007,** *7* (1), 16.

70. Fox, R. J.; Davis, S. C.; Mundorff, E. C.; Newman, L. M.; Gavrilovic, V.; Ma, S. K.; Chung, L. M.; Ching, C.; Tam, S.; Muley, S.; Grate, J.; Gruber, J.; Whitman, J. C.; Sheldon, R. A.; Huisman, G. W., Improving catalytic function by ProSAR-driven enzyme evolution. *Nat Biotech* **2007,** *25*, 338.

71. Cheng, F.; Zhu, L.; Schwaneberg, U., Directed evolution 2.0: improving and deciphering enzyme properties. *Chem. Commun.* **2015,** *51* (48), 9760-9772.

72. Bornscheuer, U. T.; Kazlauskas, R. J., Catalytic Promiscuity in Biocatalysis: Using Old Enzymes to Form New Bonds and Follow New Pathways. *Angewandte Chemie International Edition* **2004,** *43* (45), 6032-6040.

73. Khersonsky, O.; Tawfik, D. S., Enzyme Promiscuity: A Mechanistic and Evolutionary Perspective. *Annu. Rev.Biochem.* **2010,** *79* (1), 471-505.

74. Kiss, G.; Çelebi-Ölçüm, N.; Moretti, R.; Baker, D.; Houk, K. N., Computational Enzyme Design. *Angewandte Chemie International Edition* **2013,** *52* (22), 5700-5725.

75. Siegel, J. B.; Zanghellini, A.; Lovick, H. M.; Kiss, G.; Lambert, A. R.; St.Clair, J. L.; Gallaher, J. L.; Hilvert, D.; Gelb, M. H.; Stoddard, B. L.; Houk, K. N.; Michael, F. E.; Baker, D., Computational Design of an Enzyme Catalyst for a Stereoselective Bimolecular Diels-Alder Reaction. *Science* **2010,** *329* (5989), 309-313.

76. Röthlisberger, D.; Khersonsky, O.; Wollacott, A. M.; Jiang, L.; DeChancie, J.; Betker, J.; Gallaher, J. L.; Althoff, E. A.; Zanghellini, A.; Dym, O.; Albeck, S.; Houk, K. N.; Tawfik, D. S.; Baker, D., Kemp elimination catalysts by computational enzyme design. *Nature* **2008,** *453*, 190.

77. Giger, L.; Caner, S.; Obexer, R.; Kast, P.; Baker, D.; Ban, N.; Hilvert, D., Evolution of a designed retro-aldolase leads to complete active site remodeling. *Nat. Chem. Biol.* **2013,** *9*, 494.

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

79. Kung, H. F.; Redfield, B.; Treadwell, B. V.; Eskin, B.; Spears, C.; Weissbach, H., DNA-directed in vitro synthesis of beta-galactosidase. Studies with purified factors. *J. Biol. Chem.* **1977,** *252* (19), 6889-94.

80. Stiege, W.; Erdmann, V. A., The potentials of the in vitro protein biosynthesis system. *J. Biotechnol.* **1995,** *41* (2), 81-90.

81. Shimizu, Y.; Inoue, A.; Tomari, Y.; Suzuki, T.; Yokogawa, T.; Nishikawa, K.; Ueda, T., Cell-free translation reconstituted with purified components. *Nat Biotech* **2001,** *19* (8), 751-755.

82. Spirin, A. S.; Baranov, V. I.; Ryabova, L. A.; Ovodov, S. Y.; Alakhov, Y. B., A continuous cell-free translation system capable of producing polypeptides in high yield. *Science* **1988,** *242* (4882), 1162-1164.

83. Moore, S. D.; Sauer, R. T., The tmRNA system for translational surveillance and ribosome rescue. *Annu. Rev. Biochem.* **2007,** *76*, 101-124.

84. Jewett, M. C.; Calhoun, K. A.; Voloshin, A.; Wuu, J. J.; Swartz, J. R., An integrated cell-free metabolic platform for protein production and synthetic biology. *Mol. Syst. Biol.* **2008,** *4* (1), 220.

85. Dudley, Q. M.; Karim, A. S.; Jewett, M. C., Cell-free metabolic engineering: Biomanufacturing beyond the cell. *Biotechnol. J.* **2015,** *10* (1), 69-82.

86. Shimizu, Y.; Kanamori, T.; Ueda, T., Protein synthesis by pure translation systems. *Methods* **2005,** *36* (3), 299-304.

87. Li, J.; Zhang, C.; Huang, P.; Kuru, E.; Forster-Benson, E. T. C.; Li, T.; Church, G. M., Dissecting limiting factors of the Protein synthesis Using Recombinant Elements (PURE) system. *Translation* **2017,** *5* (1).

88. Li, J.; Gu, L.; Aach, J.; Church, G. M., Improved Cell-Free RNA and Protein Synthesis System. *PLoS One* **2014,** *9* (9).

89. Kannan, K.; Kanabar, P.; Schryer, D.; Florin, T.; Oh, E.; Bahroos, N.; Tenson, T.; Weissman, J. S.; Mankin, A. S., The general mode of translation inhibition by macrolide antibiotics. *Proc Natl Acad Sci USA* **2014,** *111* (45), 15958-15963.

90. Josephson, K.; Hartman, M. C. T.; Szostak, J. W., Ribosomal Synthesis of Unnatural Peptides. *J. Am. Chem. Soc.* **2005,** *127* (33), 11727-11735.

91. Forster, A. C.; Tan, Z.; Nalam, M. N. L.; Lin, H.; Qu, H.; Cornish, V. W.; Blacklow, S. C., Programming peptidomimetic syntheses by translating genetic codes designed *de novo*. *Proc Natl Acad Sci USA* **2003,** *100* (11), 6353-6357.

92. Lipovsek, D.; Plückthun, A., In-vitro protein evolution by ribosome display and mRNA display. *J. Immunol. Methods* **2004,** *290* (1), 51-67.

93. Fen, C. X.; Coomber, D. W.; Lane, D. P.; Ghadessy, F. J., Directed Evolution of p53 Variants with Altered DNA-binding Specificities by In Vitro Compartmentalization. *J Mol Biol* **2007,** *371* (5), 1238-1248.

94. Daugherty, A. B.; Horton, J. R.; Cheng, X.; Lutz, S., Structural and Functional Consequences of Circular Permutation on the Active Site of Old Yellow Enzyme. *ACS Catalysis* **2014**, 892-899.

95. Niwa, T.; Kanamori, T.; Ueda, T.; Taguchi, H., Global analysis of chaperone effects using a reconstituted cell-free translation system. *Proc Natl Acad Sci USA* **2012,** *109* (23), 8937-8942.

96. Niwa, T.; Ying, B.-W.; Saito, K.; Jin, W.; Takada, S.; Ueda, T.; Taguchi, H., Bimodal protein solubility distribution revealed by an aggregation analysis of the entire ensemble of Escherichia coli proteins. *Proc Natl Acad Sci USA* **2009,** *106* (11), 4201-4206.

97. Carlson, E. D.; Gan, R.; Hodgman, C. E.; Jewett, M. C., Cell-free protein synthesis: Applications come of age. *Biotechnol. Adv.* **2012,** *30* (5), 1185-1194.

98. Kuruma, Y.; Ueda, T., The PURE system for the cell-free synthesis of membrane proteins. *Nat. Protoc.* **2015,** *10*, 1328.

99. Matsubayashi, H.; Ueda, T., Purified cell-free systems as standard parts for synthetic biology. *Curr. Opin. Chem. Biol.* **2014,** *22*, 158-162.

100. Joosten, V.; van Berkel, W. J. H., Flavoenzymes. *Curr. Opin. Chem. Biol.* **2007,** *11* (2), 195-202.

101. Jortzik, E.; Wang, L.; Ma, J.; Becker, K., Flavins and Flavoproteins: Applications in Medicine. In *Flavins and Flavoproteins: Methods and Protocols*, Weber, S.; Schleicher, E., Eds. Springer New York: New York, NY, 2014; pp 113-157.

102. Chen, H.; O'Connor, S.; Cane, D. E.; Walsh, C. T., Epothilone biosynthesis: assembly of the methylthiazolylcarboxy starter unit on the EpoB subunit. *Chem. Biol.* **2001,** *8* (9), 899-912.

103. Martínez, A. T.; Ruiz-Dueñas, F. J.; Camarero, S.; Serrano, A.; Linde, D.; Lund, H.; Vind, J.; Tovborg, M.; Herold-Majumdar, O. M.; Hofrichter, M.; Liers, C.; Ullrich, R.; Scheibner, K.; Sannia, G.; Piscitelli, A.; Pezzella, C.; Sener, M. E.; Kılıç, S.; van Berkel, W. J. H.; Guallar, V.; Lucas, M. F.; Zuhse, R.; Ludwig, R.; Hollmann, F.; Fernández-Fueyo, E.; Record, E.; Faulds, C. B.; Tortajada, M.; Winckelmann, I.; Rasmussen, J.-A.; Gelo-Pujic, M.; Gutiérrez, A.; del Río, J. C.; Rencoret, J.; Alcalde, M., Oxidoreductases on their way to industrial biotransformations. *Biotechnol. Adv.* **2017,** *35* (6), 815-831.

104. Chaiyen, P.; Scrutton, N. S., Special Issue: Flavins and Flavoproteins. *FEBS J* **2015,** *282* (16), 3001-3002.

105. Teufel, R.; Miyanaga, A.; Michaudel, Q.; Stull, F.; Louie, G.; Noel, J. P.; Baran, P. S.; Palfey, B.; Moore, B. S., Flavin-mediated dual oxidation controls an enzymatic Favorskii-type rearrangement. *Nature* **2013,** *503*, 552.

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

107. Yeh, E.; Garneau, S.; Walsh, C. T., Robust in vitro activity of RebF and RebH, a two-component reductase/halogenase, generating 7-chlorotryptophan during rebeccamycin biosynthesis. *Proc Natl Acad Sci USA* **2005,** *102* (11), 3960-3965.

108. Xu, F., Applications of oxidoreductases: Recent progress. *Industrial Biotechnology* **2005,** *1* (1), 38-50.

109. Ko, H.-S.; Yokoyama, Y.; Ohno, N.; Okadome, M.; Amachi, S.; Shinoyama, H.; Fujii, T., Purification and characterization of intracellular and extracellular, thermostable and alkali-tolerant alcohol oxidases produced by a thermophilic fungus, Thermoascus aurantiacus NBRC 31693. *J. Biosci. Bioeng.* **2005,** *99* (4), 348-353.

110. Hefti, M. H.; Vervoort, J.; van Berkel, W. J. H., Deflavination and reconstitution of flavoproteins. *Eur. J. Biochem.* **2003,** *270* (21), 4227-4242.

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

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

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

114. Mattevi, A., To be or not to be an oxidase: challenging the oxygen reactivity of flavoenzymes. *Trends Biochem. Sci.* **2006,** *31* (5), 276-283.

115. Piano, V.; Palfey, B. A.; Mattevi, A., Flavins as Covalent Catalysts: New Mechanisms Emerge. *Trends Biochem. Sci.* **2017,** *42* (6), 457-469.

116. 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.

117. Teufel, R.; Stull, F.; Meehan, M. J.; Michaudel, Q.; Dorrestein, P. C.; Palfey, B.; Moore, B. S., Biochemical Establishment and Characterization of EncM's Flavin-N5-oxide Cofactor. *J. Am. Chem. Soc.* **2015,** *137* (25), 8078-8085.

118. Tanner, J. J.; Boechi, L.; Andrew McCammon, J.; Sobrado, P., Structure, mechanism, and dynamics of UDP-galactopyranose mutase. *Arch. Biochem. Biophys.* **2014,** *544*, 128-141.

119. Mehra-Chaudhary, R.; Dai, Y.; Sobrado, P.; Tanner, J. J., In Crystallo Capture of a Covalent Intermediate in the UDP-Galactopyranose Mutase Reaction. *Biochemistry* **2016,** *55* (6), 833-836.

120. Razeto, A.; Mattiroli, F.; Carpanelli, E.; Aliverti, A.; Pandini, V.; Coda, A.; Mattevi, A., The Crucial Step in Ether Phospholipid Biosynthesis: Structural Basis of a Noncanonical Reaction Associated with a Peroxisomal Disorder. *Structure 15* (6), 683-692.

121. Myllykallio, H.; Lipowski, G.; Leduc, D.; Filee, J.; Forterre, P.; Liebl, U., An Alternative Flavin-Dependent Mechanism for Thymidylate Synthesis. *Science* **2002,** *297* (5578), 105-107.

122. Koehn, E. M.; Perissinotti, L. L.; Moghram, S.; Prabhakar, A.; Lesley, S. A.; Mathews, I. I.; Kohen, A., Folate binding site of flavin-dependent thymidylate synthase. *Proc Natl Acad Sci USA* **2012,** *109* (39), 15722-15727.

123. Payne, K. A. P.; White, M. D.; Fisher, K.; Khara, B.; Bailey, S. S.; Parker, D.; Rattray, N. J. W.; Trivedi, D. K.; Goodacre, R.; Beveridge, R.; Barran, P.; Rigby, S. E. J.; Scrutton, N. S.; Hay, S.; Leys, D., New cofactor supports α,β-unsaturated acid decarboxylation via 1,3-dipolar cycloaddition. *Nature* **2015,** *522*, 497.

124. Marshall, S. A.; Fisher, K.; Ní Cheallaigh, A.; White, M. D.; Payne, K. A. P.; Parker, D. A.; Rigby, S. E. J.; Leys, D., Oxidative Maturation and Structural Characterization of Prenylated FMN Binding by UbiD, a Decarboxylase Involved in Bacterial Ubiquinone Biosynthesis. *J. Biol. Chem.* **2017,** *292* (11), 4623-4637.

125. White, M. D.; Payne, K. A. P.; Fisher, K.; Marshall, S. A.; Parker, D.; Rattray, N. J. W.; Trivedi, D. K.; Goodacre, R.; Rigby, S. E. J.; Scrutton, N. S.; Hay, S.; Leys, D., UbiX is a flavin prenyltransferase required for bacterial ubiquinone biosynthesis. *Nature* **2015,** *522*, 502.

126. Leys, D.; Scrutton, N. S., Sweating the assets of flavin cofactors: new insight of chemical versatility from knowledge of structure and mechanism. *Curr. Opin. Struct. Biol.* **2016,** *41*, 19-26.

127. Warburg, O.; Christian, W., Ein zweites sauerstoffübertragendes Ferment und sein Absorptionsspektrum. *Naturwissenschaften* **1932,** *20* (37), 688-688.

128. Haas, E., Isolation of a new yellow enzyme. *Biochem. Ztschr.* **1938,** *298*, 378-390.

129. Williams, R. E.; Bruce, N. C., 'New uses for an Old Enzyme' – the Old Yellow Enzyme family of flavoenzymes. *Microbiology* **2002,** *148* (6), 1607-1614.

130. Fu, Y.; Castiglione, K.; Weuster-Botz, D., Comparative characterization of novel ene-reductases from cyanobacteria. *Biotechnol. Bioeng.* **2013,** *110* (5), 1293-1301.

131. Gao, X.; Ren, J.; Wu, Q.; Zhu, D., Biochemical characterization and substrate profiling of a new NADH-dependent enoate reductase from Lactobacillus casei. *Enzyme. Microb. Technol.* **2012,** *51* (1), 26-34.

132. Nizam, S.; Verma, S.; Borah, N. N.; Gazara, R. K.; Verma, P. K., Comprehensive genome-wide analysis reveals different classes of enigmatic old yellow enzyme in fungi. *Scientific Reports* **2014,** *4*, 4013.

133. Binks, P. R.; French, C. E.; Nicklin, S.; Bruce, N. C., Degradation of pentaerythritol tetranitrate by Enterobacter cloacae PB2. *Appl. Environ. Microbiol.* **1996,** *62* (4), 1214-1219.

134. Hall, M.; Stueckler, C.; Kroutil, W.; Macheroux, P.; Faber, K., Asymmetric Bioreduction of Activated Alkenes Using Cloned 12-Oxophytodienoate Reductase Isoenzymes OPR-1 and OPR-3 from Lycopersicon esculentum (Tomato): A Striking Change of Stereoselectivity. *Angewandte Chemie International Edition* **2007,** *46* (21), 3934-3937.

135. Müller, A.; Hauer, B.; Rosche, B., Asymmetric alkene reduction by yeast old yellow enzymes and by a novel Zymomonas mobilis reductase. *Biotechnol. Bioeng.* **2007,** *98* (1), 22-29.

136. Kitzing, K.; Fitzpatrick, T. B.; Wilken, C.; Sawa, J.; Bourenkov, G. P.; Macheroux, P.; Clausen, T., The 1.3 Å Crystal Structure of the Flavoprotein YqjM Reveals a Novel Class of Old Yellow Enzymes. *J. Biol. Chem.* **2005,** *280* (30), 27904-27913.

137. Oberdorfer, G.; Steinkellner, G.; Stueckler, C.; Faber, K.; Gruber, K., Stereopreferences of Old Yellow Enzymes: Structure Correlations and Sequence Patterns in Enoate Reductases. *ChemCatChem* **2011,** *3* (10), 1562-1566.

138. Toogood, H. S.; Gardiner, J. M.; Scrutton, N. S., Biocatalytic Reductions and Chemical Versatility of the Old Yellow Enzyme Family of Flavoprotein Oxidoreductases. *ChemCatChem* **2010,** *2* (8), 892-914.

139. Fox, K. M.; Karplus, P. A., Old yellow enzyme at 2 A resolution: overall structure, ligand binding, and comparison with related flavoproteins. *Curr. Biol.* **1994,** *2* (11), 1089-1105.

140. Adalbjörnsson, B. V.; Toogood, H. S.; Fryszkowska, A.; Pudney, C. R.; Jowitt, T. A.; Leys, D.; Scrutton, N. S., Biocatalysis with Thermostable Enzymes: Structure and Properties of a Thermophilic 'ene'-Reductase related to Old Yellow Enzyme. *ChemBioChem* **2010,** *11* (2), 197-207.

141. Kohli, R. M.; Massey, V., The Oxidative Half-reaction of Old Yellow Enzyme: The Role of Tyrosine 196. *J. Biol. Chem.* **1998,** *273* (49), 32763-32770.

142. Opperman, D. J.; Sewell, B. T.; Litthauer, D.; Isupov, M. N.; Littlechild, J. A.; van Heerden, E., Crystal structure of a thermostable Old Yellow Enzyme from Thermus scotoductus SA-01. *Biochem. Biophys. Res. Commun.* **2010,** *393* (3), 426-431.

143. Khan, H.; Barna, T.; Bruce, N. C.; Munro, A. W.; Leys, D.; Scrutton, N. S., Proton transfer in the oxidative half-reaction of pentaerythritol tetranitrate reductase. *FEBS J.* **2005,** *272* (18), 4660-4671.

144. Brown, B. J.; Deng, Z.; Karplus, P. A.; Massey, V., On the Active Site of Old Yellow Enzyme: Role of Histidine 191 and Asparagine 194. *J. Biol. Chem.* **1998,** *273* (49), 32753-32762.

145. Theorell, H.; Yagi, K.; Ludwig, G. D.; Egami, F., Effect of Flavin Monosulphate on Old Yellow Enzyme. *Nature* **1957,** *180*, 922.

146. Nakamura, T.; Yoshimura, J.; Ogura, Y., Action Mechanism of the Old Yellow Enzyme. *The Journal of Biochemistry* **1965,** *57* (4), 554-564.

147. Fryszkowska, A.; Toogood, H. S.; Mansell, D.; Stephens, G.; Gardiner, J. M.; Scrutton, N. S., A surprising observation that oxygen can affect the product enantiopurity of an enzyme-catalysed reaction. *FEBS J.* **2012,** *279* (22), 4160-4171.

148. Hall, M.; Stueckler, C.; Hauer, B.; Stuermer, R.; Friedrich, T.; Breuer, M.; Kroutil, W.; Faber, K., Asymmetric Bioreduction of Activated C=C Bonds UsingZymomonas mobilis NCR Enoate Reductase and Old Yellow Enzymes OYE 1–3 from Yeasts. *Eur. J. Org. Chem.* **2008,** *2008* (9), 1511-1516.

149. Amato, E. D.; Stewart, J. D., Applications of protein engineering to members of the old yellow enzyme family. *Biotechnol. Adv.* **2015,** *33* (5), 624-631.

150. Reich, S.; Hoeffken, H. W.; Rosche, B.; Nestl, B. M.; Hauer, B., Crystal structure determination and mutagenesis analysis of the ene reductase NCR. *ChemBioChem* **2012,** *13* (16), 2400-7.

151. Hulley, M. E.; Toogood, H. S.; Fryszkowska, A.; Mansell, D.; Stephens, G. M.; Gardiner, J. M.; Scrutton, N. S., Focused directed evolution of pentaerythritol tetranitrate reductase by using automated anaerobic kinetic screening of site-saturated libraries. *ChemBioChem* **2010,** *11* (17), 2433-47.

152. Toogood, H. S.; Fryszkowska, A.; Hulley, M.; Sakuma, M.; Mansell, D.; Stephens, G. M.; Gardiner, J. M.; Scrutton, N. S., A site-saturated mutagenesis study of pentaerythritol

tetranitrate reductase reveals that residues 181 and 184 influence ligand binding, stereochemistry and reactivity. *ChemBioChem* **2011,** *12* (5), 738-49.

153. Bougioukou, D. J.; Kille, S.; Taglieber, A.; Reetz, M. T., Directed Evolution of an Enantioselective Enoate-Reductase: Testing the Utility of Iterative Saturation Mutagenesis. *Adv. Synth. Catal.* **2009,** *351* (18), 3287-3305.

154. Horita, S.; Kataoka, M.; Kitamura, N.; Nakagawa, T.; Miyakawa, T.; Ohtsuka, J.; Nagata, K.; Shimizu, S.; Tanokura, M., An engineered old yellow enzyme that enables efficient synthesis of (4R,6R)-Actinol in a one-pot reduction system. *ChemBioChem* **2015,** *16* (3), 440-5.

155. Reich, S.; Kress, N.; Nestl, B. M.; Hauer, B., Variations in the stability of NCR ene reductase by rational enzyme loop modulation. *J. Struct. Biol.* **2014,** *185* (2), 228-233.

156. Reich, S.; Nestl, B. M.; Hauer, B., Loop-Grafted Old Yellow Enzymes in the Bienzymatic Cascade Reduction of Allylic Alcohols. *ChemBioChem* **2016,** *17* (7), 561-565.

157. Brown, B. J.; Hyun, J. W.; Duvvuri, S.; Karplus, P. A.; Massey, V., The role of glutamine 114 in old yellow enzyme. *J Biol Chem* **2002,** *277* (3), 2138-45.

158. Karplus, P. A.; Fox, K. M.; Massey, V., Structure-function relations for old yellow enzyme. *The FASEB Journal* **1995,** *9*, 1518-26.

159. Massey, V.; Schopfer, L. M., Reactivity of Old Yellow Enzyme with a-NADPH and Other Pyridine Nucleotide Derivatives. *J. Biol. Chem.* **1986,** *261* (3), 1215-1222.

160. Vaz, A. D. N.; Chakroborty, S.; Massey, V., Old Yellow Enzyme: Aromatization of Cyclic Enones and the Mechanism of a Novel Dismutation Reaction. *Biochemistry* **1995,** *34*, 4246-4256.

161. Xu, D.; Kohli, R. M.; Massey, V., The role of threonine 27 in flavin reactivity of old yellow enzyme. *Proc Natl Acad Sci USA* **1999,** *96*, 3556-3561.

162. Swiderska, M. A.; Stewart, J. D., Stereoselective enone reductions by Saccharomyces carlsbergensis old yellow enzyme. *J. Mol. Catal. B: Enzym.* **2006,** *42* (1-2), 52-54.

163. Padhi, S. K.; Bougioukou, D.; Stewart, J. D., Site-Saturation Mutagenesis of Tryptophan 116 of *Saccharomyces pastorianus* Old Yellow Enzyme Uncovers Stereocomplementary Variants. *J. Am. Chem. Soc.* **2009,** *131*, 3271-3280.

164. Pompeu, Y. A.; Sullivan, B.; Stewart, J. D., X-ray Crystallography Reveals How Subtle Changes Control the Orientation of Substrate Binding in an Alkene Reductase. *ACS Catalysis* **2013,** *3* (10), 2376-2390.

165. Walton, A. Z.; Sullivan, B.; Patterson-Orazem, A. C.; Stewart, J. D., Residues Controlling Facial Selectivity in an Alkene Reductase and Semirational Alterations to Create Stereocomplementary Variants. *ACS Catalysis* **2014,** *4* (7), 2307-2318.

166. Daugherty, A. B.; Govindarajan, S.; Lutz, S., Improved biocatalysts from a synthetic circular permutation library of the flavin-dependent oxidoreductase old yellow enzyme. *J. Am. Chem. Soc.* **2013,** *135* (38), 14425-32.

167. Daugherty, A. B.; Horton, J. R.; Cheng, X.; Lutz, S., Structural and Functional Consequences of Circular Permutation on the Active Site of Old Yellow Enzyme. *ACS Catalysis* **2015,** *5* (2), 892-899.

168. Baeyer, A.; Villiger, V., Einwirkung des Caro'schen Reagens auf Ketone. *Berichte der deutschen chemischen Gesellschaft* **1899,** *32* (3), 3625-3633.

169. Krow, G. R., The Baeyer–VilligerOxidation of Ketones and Aldehydes. In *Organic Reactions*, John Wiley & Sons, Inc.: 2004.

170. Criegee, R.; Schnorrenberg, W.; Becke, J., Zur Konstitution von Ketonperoxyden. *Justus Liebigs Ann. Chem.* **1949,** *565* (1), 7-21.

171. ten Brink, G. J.; Arends, I. W. C. E.; Sheldon, R. A., The Baeyer–Villiger Reaction: New Developments toward Greener Procedures. *Chem. Rev.* **2004,** *104* (9), 4105-4124.

172. Deslongchamps, P., Stereoelectronic effects in organic chemistry. Pergamon Press: Oxford, 1983; pp 313-314.

173. Noyori, R.; Sato, T.; Kobayashi, H., Remote substituent effects in the Baeyer-Villiger oxidation I. through-bond gamma substituent effect on the regioselectivity. *Tetrahedron Lett.* **1980,** *21*, 2569-2572.

174. 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.

175. Goodman, R. M.; Kishi, Y., Experimental Support for the Primary Stereoelectronic Effect Governing Baeyer−Villiger Oxidation and Criegee Rearrangement. *J. Am. Chem. Soc.* **1998,** *120* (36), 9392-9393.

176. Leisch, H.; Morley, K.; Lau, P. C. K., Baeyer−Villiger Monooxygenases: More Than Just Green Chemistry. *Chem. Rev.* **2011,** *111* (7), 4165-4222.

177. Turfitt, G. E., The microbiological degradation of steroids: 4. Fission of the steroid molecule. *Biochem. J.* **1948,** *42* (3), 376-383.

178. Griffin, M.; Trudgill, P. W., Purification and properties of cyclopentanone oxygenase of Pseudomonas NCIB 9872. *Eur. J. Biochem.* **1976,** *63*.

179. Trudgill, P. W., Cyclohexanone 1,2-Monooxygenase from Acinetobacter NCIMB 9871. *Methods Enzymol.* **1990,** *188* (1986), 70-77.

180. Balke, K.; Kadow, M.; Mallin, H.; Saß, S.; Bornscheuer, U. T.; Sa; Bornscheuer, U. T., Discovery, application and protein engineering of Baeyer-Villiger monooxygenases for organic synthesis. *Org. Biomol. Chem.* **2012,** *10* (31), 6249-6265.

181. De Gonzalo, G.; Mihovilovic, M. D.; Fraaije, M. W., Recent Developments in the Application of Baeyer-Villiger Monooxygenases as Biocatalysts. *ChemBioChem* **2010**.

182. 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.

183. 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* (1), 43-47.

184. 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. Microbiol. Biotechnol.* **2008,** *77*.

185. 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.

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

187. 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.

188. 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.

189. Mirza, I. A.; Yachnin, B. J.; Wang, S.; Grosse, S.; Bergeron, H.; Imura, A.; Iwaki, H.; Hasegawa, Y.; Lau, P. C. K.; 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.

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

191. Yachnin, B. J.; Sprules, T.; McEvoy, M. B.; Lau, P. C. K.; 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.

192. Zhang, Z. G.; Parra, L. P.; Reetz, M. T., Protein engineering of stereoselective Baeyer-Villiger monooxygenases. *Chemistry - A European Journal* **2012**.

193. Taschner, M. J.; Black, D. J.; Chen, Q.-Z., The enzymatic baeyer-villiger oxidation: A study of 4-substituted cyclohexanones. *Tetrahedron: Asymmetry* **1993,** *4* (6), 1387-1390.

194. Reetz, M. T.; Brunner, B.; Schneider, T.; Schulz, F.; Clouthier, C. M.; Kayser, M. M., Directed Evolution as a Method To Create Enantioselective Cyclohexanone Monooxygenases for Catalysis in Baeyer–Villiger Reactions. *Angewandte Chemie International Edition* **2004,** *43* (31), 4075-4078.

195. Mihovilovic, M. D.; Rudroff, F.; Winninger, A.; Schneider, T.; Schulz, F.; Reetz, M. T., Microbial Baeyer–Villiger Oxidation: Stereopreference and Substrate Acceptance of Cyclohexanone Monooxygenase Mutants Prepared by Directed Evolution. *Org. Lett.* **2006,** *8* (6), 1221-1224.

196. van Beek, H. L.; Romero, E.; Fraaije, M. W., Engineering Cyclohexanone Monooxygenase for the Production of Methyl Propanoate. *ACS Chem. Biol.* **2016**.

197. 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.

198. Bong, Y.; Clay, M.; Collier, S.; Mijts, B.; Vogel, M.; Zhang, X.; Zhu, J.; Nazor, J.; Smith, D.; Song, S., Engineered cylohexanone monooxygenases for synthesis of prazole compounds. *USA WO2011071982 A* **2011,** *2*, 2011.

199. Bučko, M.; Gemeiner, P.; Schenkmayerová, A.; Krajčovič, T.; Rudroff, F.; Mihovilovič, M. D., Baeyer-Villiger oxidations: biotechnological approach. *Appl. Microbiol. Biotechnol.* **2016,** *100* (15), 6585-6599.

200. Fraaije, M. W.; Wu, J.; Heuts, D. P. H. M.; 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*.

201. Malito, E.; Alfieri, A.; Fraaije, M. W.; Mattevi, A., Crystal structure of a Baeyer-Villiger monooxygenase. *Proc Natl Acad Sci USA* **2004,** *101*.

202. 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* (7-8), 979-986.

203. Reetz, M. T.; Wu, S., Greatly reduced amino acid alphabets in directed evolution: making the right choice for saturation mutagenesis at homologous enzyme positions. *Chem. Commun.* **2008**.

204. Reetz, M. T.; Wu, S., Laboratory Evolution of Robust and Enantioselective Baeyer−Villiger Monooxygenases for Asymmetric Catalysis. *J. Am. Chem. Soc.* **2009,** *131* (42), 15424-15432.

205. 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. *Advanced Synthesis and Catalysis* **2007**.

206. van Beek, H. L.; Gonzalo, G. d.; Fraaije, M. W., Blending Baeyer–Villiger monooxygenases: using a robust BVMO as a scaffold for creating chimeric enzymes with novel catalytic properties. *Chem. Commun.* **2012**.

207. Wu, S.; Acevedo, J. P.; Reetz, M. T., Induced allostery in the directed evolution of an enantioselective Baeyer-Villiger monooxygenase. *Proc Natl Acad Sci USA* **2010**.

208. Dudek, H. M.; de Gonzalo, G.; Torres Pazmiño, D. E.; Stępniak, P.; Wyrwicz, L. S.; Rychlewski, L.; Fraaije, M. W., Mapping the Substrate Binding Site of Phenylacetone Monooxygenase from Thermobifida fusca by Mutational Analysis. *Appl. Environ. Microbiol.* **2011,** *77* (16), 5730-5738.

209. 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.* **2015,** *112* (7), 1354-1364.

210. Balke, K.; Schmidt, S.; Genz, M.; Bornscheuer, U. T., Switching the Regioselectivity of a Cyclohexanone Monooxygenase toward (+)-trans-Dihydrocarvone by Rational Protein Design. *ACS Chem. Biol.* **2016,** *11* (1), 38-43.

211. Clouthier, C. M.; Kayser, M. M.; Reetz, M. T., Designing New Baeyer–Villiger Monooxygenases Using Restricted CASTing. *The Journal of Organic Chemistry* **2006,** *71* (22), 8431-8437.

212. Kirschner, A.; Bornscheuer, U. T., Directed evolution of a Baeyer–Villiger monooxygenase to enhance enantioselectivity. *Appl. Microbiol. Biotechnol.* **2008,** *81* (3), 465-472.

213. Balke, K.; Bäumgen, M.; Bornscheuer, U., Controlling the Regioselectivity of Baeyer-Villiger Monooxygenases by Mutation of Active Site Residues. *ChemBioChem* **2017**.

214. Chen, K.; Wu, S.; Zhu, L.; Zhang, C.; Xiang, W.; Deng, Z.; Ikeda, H.; Cane, D. E.; Zhu, D., Substitution of a Single Amino Acid Reverses the Regiospecificity of the Baeyer-Villiger

Monooxygenase PntE in the Biosynthesis of the Antibiotic Pentalenolactone. *Biochemistry* **2016**.

215. 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.

216. Li, G.; Fürst, M. J. L. J.; Mansouri, H. R.; Ressmann, A. K.; Ilie, A.; Rudroff, F.; Mihovilovic, M. D.; Fraaije, M. W.; Reetz, M. T., Manipulating the stereoselectivity of the thermostable Baeyer–Villiger monooxygenase TmCHMO by directed evolution. *Org. Biomol. Chem.* **2017**.

217. Romero, E.; Castellanos, J. R. G.; Mattevi, A.; Fraaije, M. W., Characterization and Crystal Structure of a Robust Cyclohexanone Monooxygenase. *Angewandte Chemie International Edition* **2016,** *55* (51), 15852-15855.

218. Fürst, M. J. L. 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**.

219. 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.

220. 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), 1-16.

221. Fink, M. J.; Rial, D. V.; Kapitanova, P.; Lengar, A.; Rehdorf, J.; Cheng, Q.; Rudroff, F.; Mihovilovic, M. D., Quantitative Comparison of Chiral Catalysts Selectivity and Performance:

A Generic Concept Illustrated with Cyclododecanone Monooxygenase as Baeyer-Villiger Biocatalyst. *Adv. Synth. Catal.* **2012,** *354* (18), 3491-3500.

222. Fink, M. J.; Fischer, T. C.; Rudroff, F.; Dudek, H.; Fraaije, M. W.; Mihovilovic, M. D., Extensive substrate profiling of cyclopentadecanone monooxygenase as Baeyer-Villiger biocatalyst reveals novel regiodivergent oxidations. *J. Mol. Catal. B: Enzym.* **2011,** *73* (1-4), 9-16.

Chapter 2: RApid Parallel Protein EvaluatoR (RAPPER), from gene to enzyme function in one day

Previously published in: Chem. Commun., 2015, 51, 122-124

Abstract

Cell-free transcription–translation systems offer an effective and versatile platform to explore the impact of genetic variations on protein function. We have developed a protocol for preparing linear, mutagenic DNA templates for direct use in the PURE system, enabling the fast and semi-quantitative evaluation of amino acid variations on catalytic activity and stereo-selectivity in native and engineered variants of Old Yellow Enzyme.

Results and Discussion

Tailoring biocatalysts by directed evolution or protein design is traditionally performed through manipulation of the corresponding nucleic acid sequence. The mutated gene is then transformed into a host organism for heterologous expression, followed by purification and functional evaluation via enzyme assay. Overall, the established process is robust but laborious and time-consuming. Cell-free transcription–translation systems offer an attractive alternative, not only eliminating the need for DNA transformation, culture growth and induction but also offering easy control over reaction conditions and the ability to supplement the reaction mixture.[1-2] The use of PURE (Protein synthesis Using Recombinant Elements), a chemically defined transcription–translation system first reported by Ueda and coworkers,[3-4] is particularly attractive as it minimizes side-reactions and reduces background activity. Previously, we have successfully employed PURE for the evaluation of circular permutation (CP) libraries of the NADPH-dependent flavoprotein Old Yellow Enzyme from Saccharomyces pastoranius (OYE1).[5] These studies relied on synthetic gene libraries, cloned in a circular DNA vector (pET-14b), as templates for protein synthesis in PURE. To further accelerate the discovery process for novel biocatalysts, we adapted our experimental protocol to directly use linear DNA products from polymerase chain reactions as templates for protein synthesis, hence allowing for assembly of (mutant) genes and their

subsequent biocatalytic assessment in one day. We herein report on the validation of our new protocol named RAPPER (RApid Parallel Protein EvaluatoR) with established OYEs, as well as an initial exploration of OYE variants for improved catalytic activity and altered stereo-selectivity.

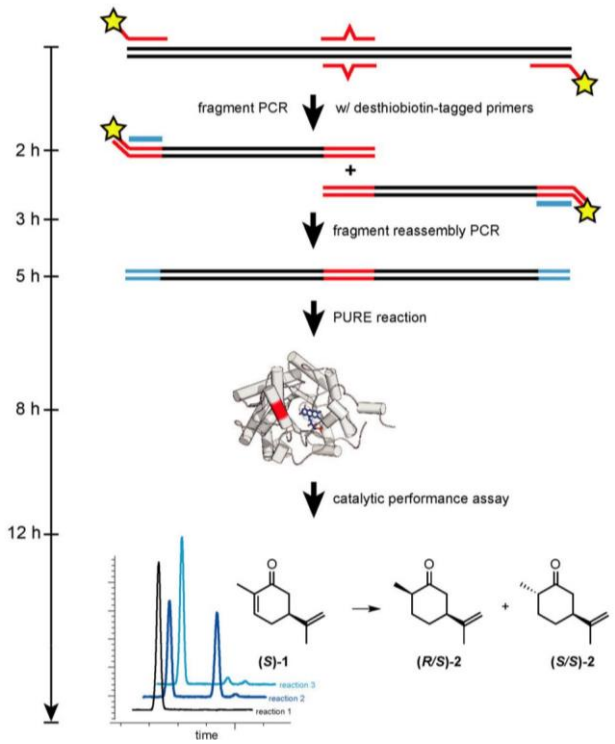


Figure 2.1 Overview of key steps (mutagenesis, protein synthesis, and enzyme assay) for the RAPPER protocol.

The overall strategy for RAPPER is outlined in Fig. 2.1 and consists of three generic steps: mutagenesis, protein synthesis, and enzyme assay. For our specific study of OYE1 variants, the initial step consisted of a primer overlap extension PCR[6] to introduce one or more mutations in the template DNA. Next, the modified DNA served as template for *in vitro* transcription– translation by PURE. Finally, the PURE reaction was mixed directly with a stock solution containing substrate, as well as necessary cofactors and buffer salts. Following a predetermined incubation time, the enzyme assay was quenched and product formation was detected by GC analysis.

To implement the three-step protocol of RAPPER, we first explored the effects of the regions flanking the gene of interest. Minimizing the size of up and downstream regions in the linear DNA template is desirable to simplify the PCR conditions, yet poses the risk of eliminating control elements for effective transcription and translation of the target protein. To probe for the shortest DNA template without diminishing effects in the subsequent enzyme activity assay, we prepared a series of five truncated DNA sequences based on the

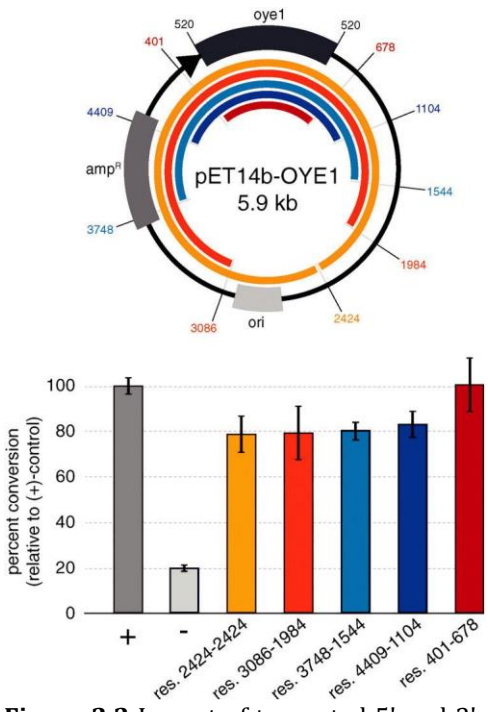


Figure 2.2 Impact of truncated 5' and 3'-regions in linear DNA templates on transcription-translation efficiency. (+) control: circular pET14b-OYE1; (-) control: pET14b-DHFR.

pET14b-OYE1 vector (Fig. 2.2). These linear DNA templates ranged in size from ~5.9 kb (res. 2424–2424) to ~1.5 kb (res. 401–678). The efficiency of protein synthesis with these linear DNA fragments by PURE was assessed *via* OYE reductase activity assay, measuring the percent conversion of (*S*)-carvone (**1**) to (*R*/*S*) and (*S*/*S*)-dihydrocarvone (**2**). Our results show that the enzyme activity for linear DNA templates are only slightly lower, yielding 80 to 90% conversion relative to circular DNA (+control). Furthermore, the data indicates no significant difference in the production of active enzyme as a function of DNA template size. These findings suggest no critical DNA sequence elements controlling protein synthesis beyond the roughly 150 base pairs flanking the gene of interest which encode for known elements including the RBS, as well as the T7 promoter and termination sites. In summary, linear DNA templates encoding minimal transcriptional and translational control elements express OYE at levels comparable to plasmid templates.

Next, we addressed compatibility issues between the mutagenesis step and PURE protocol. Dramatically fluctuating enzyme activity in replica experiments of individual OYE variants were noted in early experiments. Interestingly, the problem was not seen with native DNA templates or libraries generated by random mutagenesis (data not shown). The poor reproducibility was traced back to the agarose gel purification step in the primer overlap extension protocol. Gel purification of DNA fragments after the first PCR is necessary to remove the full-length DNA template, yet causes sample contamination with

inhibitors of translation (NEB, personal communication). The use of desthiobiotin-tagged primers offered a simple alternative for effectively removing the DNA template by washing the PCR products over immobilized streptavidin (Fig. 2.1). Quantitative recovery of the desthiobiotin-tagged DNA fragments was possible by rinsing with elution buffer containing biotin.

Following the adjustments to the RAPPER protocol, we prepared a series of OYE variants to validate the method and to explore the functional impact of a few specific amino acid substitutions in native OYE1 and cpOYE303. cpOYE303 is a circular permutated OYE variant whose new protein sequence termini are located in the loop β6 region near the active site.[5] Its enhanced ene-reductase activity is thought to derive from increased local backbone flexibility which improves active site accessibility and benefits rate-limiting conformational changes. Single and double mutations at amino acid positions 116, 250, 295, and 375 (all residue numbering based on OYE1) were introduced by primer overlap extension PCR. For validation purposes, aliquots of each PCR were directly processed by the PURE system, followed by detection of enzyme activity based on reduction of **1** (Fig. 2.3, "yellow" data). Separately, samples of the PCR products were also cloned into pET-14b and transformed into *E. coli* BL21(DE3) for heterologous expression. Following protein isolation, the purified enzymes were characterized by the same ene-reductase activity assay (Fig. 2.3, "blue" data). The catalytic conversion and diastereo-selectivity data for OYE variants based on RAPPER

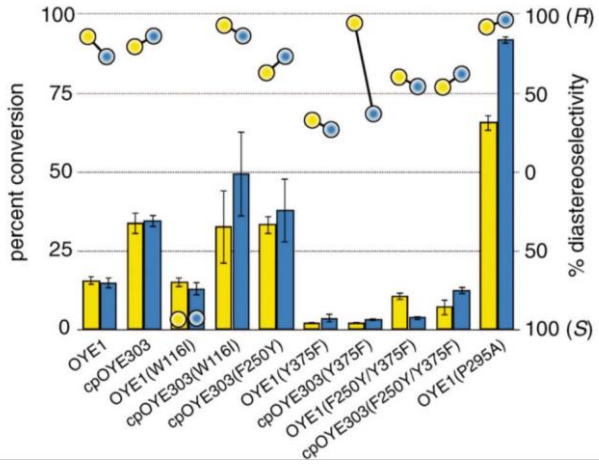


Figure 2.3 Comparison of activity (columns) and diastereo-selectivity (circles) data for reduction of **1** by selected OYE variants by RAPPER (yellow) and purified enzymes prepared by conventional methods using heterologous expression in *E. coli* host cells (blue).

and purified enzymes are summarized in Fig. 2.3. Over all, the results from the two enzyme activity assays are consistent, supporting the notion that RAPPER is a reliable predictor of catalytic performance for individual variants. Similarly, our cell-free screening system accurately captures the diastereo-selectivity. Measuring reduction of **1** to either (*R*/*S*)-**2** or (*S*/*S*)-**2**, the performance of OYE variants detected with RAPPER and the classical enzyme assay matches for all but one enzyme variant. We attribute the discrepancy in the data for cpOYE303(Y375F) to the enzyme's poor overall activity which results in low signal-to-noise and inaccurate peak integration.

Beyond validation experiments, RAPPER also allowed for an initial evaluation of specific amino acid changes on the catalytic performance and diastereo-selectivity of OYE1 and cpOYE303. We first tested the effects of W116I on the biocatalytic performance of cpOYE303. In previous work, Stewart and coworkers showed a complete reversion of OYE1 diastereo-selectivity for the reduction of **1**.[8] While our assay accurately reproduced Stewart's results in OYE1, the same W116I change in cpOYE303 did not result in any detectable diastereo-selectivity change (Fig. 2.3). Crystallographic data for cpOYE303 suggest that slight conformational changes at Y375 might explain the selectivity shift.[9] These findings also raised questions as to the possible role of neighboring residues in the active site and in the access tunnel, which can affect the pre-orientation of substrate. Hence we evaluated the effects of amino acid changes in positions F250 and Y375 (Fig. 2.4). Two single-site variants (F250Y and Y375F) and the double mutant (F250Y/Y375F) were prepared in OYE1 and cpOYE303, respectively.

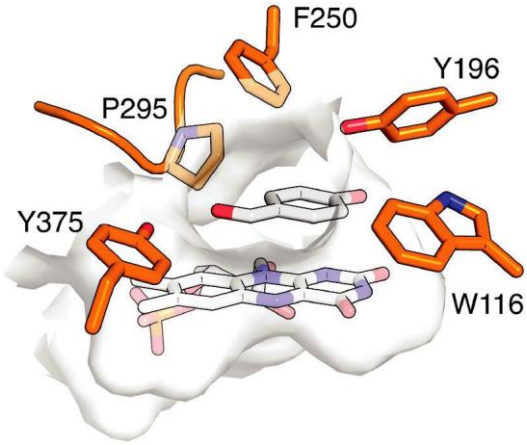


Figure 2.4 Active site of OYE1 (PDB 1K03)[7] with bound FMN and *p*-hydroxybenzaldehyde. Key active site residues are highlighted.

While the F250Y substitution showed little effect on activity and selectivity, Y375F proved highly detrimental to enzyme function. In contrast, combining the two amino acid changes resulted in recovery of activity at diminished diastereo-selectivity. Our initial screening data encourage the investigation of additional amino acid substitutions in these two positions to search for enzymes with altered selectivity without loss of catalytic activity.

In a final experiment, we used RAPPER to test an idea regarding the role of protein backbone flexibility in the loop β6 region of OYE1. Loop β6 forms a lid over the active site and is thought to be involved in the rate-limiting conformational change during biocatalysis. Consistent with this hypothesis, termini relocation into this region of the protein structure by circular permutation, which increases local conformational flexibility, resulted in over ten-fold improved catalytic activity for reduction of **1**.[5] An alternative strategy for probing the functional role of loop β6 is the substitution of alanine for proline 295, located at the tip of the loop. The amino acid change was predicted to reduce the region's conformational constrains, in turn translating into enhanced catalytic performance. Activity measurements in RAPPER and control experiments with the purified enzyme indicate significant functional gains that exceed even the increased activity observed for cpOYE303.

The modification of the primer overlap extension PCR protocol allows for direct use of linear DNA product as template in the PURE *in vitro* transcription–translation system. Such an adaptation allows for the rapid functional evaluation of one or more amino acid substitutions in target proteins. Our validation experiments with OYE variants demonstrate that the results from this approach are semi-quantitative in regard to catalytic activity and stereo-selectivity, hence presenting an effective method for functional prescreening of enzyme variants. RAPPER should be broadly applicable for proteins usually expressed in *E. coli* hosts. The method is also scalable for parallel screening of protein libraries in 96-well microtiter plate format. We consider RAPPER a useful tool to rapidly assess the function of

enzyme variants without the need for lengthy procedures of cloning, transformation, culture growth and expression.

Materials and Methods

Materials: PURExpress kits were purchased from New England Biolabs (Ipswich, MA). *Pfu* DNA polymerase (Stratagene, La Jolla, CA) in combination with the manufacturer's buffer was used for PCR amplification. All other enzymes were purchased from New England Biolabs. Oligonucleotide primers were ordered from Integrated DNA Technologies (Coralville, IA). Reagents, substrates, and reference materials were obtained from Sigma-Aldrich (St. Louis, MO) unless otherwise indicated.

Preparation of linear DNA templates: Linear DNA sequences as templates for *in vitro* transcription- translation were prepared by PCR amplification using site-specific oligonucleotide primers (listed below) and pET14b-OYE1 as template.

res. 2424-2424:	for: 5'-GCTGCATGTGTCAGAGGTTTTCAC-3'
	rev: 5'-GGAGACGGTCACAGCTTGTCTGT-3'
res. 3086-1984:	for: 5'-ATGCAGATCCGGAACATAATGGTGCAG-3'
	rev: 5'-AGCTGGGCTGTGTGCACGAACC-3'
res. 3748-1544:	for: 5'-GCGGGGAGGCAGACAAGGTATAGGG-3'
	rev: 5'-CCACGCTCACCGGCTCCAGATTTA-3'
res. 4409-1104:	for: 5'-GCCGAAAATGACCCAGAGCGCTG-3'
	rev: 5'-GGCAAAATGCCGCAAAAAAGGGAAT-3'
res. 401-678:	for: 5'-CTCGATCCCGCGAAATTAATACGACT-3' (T7 promoter)
	rev: 5'-CAGCAAAAAACCCCTCAAGACCCG-3' (T7 terminator)

Mutagenesis by primer overlap extension PCR: Mutations in specific locations of *oye1* were introduced by primer overlap extension PCR, using either pET14b-OYE1 or pET14b-cpOYE303 (5) as templates. To facilitate PCR product purification without agarose gel electrophoresis, the flanking T7 promoter and T7 terminator primers (see sequence above) were modified at the 5'-end by a desthiobiotin moiety on a PEG linker. These tagged primers were used in combination with the appropriate mutagenic primers listed below for the preparation of individual DNA fragments by PCR amplification (Fig. 2.1).

OYE_W116I: for: 5'-GGGTTCAGTTAATAGTTTTGGGTTG-3'

rev: 5'-CAACCCAAAACTATTAACTGAACCC-3'

OYE_F250Y: for: 5'-CCCCATACGGTGTTTACAACAGTATGTCTG-3'

rev: 5'-CAGACATACTGTTGTAAACACCGTATGGGG-3'

OYE_P295A: for: 5'-CGTGTAACTAACGCATTCTTGACTGAAG-3'

rev: 5'-CTTCAGTCAAGAATGCGTTAGTTACACG-3'

OYE_Y375F: for: 5'-GACAGAGATACTTTCTTCCAGATGTCTGCTC-3'

rev: 5'-GAGCAGACATCTGGAAGAAAGTATCTCTGTC-3'

For DNA fragment purification, the 50 µl-PCR reactions were diluted with 50 µl water and mixed with high-capacity streptavidin agarose beads (Pierce, Rockford, IL) in binding buffer (5 mM Tris-HCl (pH 7.5), 0.5 mM EDTA, 1 mM NaCl). After incubation on a rocking platform at 21 ˚C for 15 min, the suspension was centrifuged (8000g, 2 min) and the supernatant discarded. The streptavidin beads were washed with 3 x 100 µL of binding buffer, followed by incubation with 50 µL elution buffer (4 mM biotin, 10 mM Tris-HCl (pH 8.5)) at 21 ˚C for 10 min. Purified DNA fragments were recovered in clear supernatant after sample centrifugation, quantified spectrophotometrically by absorbance at 260 nm and stored at 4 ˚C. For recombination of DNA fragments into the full-length *oye1* genes,

approximately 40 ng of each fragment were used as templates for PCR amplification with standard T7 forward and reverse primers. PCR products were purified with QIAquick PCR Purification Kit (Qiagen, Valencia, CA).

Cloning of oye1 variants: As part of the RAPPER validation experiments, a separate aliquot of linear DNA products used with PURExpress was also cloned into pET-14b for DNA sequencing and traditional heterologous protein expression in an *E. coli* host organism. Briefly, full-length *oye1* genes obtained by primer overlap extension PCR were digested with *XhoI* and *XbaI*, followed by T4 DNA ligase catalyzed insertion into pET-14b, linearized with the same two restriction endonucleases. Ligation products were transformed into E. coli host and processed for protein overexpression and purification as previously described [5].

PURExpress In Vitro Transcription-Translation: Full-length *oye1* genes were used as linear DNA templates for *in vitro* transcription-translation (IVTT). Briefly, IVTT reactions were assembled using the PURExpress *in vitro* protein synthesis kit (PURE) following manufactures protocol with a few adjustments (see below) optimized for our application. Reactions were assembled on a 10-µL scale containing PURExpress solution A, PURExpress solution B, 100 µM FMN, 10 units Murine RNase Inhibitor (NEB), 100 ng DNA template, and nuclease-free H2O. All reactions were run simultaneously with native OYE1 (positive control) and dihydrofolate reductase (DHFR; negative control; provided by NEB) under the same experimental conditions. The reaction mixtures were incubated at 37°C for 2.5 hours to allow for protein synthesis, followed by cooling to 4 ˚C.

Enzyme activity assay: The ene-reductase activity of enzymes produced by IVTT or traditional heterologous expression/purification (referred to as “purified enzyme”) was

measured under anaerobic conditions (Coy Laboratory Products, Grass Lake, MI) at ambient temperature using glucose dehydrogenase (GDH) from *Thermoplasma acidophilum* for NADPH regeneration. A reaction stock solution was prepared, containing 10 mM (*S*)-carvone (**1**), 200 µM NADP+, GDH (2 or 5 units for IVTT or purified enzyme reactions, respectively), 100 mM glucose in 50 mM Tris-HCl (pH 7.5).

For activity assays of IVTT reactions, 20 µL of reaction stock was added directly to the IVTT reaction mixture (10 µL). To assay purified enzyme, the OYE1 variant (final concentration: 250 nM) was added to 500 µL of reaction stock solution. After 4 h reaction time, a 30-µl aliquot of the assay solution was quenched by mixing thoroughly with 30 µL of ethyl acetate containing 1 mM cyclohexanone as internal standard. A sample of the organic phase was injected onto an Agilent Technologies 6850 GC instrument equipped with a chiral CycloSil-B column (30 m x 0.32 mm / 0.25 µm, Agilent, Santa Clara, CA) using hydrogen as a carrier gas (flow rate 1.8 mL/min) and an FID detector (detector temperature 200 ˚C, split ratio 25:1). The temperature program for the GC was as follows: 90 ˚C, hold 5 min, then 1 ˚C/min to 120 ˚C (retention time: **1** = 27.35 min, (1R,4S)-**2** = 22.9 min, (1S,4S)-**2** = 22.6 min). The percent conversions and diasteriomeric excess was calculated from substrate and product integration areas and were quantified using standard curves generated using known amounts of the substrate and product.

Bibliography

1. Shimizu, Y.; Kuruma, Y.; Ying, B.-W.; Umekage, S.; Ueda, T., Cell-free translation systems for protein engineering. *FEBS J.* **2006,** *273* (18), 4133-4140.

2. Carlson, E. D.; Gan, R.; Hodgman, C. E.; Jewett, M. C., Cell-free protein synthesis: Applications come of age. *Biotechnol. Adv.* **2012,** *30* (5), 1185-1194.

3. Shimizu, Y.; Inoue, A.; Tomari, Y.; Suzuki, T.; Yokogawa, T.; Nishikawa, K.; Ueda, T., Cell-free translation reconstituted with purified components. *Nat Biotech* **2001,** *19* (8), 751-755.

4. Ying, B.-W.; Ueda, T., Protein Generation Using a Reconstituted System. In *Protein Engineering Handbook*, Wiley-VCH Verlag GmbH & Co. KGaA: 2008; pp 515-535.

5. Daugherty, A. B.; Govindarajan, S.; Lutz, S., Improved biocatalysts from a synthetic circular permutation library of the flavin-dependent oxidoreductase old yellow enzyme. *J. Am. Chem. Soc.* **2013,** *135* (38), 14425-32.

6. Horton, R. M.; Cai, Z.; Ho, S. N.; Pease, L. R., Gene Splicing by Overlap Extension: Tailor-Made Genes Using the Polymerase Chain Reaction. *BioTechniques* **1990,** *8* (5), 528-535.

7. Brown, B. J.; Hyun, J. W.; Duvvuri, S.; Karplus, P. A.; Massey, V., The role of glutamine 114 in old yellow enzyme. *J Biol Chem* **2002,** *277* (3), 2138-45.

8. Padhi, S. K.; Bougioukou, D.; Stewart, J. D., Site-Saturation Mutagenesis of Tryptophan 116 of *Saccharomyces pastorianus* Old Yellow Enzyme Uncovers Stereocomplementary Variants. *J. Am. Chem. Soc.* **2009,** *131*, 3271-3280.

9. Daugherty, A. B.; Horton, J. R.; Cheng, X.; Lutz, S., Structural and Functional Consequences of Circular Permutation on the Active Site of Old Yellow Enzyme. *ACS Catalysis* **2014**, 892-899.

Chapter 3: Cell-free protein engineering of Old Yellow Enzyme 1 from *Saccharomyces pastorianus*

Previously published in: Tetrahedron, 2016, 72, 7282-7287

Abstract

In protein engineering, cell-free transcription/translation of linear mutagenic DNA templates can tremendously accelerate and simplify the screening of enzyme variants. Using the RApid Parallel Protein EvaluatoR (RAPPER) protocol, we have evaluated the impact of amino acid substitutions and loop truncations on substrate specificity and stereoselectivity of Old Yellow Enzyme 1 from *Saccharomyces pastorianus*. Our study demonstrates the benefit of systematically assessing amino acid variations including substrate profiling to explore sequence-function space.

Introduction

Enzymes are highly efficient (bio)catalysts for asymmetric synthesis of enantiomerically pure compounds, typically displaying high enantio, regio and chemo-selectivity and performing under benign environmental conditions.[1] Among the variety of biocatalytic conversions, the asymmetric reduction of alkenes by members of the Old Yellow Enzyme (OYE) family is widely used to generate up to two stereogenic centers. OYEs are flavin mononucleotide (FMN)-containing NADPH-dependent oxidoreductases. Mechanistically, the reaction proceeds via a Michael-type hydride transfer step from the reduced flavin cofactor to the substrate, followed by stereospecific protonation to yield the trans-product (Fig. 3.1). The FMN is then reduced by NADPH, regenerating the flavin cofactor and completing the catalytic cycle.

Contributing to their popularity in biotechnology and the pharmaceutical industry, OYEs tolerate a wide range of α,β-unsaturated ketones, aldehydes and carboxylic acid derivatives, as well as nitroalkenes, nitriles, and nitroaromatics as substrates.[3-11] Nevertheless, the need for increased catalytic activity and stereo-selectivity of OYEs for specific substrates have inspired protein engineers to improve upon the natural diversity of these biocatalysts. A number of structure-guided site-directed and site-saturation mutagenesis studies, as well as directed evolution experiments have yielded OYE variants with improved catalytic properties.[12-20] These engineering efforts also identified several experimental challenges and limitations related to heterologous expression and test substrates and have led us to explore new strategies for evaluating libraries of OYE variants from *Saccharomyces pastorianus* (OYE1).[21] Specifically, we switched to *in vitro* transcription/translation using the chemically-defined 'protein synthesis using recombinant elements' (PURE) system.[22] Following initial studies in PURE to eliminate cellular background reaction(s) and accelerate library screening, we continued to develop the cell-free concept to include the use of linear DNA as template for protein synthesis, enabling direct use of PCR products for functional evaluation of the

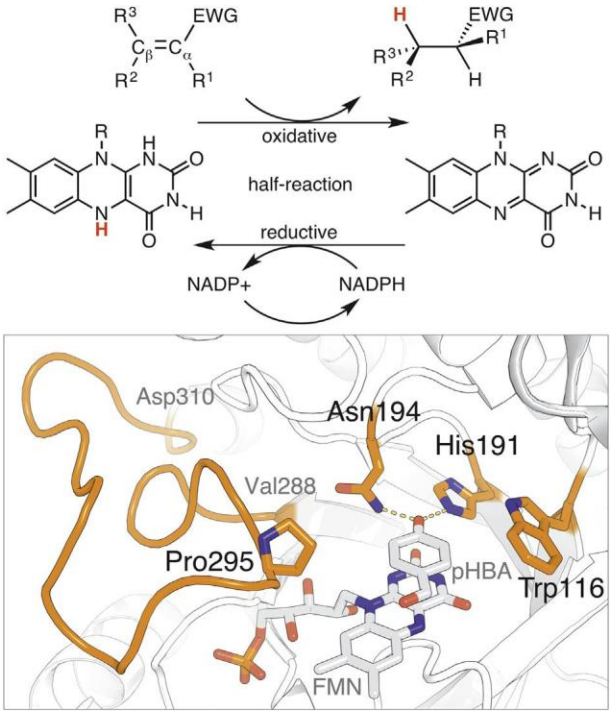


Figure 3.1 OYE-catalyzed trans-hydrogenation of alkenes activated by an electron-withdrawing group (EWG). The asymmetric reduction of substrate via hydride transfer (shown in red) as part of the oxidative half-reaction (in reference to flavin cofactor) is coupled with cofactor regeneration by NADPH. The crystal structure of OYE1 with substrate analog *p*-hydroxybenzaldehyde (pHBA) bound in the active site (PDB: 1K03)[2] shows the substrate's hydrogen bonding interactions with H191 and N194. The positions of W116 and P295, as well as the loop β6 region (V288-D310) are highlighted.

corresponding native and engineered enzymes through a protocol named RAPPER (RApid Parallel Protein EvaluatoR).[23]

We have now started to apply RAPPER for systematic protein engineering studies of OYE1, as well as to expand the RAPPER protocol to include multi-site mutagenesis, InDels variations, and substrate profiling. Besides validating the impact of amino acid replacements at position W116 in OYE1 on catalytic activity and stereoselectivity, we are employing RAPPER to study the functional contributions of the loop β6 region in OYE1 by alanine-scanning and site-specific mutagenesis, as well as loop truncation. Finally, activity data for individual amino acid changes of two key active site residues at positions H191 and N194 were measured.

Results and Discussion

A first set of systematic amino acid replacements using RAPPER focused on position W116 in OYE1. The tryptophan residue lines the active site binding pocket and substitutions at this position have been shown to be highly effective in reversing the enzyme's diastereoselectivity for reduction of (*S*)-carvone.[13, 18] To validate these findings in RAPPER, we employed the previously reported two-step fragment overlap extension PCR to create linear template DNA followed by *in vitro* transcription/translation to prepare five W116x variants (x=A, V, I, N, and F).[23] Cell-free protein synthesis was performed at 10-µL reaction scale to minimize sample evaporation and reduce variability in activity assays. However, protein expression levels were sufficiently high to split the reaction mixture after protein synthesis into 2-µL aliquots for the subsequent assessment of catalytic activity, significantly broadening and accelerating enzyme characterization by allowing for parallel testing of up to four substrates. As summarized in Fig. 3.2, each W116x variant could be evaluated for reduction of (*S*)-carvone, methyl-2-(hydroxymethyl) acrylate, geranial, and neral in a single experiment. In comparison to wild type OYE1, the rates of conversion for selected variants

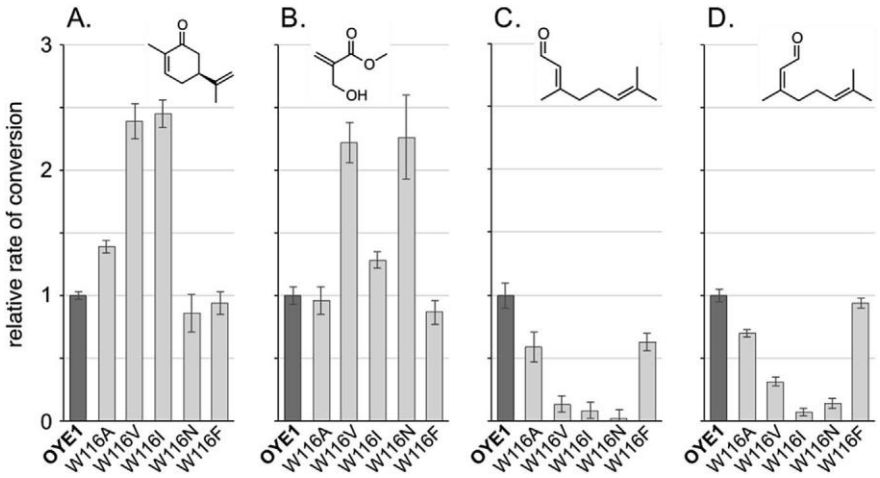


Figure 3.2 Rate of conversion for selected W116 variants of OYE1 relative to wild type enzyme. a) (*S*)-carvone, b) methyl-2-(hydroxymethyl)acrylate, c) geranial, and d) neral by RAPPER. Full experimental data including observed changes in stereoelectivity are summarized in Table 3.S1. The results were consistent with previously reported data.[18, 20]

showed greater than two-fold improvements for (*S*)-carvone and methyl-2-(hydroxymethyl)acrylate, respectively, while activity for geranial and neral was generally lower. Our routine analysis of assay products by chiral GC further confirmed the reversed stereoselectivity of selected W116x variants for (*S*)-carvone, consistent with previous literature reports.[18] Overall, these single-time point measurements do not provide sufficient information to rationalize the underlying cause for the performance differences of our W116x variants with these substrates. Instead, they offer a rapid approach to establish substrate profiles for native and engineered enzymes, helping to identify lead candidates for further in-depth characterization and exploring the generality of amino acid changes in engineered biocatalysts.

A second region of interest in OYE1 is loop β6. The extended loop region stretches from amino acid residues 288 to 310, constituting a lid that caps the active site binding pocket (Fig. 3.1). Previous circular permutation studies have shown significant catalytic rate enhancements upon cleavage of peptide linkages in this loop.[21, 24] These functional changes were explained through changes in the loop's conformational flexibility. To further probe

the role of loop β6 on enzyme function, we conducted alanine-scanning mutagenesis of the entire region and measured its impact on (*S*)-carvone reduction. Using the RAPPER protocol enabled parallel processing of all 24 samples (wild type plus 23 variants) in triplicate. The subsequent activity assay showed variable positional effects of alanine substitutions on enzyme function (Fig. 3.3). By far the most significant improvement resulted from replacement of proline 295, located at the very tip of the loop region. Besides impacting the neighboring F296 which lines the active site and undergoes a conformational change upon substrate binding,[18, 24] the release of conformational restraints upon alanine substitution of the cyclic proline could explain the >3-fold activity gain. Subsequent time-course studies with purified P295A variant confirmed the observed functional gains (Fig. 3.S1). Consistent with our interpretation of greater conformational flexibility of the entire loop β6 region as a consequence of the P295A substitution, Tanokura and co-workers recently reported similar observations in their studies of Old Yellow Enzyme from *Candida macedoniensis*.[25] The equivalent loop region in *C. macedoniensis* OYE shows significant movement upon crystallization in the presence or absence of substrate inhibitor *p*-hydroxybenzaldehyde and introduction of glycine at P295 resulted in rate increases for ketoisophorone reduction. These observations naturally raised the question whether P295A is the most beneficial amino acid change in OYE1 or whether alternative side chains could further modulate enzyme activity.

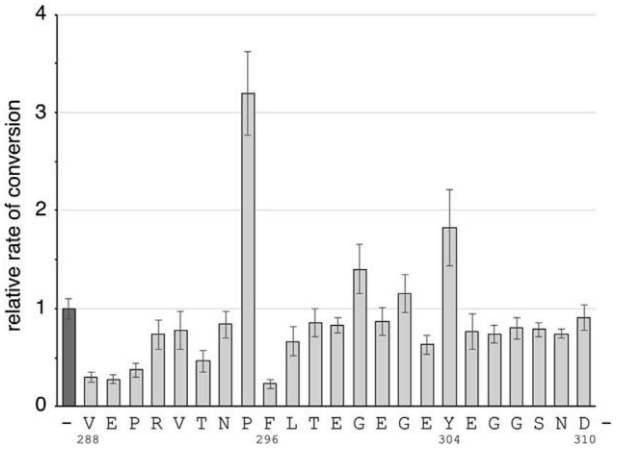


Figure 3.3 Relative rates of conversion for (*S*)-carvone by OYE1 variants generated via alanine-scanning mutagenesis of residues in loop β6 region (positions 288-310) by RAPPER. Full experimental data including observed changes in stereoelectivity are summarized in Table 3.S2.

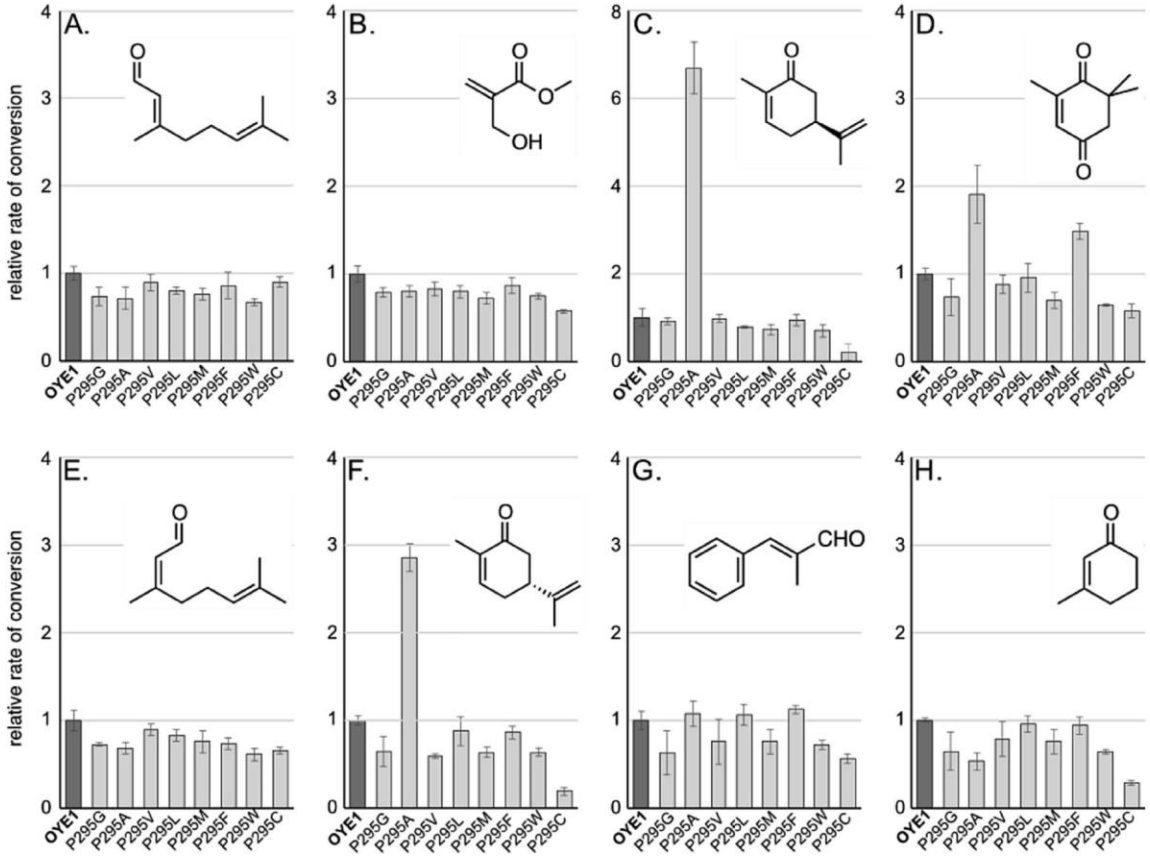


Figure 3.4 Substrate profiles for selected P295 variants of OYE1 by RAPPER. Relative rates of conversion were based on wild type enzyme activity for the respective substrate. a) geranial, b) methyl-2-(hydroxymethyl)acrylate, c) (*S*)-carvone, d) ketoisophorone, e) neral, f) (*R*)-carvone, g) α-methyl-trans-cinnamaldehyde and h) 3-methyl-2-cyclohexenone. Full experimental data are summarized in Table 3.S3.

To further study the functional role of P295, we prepared a small, secondary library of P295x variants (x=G, A, V, L, M, F, W, and C) to explore the impact of hydrophobic side chain size at this position. As for the previous W116x library, we employed RAPPER and screened all library members for activity with multiple substrates (Fig. 3.4). Besides (*S*)-carvone, methyl-2-(hydroxymethyl) acrylate, geranial, and neral, we doubled the set of test substrates by adding (*R*)-carvone, ketoisophorone, α-methyl-trans-cinnamaldehyde and 3-methyl-2-cyclohexenone. This expanded set substrates could again be tested in parallel by doubling the volume of cell-free protein synthesis to 20-µL and subsequently splitting the reaction mixture into 8 2-µL activity assays. Our findings confirm the substantive functional gains of OYE variant P295A of three and seven-fold for (*R*) and (*S*)-carvone, respectively.

Two-fold activity gains were also observed with ketoisophorone. Interestingly, the anticipated functional gains in variant P295G due to even greater conformational flexibility did not materialize. Furthermore, the functional benefits of P295A seem to only apply to a subset of substrates (Fig. 3.4 C, D, F). For the other five test substrates (Fig. 3.4 A, B, E, G, H), none of the substitutions in our P295x library seem to have any significant functional consequence. The distinct behavior of OYE variants with different substrates could be explained by variations in the enzymes' rate-determining step. Our results are consistent with conformational changes in loop β6 during alkene substrate binding as the rate-limiting step for overall turnover of carvones and ketoisophorone (Fig. 3.4 C, D, F). In contrast, changes to the loop dynamics seem irrelevant to enzyme-substrate pairs such as geranial, neral, and methyl-2-(hydroxymethyl)acrylate where other steps along the reaction coordinate are slower than the conformational changes in the loop β6 region. While testing this hypothesis requires extensive steady-state and rapid enzyme kinetic measurements (Orru & Lutz, unpublished results), these findings illustrate the potentially highly specific effects of individual amino acid substitutions in connection with particular substrates. Furthermore, they stress the importance of evaluating enzyme variants with multiple substrates to assess the broader impact of amino acid substitutions.

Beyond exploring the functional consequences of amino acid replacements in enzymes, the impact of amino acid insertions and deletions can offer an additional dimension in protein engineering.[26-27] Such a strategy is particularly attractive for altering loop size and structure, hence we prepared a small set of OYE1 variants with deletions of up to five amino acid residues in loop β6. Centered around position F296 at the loop tip, the RAPPER protocol was used to generate three variants; Δ1 (F296 deletion), Δ3 (F296-T298 deletion) and Δ5 (P295-E299 deletion) and functional changes were assayed on four substrates (Fig. 3.5). Our results unfortunately showed little or no significant benefit of

such loop changes. The rates of reduction of the three deletion variants for (*S*)-carvone, geranial, and neral are consistently slightly lower than native OYE1 (Fig. 5 A, C, D). For methyl-2-(hydroxymethyl)acrylate (Fig. 3.5B), variant Δ1 shows a rather dramatic decline in activity, yet the removal of additional residues in Δ3 and Δ5 largely restores catalytic activity to levels seen for the other substrates.

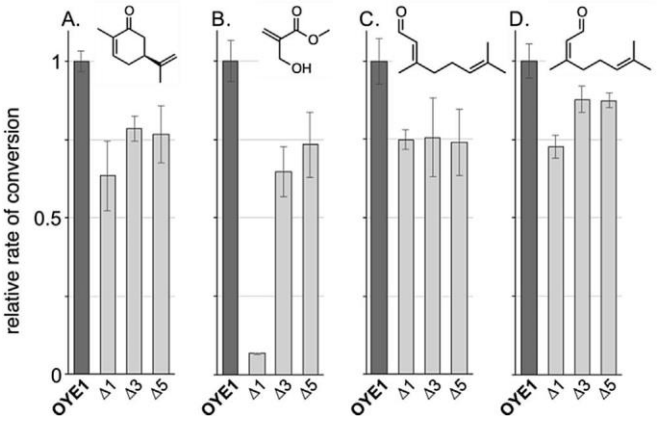


Figure 3.5 Impact of deletions in loop β6 region of OYE1 on catalytic activity. Changes in relative rates of conversion upon elimination of residues F296 (Δ1), F296-T298 (Δ3), or P295-E299 (Δ5) were assessed on four substrates: a) (*S*)-carvone, b) methyl-2-(hydroxymethyl)acrylate, c) geranial, and d) neral by RAPPER.

Finally, we applied RAPPER to explore amino acid substitution effects at positions H191 and N194, two key active site residues in OYE1. The side chains of these two amino acids typically form hydrogen bonding interactions with the substrate's electron-withdrawing group, facilitating a shift in electron density at the neighboring alkene, which activates its β-position for hydride transfer from the nearby reduced flavin cofactor (Fig. 3.1).[28] Early studies by Massey and co-workers successfully swapped amino acids at positions 191 and 194, which proved detrimental to catalytic activity.[29] More recently, the Bommarius group reported on the ene-reductase activity (or more precisely the absence thereof) of an H191A/N194A variant of *Kluyveromyces lactis* OYE, a close homolog of OYE1.[30] Independent of these results, we screened two single-site OYE1 libraries (H191x where x=A, L, Q, S, C, N or T; and N194x where x=A, L, Q, S, C, H or T). The chosen amino acid substitutions enable broader sampling of size and hydrogen bonding capacity on enzyme function. The catalytic activities for these variants were assessed with (*S*)-carvone and methyl-2-(hydroxymethyl)acrylate (Fig. 3.6). For the reduction of (*S*)-carvone, individual

amino acid replacements at H191x proved highly detrimental (Fig. 3.6A). Except for residual activity levels with H191A, all tested substitutions at this position abolished catalytic activity. The same reaction seems tolerant of substitutions at N194 (Fig. 3.6B). Although activity levels dropped in all variants, declines varied from minor (N194H) to moderate (N194S, Q) to major (N194A, T, C, L). These findings clearly suggest for H191 to play a key functional role in (*S*)-carvone binding and/or electronic activation and for N194 to serve in its support. Separately, we evaluated the single-variant libraries with methyl-2-(hydroxymethyl)acrylate (Fig. 6C,D) The impact of H191 and N194 variations on conversion of this second substrate differ significantly from the results obtained for (*S*)-carvone. All but one member of the two libraries (N194H) retain conversion rates of >25% relative to OYE1, suggesting that neither of these positions plays the same critical functional role. More importantly, our results demonstrate the highly substrate-dependent effects of amino acid changes in OYE1. More detailed experimental studies with these alanine variants to explore the structural and functional consequences of these substitutions are in progress.

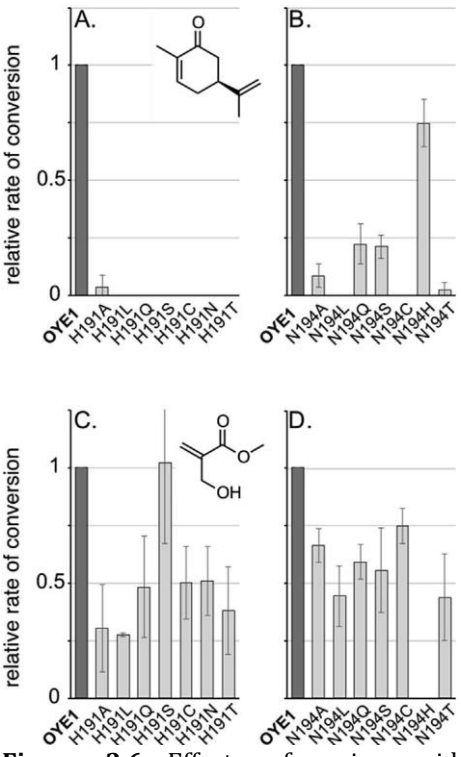


Figure 3.6 Effects of amino acid replacements of H191 and N194, key residues in the OYE1 active site, on catalytic activity. Relative rates of conversion were based on wild type enzyme activity for two substrates: a-b) (*S*)-carvone and c-d) methyl-2-(hydroxymethyl)acrylate by RAPPER. Full experimental data are summarized in Table 3.S4.

Conclusion

Protein engineering of oxidoreductases in the Old Yellow Enzyme family continue to be of great interest for developing novel biocatalysts. Our present study used small, focused

libraries to systematically probe the functional consequences of specific residues and regions in OYE1. In contrast to similar previous studies, the application of RAPPER enables faster and more efficient, semi-quantitative evaluation of enzyme variants. Beyond changes to individual amino acid positions, fragment overlap extension PCR has successfully been applied to generate gene templates for exploring functional consequences of residual InDels (insertions/ deletions). Furthermore, RAPPER's capacity to support parallel screening of engineered biocatalysts with multiple substrates enables a more comprehensive assessment of impact and contributions of individual amino acid substitutions. Our present results for OYE1 highlights the potential benefits of substrate profiling in protein engineering studies, averting the risk of misleading results from usage of a model substrate. Functional gains of amino acid replacements such as P295A or changes in stereoselectivity observed for W116I show significant substrate dependence. Similar to other high-throughput screening platforms, RAPPER does have limitations due to the small sample sizes. Protein yields for enzyme variants from 10-µL PURE reactions are too low to assess questions of expression level, foldability and stability of individual candidates in detail. Furthermore, amino acid changes may result in altered inhibitory properties and substrate binding affinity which can affect the outcome of the screening assay and complicate direct comparison of variants. Lead candidates are therefore routinely evaluated in follow-up studies using traditional biochemical and biophysical methods and these results have validated the semi-quantitative nature of RAPPER data, making the protocol useful for initial screening of enzyme libraries. In summary, the use of cell-free systems offers a simple, versatile and very effective tool to accelerate the discovery process for tailored biocatalysts.

Experimental

General information

The PURExpress® kit was purchased from New England Biolabs (Ipswitch, MA). All primers used were synthesized by Integrated DNA Technologies (Coralville, IA). All other reagents were purchased from Sigma Aldrich (St. Louis, MO) unless otherwise specified. GC spectra were obtained on an Agilent Technologies 6850 GC instrument equipped with a chiral CycloSil-B column (30 m 0.32 mm/0.25 μm, Agilent, Santa Clara, CA) using hydrogen as a carrier gas (flow rate 1.8 mL/min) and an FID detector (detector temperature 200 °C, split ratio 25:1).

Creating linear templates of OYE variants

All mutations were prepared by a two-step fragment overlap extension PCR protocol using pET14b-OYE1 as a template.[21, 23] Mutagenic primer pairs are listed in Table 3.S1 and were used in combination with universal flanking primers carrying a 5'-desthiobiotin-tag for gel-free fragment purification (forward: 5'-desthiobiotin-CTCGATCCCGCGAAATTAATACGACT-3'; reverse: 5'-desthiobiotin-CAGCAAAAAACCCCTCAAGACCCG-3'). Following the first PCR, reaction mixtures were diluted with an equal volume of distilled water and added to high-capacity streptavidin agarose beads (Pierce, Grand Island, NY) pre-equilibrated in binding buffer (5 mM Tris-HCl (pH 7.5), 0.5 mM EDTA, 1 mM NaCl). After incubation at 21 °C for 15 min, the mixture was centrifuged at 13,000 rpm for 3 min and the supernatant discarded. The streptavidin-bound fragments were washed three times with 100 μL binding buffer, followed by incubation in 50 μL elution buffer (4 mM biotin, 10 mM Tris-HCl, pH 8.5) at 21 °C for 10 min. The mixture was centrifuged at 13,000 rpm for 3 min and the clear supernatant containing fragment DNA was collected. For the second PCR to reassemble fragments into the full-length genes, aliquots of the fragment DNA were mixed with universal flanking primers (same sequence

as for the first PCR but without 5'-tags). Full-length PCR products were purified using the QIAgenPCR Purification kits (Qiagen, Valencia, CA) and quantified by UV absorbance spectroscopy at 260 nm.

In vitro transcription/translation of linear DNA

Linear DNA (PCR products) were used as templates for protein expression of selected OYE variants to assess their catalytic activities as described previously.[23] Briefly, protein synthesis was performed by *in vitro* transcription/translation (IVTT), using the PURExpress system and reaction products were directly tested for ene-reductase activity with selected substrates of interest. With each set of experiments, wild type OYE1 was run as a positive control while dihydrofolate reductase (provided by NEB) served as negative control. Typically, protein synthesis was terminated after 2.5 h incubation at 37 °C by storing the mixture on ice.

Large-scale protein expression and purification

To obtain larger quantities of purified OYE variants, PCR products were cloned into pET-14b (Novagen) via *NcoI*/*XhoI* restriction sites and the plasmid DNA transformed into *E. coli* BL21(DE3) pLysS cells. Individual colonies were grown overnight in 5 mL of 2YT media containing ampicillin (100 µg/mL) at 37 °C. An aliquot of the overnight culture was then used to inoculate 250 mL of 2YT containing ampicillin (100 µg/mL). Following incubation at 37 °C until the culture reached an OD(600) of 0.7, heterologous protein expression was induced by addition of IPTG (final concentration: 0.4 mM) for 18 h at 20 °C. Cells were pelleted by centrifugation at 4 °C, 4000 g for 20 min, the clear supernatant removed and pellets stored at -20 °C until further usage.

For protein purification, a cell pellet was resuspended in 24 mL of buffer A (40 mM Tris-HCl (pH 8.0), 20 mM NaCl) and mixed with 200 µL protease inhibitor cocktail (Sigma) and 20 µL of benzonase (Novagen). After 30 min storage on ice, cells were lysed by sonication. The suspension was centrifuged at 4 °C, 10,000 g for 30 min and the clear supernatant was loaded on a HiTrap Q FF (5 mL) anion-exchange column, pre-equilibrated in buffer A. The resin was washed with two column volumes (CVs) buffer A, followed by a linear gradient over 10 CVs to 100% buffer B (40 mM Tris-HCl (pH 8.0), 1 M NaCl). Protein elusion was monitored by UV absorption (280/460 nm) and product fractions were collected, followed by concentration and buffer-exchange into buffer C (40 mM Tris-HCl (pH 8.0), 300 mM NaCl) using a Millipore filter unit (MWCO: 10 kDa). The protein sample was further purified via size exclusion chromatography (Superdex 200, 10/300 GL column) in buffer C. Product fractions were pooled, concentrated as described above, and analyzed by SDS-PAGE.

Enzymatic activity assay for OYE1 variants

Ene-reductase activity assays were performed at ambient temperature under anaerobic conditions (Coy Laboratory, Grass Lake, MI). For preparing reaction stock solution, individual substrates (see below) were dissolved in 50 mM Tris-HCl (pH 7.5), supplemented with 200 mM $NADP^+$, 100 mM glucose, and glucose dehydrogenase (GDH) from *Thermoplasma acidophilum* (2 units for IVTT reactions, 5 units for purified enzyme reactions). Individual substrate concentrations were chosen to ensure v_{max} conditions or maximum solubility. At these substrate levels, reaction times were adjusted for 10-50% substrate conversion.

To assay OYE activity in IVTT experiments, 20 µL of the reaction stock solution was mixed with 2-10 µL IVTT reaction mixture and the total assay volume was adjusted to 30 µL

with 50 mM Tris-HCl (pH 7.5). To assay purified OYE1 variants, the enzyme (final concentration: 250 nM) was added to 500 µL reaction stock solution. Reaction progress was monitored by removing 30-µL aliquots from the assay solution and quenching them by mixing thoroughly with 30 µL of ethyl acetate containing 1 mM cyclohexanone as internal standard. A sample of the organic phase was collected and analyzed by GC (protocols: see below). The enantio/diasteriomeric excess were calculated by integration of product and substrate peak areas. Relative rates of conversion for individual substrates were calculated by dividing the measured rate of conversion for OYE1 variant over the corresponding rate for wild type OYE1.

(S)-Carvone. Reaction time: 4 h; Final substrate concentration: 10 mM. GC protocol: 90 °C, hold 5 min, then 1 °C/min to 120 °C (retention times for (*S*)-carvone=27.1 min, (1*R*,4*S*)-dihydrocarvone=22.9 min, (1*S*,4*S*)-dihydrocarvone=22.6 min).

(R)-Carvone. Reaction time: 4 h; Final substrate concentration: 2 mM. GC protocol: 90 °C, hold 5 min, then 1 °C /min to 120 °C (retention times for (*R*)-carvone=27.2 min, (1*R*,4*R*)-dihydrocarvone=22.3 min, (1*S*,4*R*)-dihydrocarvone=22.5 min).

Neral. Reaction time: 6 h; Final substrate concentration: 1 mM. GC protocol: 80 °C ramped at 2.5 °C/min to 115 °C, hold for 3 min (retention times for neral=13.0 min, citronellal=8.9 min).

Geranial. Reaction time: 0.5 h; Final substrate concentration: 5 mM. GC protocol: 80 °C ramped at 2.5 °C /min to 115 °C, hold for 3 min (retention times for geranial=14.8 min, citronellal=8.9 min).

Methyl-2-(hydroxymethyl)acrylate. Reaction time: 24 h; Final substrate concentration: 1 mM. GC protocol: 60 °C, hold for 1 min, then 1 °C /min to 75 °C, hold for 4 min (retention times for methyl-2-hydroxymethyl acrylate=16.6 min, (*R*)-3-hydroxy-2-methylpropanoate=15.0 min).

Ketoisophorone. Reaction time: 2.5 h; Final substrate concentration: 200 µM. GC protocol: 150 °C, hold for 5 min (retention times for ketoisophorone=3.26 min, (*R*)-levodione=3.59 min, (*S*)-levodione=4.4 min).

3-Methyl-2-cyclohexenone. Reaction time: 24 h; Final substrate concentration: 1 mM. GC protocol: 150 °C, hold for 5 min (retention times for 3-methyl-2-cyclohexenone=2.76 min, 3-(*S*)-methylcyclohexanone=2.06 min).

α-Methyl-trans-cinnamaldehyde. Reaction time: 0.5 h; Final substrate concentration: 1 mM. GC protocol: 150 °C hold for 10 min (retention times for α-methyl-trans-cinnamaldehyde=6.28 min, 2-methyl-3-phenylpropanal=3.82 min).

Supplemental information

Oligonucleotide sequences for construction of OYE1 variants.

Primer Design: *W116x*

W116Afor: CGTTTGGGTTCAGTTAGCGGTTTTGGGTTGGGC

W116Arev: GCCCAACCCAAAACCGCTAACTGAACCCAAACG

W116Vfor: CGTTTGGGTTCAGTTAGTGGTTTTGGGTTGGGC

W116Vrev: GCCCAACCCAAAACCACTAACTGAACCCAAACG

W116Nfor: CGTTTGGGTTCAGTTAAACGTTTTGGGTTGGGC

W116Nrev: GCCCAACCCAAAACGTTTAACTGAACCCAAACG

W116Ifor: GGGTTCAGTTAATAGTTTTGGGTTG

W116Irev: CAACCCAAAACTATTAACTGAACCC

W116Ffor: GGGTTCAGTTATTCGTTTTGGGTTG

W116Frev: GAAGGGAAAACGAATAACTGAACCC

Primer Design: *loop β6 Ala-scanning*

V288Afor: GCTTTTGTTCATTTGGCTGAACCTCGTG

V288Arev: CACGAGGTTCAGCCAAATGAACAAAAGC

E289Afor: GTTCATTTGGTTGCACCTCGTGTAAC

E289Arev: GTTACACGAGGTGCAACCAAATGAAC

P290Afor: CATTTGGTTGAAGCTCGTGTAACTAACCC

P290Arev: GGGTTAGTTACACGAGCTTCAACCAAATG

R291Afor: TTGGTTGAACCTGCTGTAACTAACCCATTC

R291Arev: GAATGGGTTAGTTACAGCAGGTTCAACCAA

V292Afor: GGTTGAACCTCGTGCAACTAACCCATTC

V292Arev: GAATGGGTTAGTTGCACGAGGTTCAACC

T293Afor: GAACCTCGTGTAGCTAACCCATTCTTG

T293Arev: CAAGAATGGGTTAGCTACACGAGGTTC

N294Afor: GAACCTCGTGTAACTGCCCCATTCTTGACTG

N294Arev: CAGTCAAGAATGGGGCAGTTACACGAGGTTC

L297Afor: CTAACCCATTCGCGACTGAAGGGGAGGG

L297Arev: CCCTCCCCTTCAGTCGCGAATGGGTTAG

T298Afor: CCCATTCTTGGCTGAAGGGGAGGG

T298Arev: CCCTCCCCTTCAGCCAAGAATGGG

E299Afor: CCCATTCTTGACTGCAGGGGAGGG

E299Arev: CCCTCCCCTGCAGTCAAGAATGGG

G300Afor: CATTCTTGACTGAAGCGGAGGGTGAATAC

G300Arev: GTATTCACCCTCCGCTTCAGTCAAGAATG

E301Afor: GACTGAAGGGGCGGGTGAATACGAAGG

E301Arev: CCTTCGTATTCACCCGCCCCTTCAGTC

G302Afor: GACTGAAGGGGAGGCTGAATACGAAGG

G302Arev: CCTTCGTATTCAGCCTCCCCTTCAGTC

E303Afor: GGGGAGGGTGCATACGAAGGAGG

E303Arev: CCTCCTTCGTATGCACCCTCCCC

Y304Afor: GGGGAGGGTGAAGCCGAAGGAGGTAGC

Y304Arev: GCTACCTCCTTCGGCTTCACCCTCCCC

E305Afor: GAGGGTGAATACGCAGGAGGTAGCAAC

E305Arev: GTTGCTACCTCCTGCGTATTCACCCTC

G306Afor: GGGTGAATACGAAGCAGGTAGCAACG

G306Arev: CGTTGCTACCTGCTTCGTATTCACCC

G307Afor: GGGTGAATACGAAGGAGCTAGCAACGAT

G307Arev: ATCGTTGCTAGCTCCTTCGTATTCACCC

S308Afor: GGTGAATACGAAGGAGGTGCCAACGATTTTGTTTAC

S308Arev: GTAAACAAAATCGTTGGCACCTCCTTCGTATTCACC

N309Afor: CGAAGGAGGTAGCGCCGATTTTGTTTACTC

N309Arev: GAGTAAACAAAATCGGCGCTACCTCCTTCG

D310Afor: GGAGGTAGCAACGCTTTTGTTTACTCCATC

D310Arev: GATGGAGTAAACAAAAGCGTTGCTACCTCC

Primer Design: *P295x*

P295Gfor: CCTCGTGTAACTAACGGATTCTTGACTGAAGGG

P295Grev: CCCTTCAGTCAAGAATCCGTTAGTTACACGAGG

P295Afor: CGTGTAACTAACGCATTCTTGACTGAAG

P295Arev: CTTCAGTCAAGAATGCGTTAGTTACACG

P295Vfor: CTCGTGTAACTAACGTATTCTTGACTGAAGGG

P295Vrev: CCCTTCAGTCAAGAATACGTTAGTTACACGAG

P295Lfor: CTCGTGTAACTAACCTATTCTTGACTGAAGG

P295Lrev: CCTTCAGTCAAGAATAGGTTAGTTACACGAG

P295Mfor: CTCGTGTAACTAACATGTTCTTGACTGAAGGGG

P295Mrev: CCCCTTCAGTCAAGAACATGTTAGTTACACGAG

P295Ffor: CTCGTGTAACTAACTTTTTCTTGACTGAAGGGG

P295Frev: CCCCTTCAGTCAAGAAAAAGTTAGTTACACGAG

P295Wfor: CTCGTGTAACTAACTGGTTCTTGACTGAAGGGG

P295Wrev: CCCCTTCAGTCAAGAACCAGTTAGTTACACGAG

P295Cfor: ACTAACTGTTTCTTGACTGA

P295Crev: CAAGAAACAGTTAGTTACACG

Primer Design: *H191x/N194x*

H191Nfor: GCTGGTGCCGATGGTGTTGAAATTAACAGTGCT

H191Nrev: AGCACTGTTAATTTCAACACCATCGGCACCAGC

H191Qfor: GCTGGTGCCGATGGTGTTGAAATTCAAAGTGCT

H191Qrev: AGCACTTTGAATTTCAACACCATCGGCACCAGC

H191Sfor: GCTGGTGCCGATGGTGTTGAAATTTCCAGTGCT

H191Srev: AGCACTGGAAATTTCAACACCATCGGCACCAGC

H191Tfor: GCTGGTGCCGATGGTGTTGAAATTACCAGTGCT

H191Trev: AGCACTGGTAATTTCAACACCATCGGCACCAGC

H191Cfor: GCTGGTGCCGATGGTGTTGAAATTTGCAGTGCT

H191Crev: AGCACTGCAAATTTCAACACCATCGGCACCAGC

H191Lfor: GCTGGTGCCGATGGTGTTGAAATTCTCAGTGCT

H191Lrev: AGCACTGAGAATTTCAACACCATCGGCACCAGC

N194Hfor: AGTGCTCACGGTTACTTGTTAAACCAGTTC

N194Hrev: GAACTGGTTTAACAAGTAACCGTGAGCACT

N194Qfor: AGTGCTCAAGGTTACTTGTTAAACCAGTTC

N194Qrev: GAACTGGTTTAACAAGTAACCTTGAGCACT

N194Sfor: AGTGCTTCCGGTTACTTGTTAAACCAGTTC

N194Srev: GAACTGGTTTAACAAGTAACCGGAAGCACT

N194Tfor: AGTGCTACCGGTTACTTGTTAAACCAGTTC

N194Trev: GAACTGGTTTAACAAGTAACCGGTAGCACT

N194Cfor: AGTGCTTGCGGTTACTTGTTAAACCAGTTC

N194Crev: GAACTGGTTTAACAAGTAACCGCAAGCACT

N194Lfor: AGTGCTCTCGGTTACTTGTTAAACCAGTTC

N194Lrev: GAACTGGTTTAACAAGTAACCGAGAGCACT

H191Afor: GCTGGTGCCGATGGTGTTGAAATTGCCAGTGCT

H191Arev: AGCACTGGCAATTTCAACACCATCGGCACCAGC

N194Afor: AGTGCTGCCGGTTACTTGTTAAACCAGTTC

N194Arev: GAACTGGTTTAACAAGTAACCGGCAGCACT

Table 3.S1: Catalytic conversion of geranial, neral, methyl-2-(hydroxymethyl) acrylate (MHMA) and (*S*)-carvone by W116x variants. The percent substrate conversion (% conv.) was measured after quenching the reaction at a predetermined substrate-specific time. All experiments were performed in triplicates to determine standard deviations (STD). The wild type enzyme (OYE) was run as internal positive control and served as reference for calculating relative conversion (rel. conv.). The enantioselectivity (ee) and diastereoselectivity (de) of individual variants is listed (standard error: 10%).

	Geranial			MHMA			Neral			(*S*)-carvone		
Gene	% conv.	STD	% ee (*R*)	% conv.	STD	% ee (*R*)	% conv.	STD	% ee (*R*)	% conv.	STD	% de
OYE	4.17	0.43	>90	3.79	0.25	>90	38.43	1.75	>90	4.78	0.15	88 (*R/S*)
OYEW116A	2.39	0.51	>90	3.70	1.64	>90	27.01	1.25	>90	6.65	0.31	>90 (*S/S*)
OYEW116V	0.57	0.35	>90	8.41	0.93	>90	12.15	1.48	>90	11.41	0.99	>90 (*S/S*)
OYEW116I	0.36	0.31	>90	4.86	0.18	>90	2.90	0.98	>90	11.68	0.81	>90 (*S/S*)
OYEW116N	0.09	0.30	>90	8.51	1.23	>90	5.47	1.71	>90	4.10	0.56	>90 (*S/S*)
OYEW116F	2.59	0.24	>90	3.30	0.51	>90	36.23	0.42	>90	4.49	0.38	>90 (*R/S*)

	Geranial		MHMA		Neral		(*S*)-carvone	
Gene	rel. conv.	STD	rel. conv.	STD	rel. conv.	STD	rel. conv.	STD
OYE	1.00	0.10	1.00	0.07	1.00	0.05	1.00	0.03
OYEW116A	0.58	0.12	0.96	0.11	0.70	0.03	1.39	0.05
OYEW116V	0.13	0.07	2.22	0.16	0.32	0.03	2.39	0.14
OYEW116I	0.08	0.06	1.28	0.06	0.08	0.03	2.44	0.11
OYEW116N	0.02	0.07	2.26	0.33	0.14	0.04	0.86	0.15
OYEW116F	0.63	0.07	0.87	0.09	0.94	0.04	0.94	0.09

Table 3.S2: Catalytic conversion of (*S*)-carvone by Ala-scanning variants in loop β6 region. The percent substrate conversion (% conv.) was measured after quenching the reaction at a predetermined substrate-specific time. All experiments were performed in triplicates to determine standard deviations (STD). The wild type enzyme (OYE) was run as internal positive control and served as reference for calculating relative conversion (rel. conv.). The OYE variants retained the native enzyme's preference for formation of (*R*/*S*)-dihydrocarvone. Diastereoselectivity (de) for each variant are listed (standard error: ± 10%)

Gene	% conv.	STD	rel. conv.	STD	% de
OYE	20.6	2.1	1.00	0.10	76%
OYE-V288A	6.1	1.0	0.30	0.05	56%
OYE-E289A	5.8	1.0	0.28	0.05	53%
OYE-P290A	7.6	1.5	0.37	0.07	57%
OYE-R291A	15.1	3.0	0.73	0.15	70%
OYE-V292A	16.0	3.9	0.78	0.19	70%
OYE-T293A	9.6	2.4	0.46	0.11	63%
OYE-N294A	17.2	2.8	0.84	0.13	72%
OYE-P295A	65.7	8.8	3.19	0.43	>90%
OYE-F296A	4.8	0.9	0.23	0.04	61%
OYE-L297A	13.7	3.1	0.67	0.15	68%
OYE-T298A	17.5	2.9	0.85	0.14	73%
OYE-E299A	17.1	1.5	0.83	0.08	71%
OYE-G300A	28.9	5.2	1.40	0.25	79%
OYE-E301A	17.9	2.8	0.87	0.14	73%
OYE-G302A	23.9	4.0	1.16	0.19	75%
OYE-E303A	13.1	2.0	0.63	0.10	68%
OYE-Y304A	37.6	8.0	1.82	0.39	80%
OYE-E305A	15.9	3.7	0.77	0.18	71%
OYE-G306A	15.1	1.8	0.74	0.09	71%
OYE-G307A	16.5	2.2	0.80	0.10	74%
OYE-S308A	16.2	1.5	0.79	0.07	73%
OYE-N309A	15.3	1.0	0.75	0.05	71%
OYE-D310A	18.6	2.6	0.90	0.13	75%

Table 3.S3: Catalytic conversion of geranial, neral, methyl-2-(hydroxymethyl) acrylate (MHMA), (*S*)-carvone, (*R*)-carvone, ketoisophorone, α-methyl-trans-cinnemaldehyde and 3-methyl-2-cyclohexenone by P295x variants. The percent substrate conversion (% conv.) was measured after quenching the reaction at a predetermined substrate-specific time. All experiments were performed in triplicates to determine standard deviations (STD). The wild type enzyme (OYE) was run as internal positive control and served as reference for calculating relative conversion (rel. conv.). The enantioselectivity (ee) and diastereoselectivity of individual variants is listed (standard error: 10%).

	geranial					neral				
Gene	% conv.	STD	rel. conv.	STD	% ee (*R*)	% conv.	STD	rel. conv.	STD	% ee (*R*)
OYE	7.8	0.6	1.00	0.08	>90%	46.9	5.5	1.00	0.12	>90%
OYEP295G	5.8	0.8	0.74	0.10	>90%	34.2	0.8	0.73	0.02	>90%
OYEP295A	5.6	1.0	0.72	0.12	>90%	31.7	3.1	0.68	0.07	>90%
OYEP295V	7.0	0.7	0.90	0.09	>90%	41.9	2.9	0.89	0.06	>90%
OYEP295L	6.3	0.3	0.81	0.04	>90%	38.9	3.1	0.83	0.07	>90%
OYEP295M	6.0	0.5	0.77	0.07	>90%	35.7	5.8	0.76	0.12	>90%
OYEP295F	6.7	1.2	0.86	0.15	>90%	34.5	3.0	0.73	0.06	>90%
OYEP295W	5.2	0.3	0.67	0.04	>90%	28.7	3.4	0.61	0.07	>90%
OYEP295C	7.0	0.4	0.90	0.06	>90%	30.8	1.8	0.66	0.04	>90%

	(*S*)-carvone					(*R*)-carvone				
Gene	% conv.	STD	rel. conv.	STD	% de (*R/S*)	% conv.	STD	rel. conv.	STD	% de (*R/R*)
OYE	4.6	0.9	1.00	0.20	80%	30.2	1.6	1.00	0.05	>90%
OYEP295G	4.2	0.4	0.92	0.08	75%	19.5	5.2	0.65	0.17	>90%
OYEP295A	30.5	2.7	6.69	0.59	>90%	86.0	4.9	2.85	0.16	>90%
OYEP295V	4.4	0.4	0.97	0.09	88%	18.0	0.8	0.60	0.03	>90%
OYEP295L	3.6	0.1	0.78	0.03	88%	26.3	4.8	0.87	0.16	>90%
OYEP295M	3.3	0.6	0.72	0.13	89%	19.2	1.7	0.64	0.06	>90%
OYEP295F	4.3	0.6	0.94	0.13	>90%	26.0	2.1	0.86	0.07	>90%
OYEP295W	3.2	0.7	0.70	0.15	>90%	19.2	1.4	0.64	0.05	>90%
OYEP295C	0.9	0.8	0.20	0.18	78%	5.7	1.5	0.19	0.05	>90%

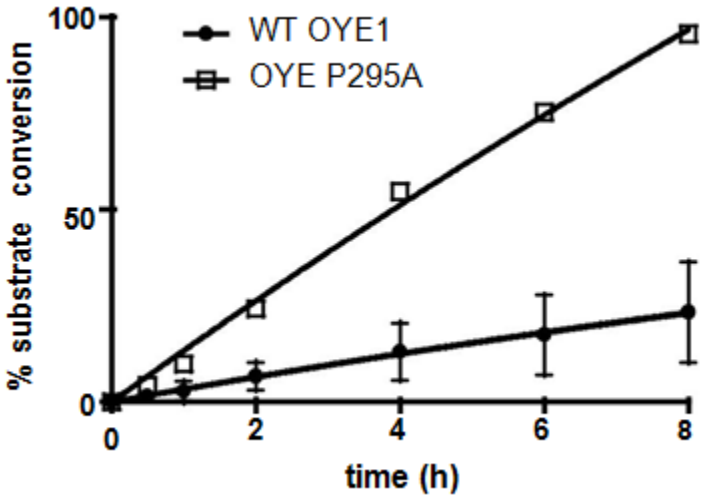
	MHMA					ketoisophorone				
Gene	% conv.	STD	rel. conv.	STD	% ee (*R*)	% conv.	STD	rel. conv.	STD	% ee (*R*)
OYE	5.7	0.5	1.00	0.09	>90%	8.9	0.6	1.00	0.07	>90%
OYEP295G	4.5	0.3	0.79	0.05	>90%	6.6	1.9	0.74	0.21	>90%
OYEP295A	4.6	0.4	0.80	0.07	>90%	17.0	3.0	1.91	0.33	>90%
OYEP295V	4.7	0.4	0.83	0.08	>90%	7.8	0.9	0.88	0.11	>90%
OYEP295L	4.5	0.4	0.80	0.08	>90%	8.5	1.5	0.95	0.17	>90%
OYEP295M	4.1	0.4	0.72	0.07	>90%	6.2	0.8	0.69	0.09	>90%
OYEP295F	5.0	0.5	0.87	0.09	>90%	13.2	8.4	1.49	0.09	41%
OYEP295W	4.2	0.2	0.74	0.03	>90%	5.8	0.1	0.65	0.01	>90%
OYEP295C	3.3	0.1	0.57	0.02	>90%	5.1	0.7	0.58	0.08	78%

	α-methyl-transcinnamaldehyde					3-methyl-2-cyclohexenone				
Gene	% conv.	STD	rel. conv.	STD	% ee (*S*)	% conv.	STD	rel. conv.	STD	% ee (*S*)
OYE	71.6	7.6	1.00	0.11	>90%	14.1	0.3	1.00	0.02	>90%
OYEP295G	45.0	17.8	0.63	0.25	>90%	9.1	3.0	0.65	0.22	>90%
OYEP295A	77.4	10.5	1.08	0.15	>90%	7.5	1.4	0.53	0.10	>90%
OYEP295V	54.2	18.3	0.76	0.26	>90%	11.1	2.7	0.79	0.19	>90%
OYEP295L	76.6	8.4	1.07	0.12	>90%	13.5	1.3	0.96	0.09	>90%
OYEP295M	54.1	10.1	0.76	0.14	>90%	10.6	1.9	0.76	0.14	>90%
OYEP295F	80.6	3.5	1.12	0.05	>90%	13.2	1.4	0.94	0.10	>90%
OYEP295W	51.5	3.7	0.72	0.05	>90%	9.0	0.3	0.64	0.02	>90%
OYEP295C	40.3	4.1	0.56	0.06	>90%	4.0	0.3	0.29	0.02	>90%

Table 3.S4: Catalytic conversion of methyl-2-(hydroxymethyl) acrylate (MHMA) and (*S*)-carvone by H191x and N194x variants. The percent substrate conversion (% conv.) was measured after quenching the reaction at a predetermined substrate-specific time. All experiments were performed in triplicates to determine standard deviations (STD). The wild type enzyme (OYE) was run as internal positive control and served as reference for calculating relative conversion (rel. conv.). Within detection limits, reduction of MHMA yielded exclusively the (*R*)-enantiomer while reduction of (*S*)-carvone produced the (*R*/*S*)-diastereomer at the listed percentages (% de) (standard error: ±10%).

	MHMA					(*S*)-carvone				
Gene	% conv.	STD	rel. conv.	STD	% ee	% conv.	STD	rel. conv.	STD	% de
OYE	57.3	0.0	1.00	0.00	>90	6.5	0.0	1.00	0.00	76%
H191A	17.3	10.8	0.30	0.19	>90	0.3	0.5	0.04	0.05	>90
H191L	15.7	0.5	0.27	0.01	>90	0.0	0.0	0.00	0.00	-
H191Q	27.6	25.6	0.48	0.45	>90	0.0	0.0	0.00	0.00	-
H191S	58.5	42.7	1.02	0.74	>90	0.0	0.0	0.00	0.00	-
H191C	28.6	9.1	0.50	0.16	>90	0.0	0.0	0.00	0.00	-
H191N	29.0	8.6	0.51	0.15	>90	0.0	0.0	0.00	0.00	-
H191T	21.6	10.9	0.38	0.19	>90	0.0	0.0	0.00	0.00	-
OYE	57.3	0.0	1.00	0.00	>90	6.5	0.0	1.00	0.00	76%
N194A	38.0	4.1	0.66	0.07	>90	0.3	0.5	0.09	0.05	>90
N194L	25.4	15.6	0.44	0.13	>90	0.0	0.0	0.00	0.00	-
N194Q	33.9	4.2	0.59	0.07	>90	1.1	0.4	0.22	0.09	80%
N194S	31.7	10.5	0.55	0.18	>90	1.1	0.3	0.21	0.05	73%
N194C	42.7	4.4	0.75	0.08	>90	0.0	0.0	0.00	0.00	-
N194H	0.0	0.0	0.00	0.00	-	3.9	0.5	0.75	0.10	>90
N194T	25.1	10.7	0.44	0.19	>90	0.2	0.3	0.02	0.03	>90

Figure 3.S1: Time-course for reduction of (*S*)-carvone by wild type OYE1 and OYE1 P295A variant. The estimated three-fold activity increase observed for the P295A variant in the RAPPER analysis was reproducible in larger-scale reaction mixtures containing heterologously expressed and purified enzyme. The percent substrate conversion (% conv.) was measured after removing and quenching 30-µl aliquots of a 500-µl reaction mixture containing 10 mM (*S*)-carvone, 200 µM $NADP^+$, 100 mM glucose, 5 U of glucose dehydrogenase, and 250 nM enzyme in 100 mM Tris-HCl (pH 7.6), 20 mM NaCl. Each experiment was performed in triplicate.



Bibliography

1. Bornscheuer, U. T.; Huisman, G. W.; Kazlauskas, R. J.; Lutz, S.; Moore, J. C.; Robins, K., Engineering the third wave of biocatalysis. *Nature* **2012,** *485* (7397), 185-94.

2. Brown, B. J.; Hyun, J. W.; Duvvuri, S.; Karplus, P. A.; Massey, V., The role of glutamine 114 in old yellow enzyme. *J Biol Chem* **2002,** *277* (3), 2138-45.

3. Brenna, E.; Crotti, M.; Gatti, F. G.; Monti, D.; Parmeggiani, F.; Powell, R. W.; Santangelo, S.; Stewart, J. D., Opposite Enantioselectivity in the Bioreduction of (Z)-β-Aryl-β-cyanoacrylates Mediated by the Tryptophan 116 Mutants of Old Yellow Enzyme 1: Synthetic Approach to (R)- and (S)-β-Aryl-γ-lactams. *Adv. Synth. Catal.* **2015,** *357* (8), 1849-1860.

4. Zhou, X.; Chow, H. L.; Wu, J. C., Bioreduction of activated alkenes by a novel "ene"-reductase from the thermophilic strainBacillus coagulansWCP10-4. *Biocatal. Biotransform.* **2014,** *32* (5-6), 267-275.

5. Toogood, H. S.; Gardiner, J. M.; Scrutton, N. S., Biocatalytic Reductions and Chemical Versatility of the Old Yellow Enzyme Family of Flavoprotein Oxidoreductases. *ChemCatChem* **2010,** *2* (8), 892-914.

6. Turrini, N. G.; Hall, M.; Faber, K., Enzymatic Synthesis of Optically Active Lactones via Asymmetric Bioreduction using Ene-Reductases from the Old Yellow Enzyme Family. *Adv. Synth. Catal.* **2015,** *357* (8), 1861-1871.

7. Amato, E. D.; Stewart, J. D., Applications of protein engineering to members of the old yellow enzyme family. *Biotechnol. Adv.* **2015,** *33* (5), 624-631.

8. Adalbjörnsson, B. V.; Toogood, H. S.; Fryszkowska, A.; Pudney, C. R.; Jowitt, T. A.; Leys, D.; Scrutton, N. S., Biocatalysis with Thermostable Enzymes: Structure and Properties of a Thermophilic 'ene'-Reductase related to Old Yellow Enzyme. *ChemBioChem* **2010,** *11* (2), 197-207.

9. Toogood, H. S.; Fryszkowska, A.; Hare, V.; Fisher, K.; Roujeinikova, A.; Leys, D.; Gardiner, J. M.; Stephens, G. M.; Scrutton, N. S., Structure-Based Insight into the Asymmetric Bioreduction of the C□C Double Bond of α,β-Unsaturated Nitroalkenes by Pentaerythritol Tetranitrate Reductase. *Adv. Synth. Catal.* **2008,** *350* (17), 2789-2803.

10. Durchschein, K.; Hall, M.; Faber, K., Unusual reactions mediated by FMN-dependent ene- and nitro-reductases. *Green Chem.* **2013,** *15* (7), 1764-1772.

11. Toogood, H. S.; Scrutton, N. S., New developments in 'ene'-reductase catalysed biological hydrogenations. *Curr. Opin. Chem. Biol.* **2014,** *19*, 107-115.

12. Bougioukou, D. J.; Kille, S.; Taglieber, A.; Reetz, M. T., Directed Evolution of an Enantioselective Enoate-Reductase: Testing the Utility of Iterative Saturation Mutagenesis. *Adv. Synth. Catal.* **2009,** *351* (18), 3287-3305.

13. Padhi, S. K.; Bougioukou, D.; Stewart, J. D., Site-Saturation Mutagenesis of Tryptophan 116 of *Saccharomyces pastorianus* Old Yellow Enzyme Uncovers Stereocomplementary Variants. *J. Am. Chem. Soc.* **2009,** *131*, 3271-3280.

14. Hulley, M. E.; Toogood, H. S.; Fryszkowska, A.; Mansell, D.; Stephens, G. M.; Gardiner, J. M.; Scrutton, N. S., Focused directed evolution of pentaerythritol tetranitrate reductase by using automated anaerobic kinetic screening of site-saturated libraries. *ChemBioChem* **2010,** *11* (17), 2433-47.

15. Hall, M.; Bommarius, A. S., Enantioenriched compounds via enzyme-catalyzed redox reactions. *Chem. Rev.* **2011,** *111* (7), 4088-110.

16. Toogood, H. S.; Fryszkowska, A.; Hulley, M.; Sakuma, M.; Mansell, D.; Stephens, G. M.; Gardiner, J. M.; Scrutton, N. S., A site-saturated mutagenesis study of pentaerythritol tetranitrate reductase reveals that residues 181 and 184 influence ligand binding, stereochemistry and reactivity. *ChemBioChem* **2011,** *12* (5), 738-49.

17. Reich, S.; Hoeffken, H. W.; Rosche, B.; Nestl, B. M.; Hauer, B., Crystal structure determination and mutagenesis analysis of the ene reductase NCR. *ChemBioChem* **2012,** *13* (16), 2400-7.

18. Pompeu, Y. A.; Sullivan, B.; Stewart, J. D., X-ray Crystallography Reveals How Subtle Changes Control the Orientation of Substrate Binding in an Alkene Reductase. *ACS Catalysis* **2013,** *3* (10), 2376-2390.

19. Reich, S.; Kress, N.; Nestl, B. M.; Hauer, B., Variations in the stability of NCR ene reductase by rational enzyme loop modulation. *J. Struct. Biol.* **2014,** *185* (2), 228-233.

20. Walton, A. Z.; Sullivan, B.; Patterson-Orazem, A. C.; Stewart, J. D., Residues Controlling Facial Selectivity in an Alkene Reductase and Semirational Alterations to Create Stereocomplementary Variants. *ACS Catalysis* **2014,** *4* (7), 2307-2318.

21. Daugherty, A. B.; Govindarajan, S.; Lutz, S., Improved biocatalysts from a synthetic circular permutation library of the flavin-dependent oxidoreductase old yellow enzyme. *J. Am. Chem. Soc.* **2013,** *135* (38), 14425-32.

22. Shimizu, Y.; Inoue, A.; Tomari, Y.; Suzuki, T.; Yokogawa, T.; Nishikawa, K.; Ueda, T., Cell-free translation reconstituted with purified components. *Nat Biotech* **2001,** *19* (8), 751-755.

23. Quertinmont, L. T.; Orru, R.; Lutz, S., RApid Parallel Protein EvaluatoR (RAPPER), from gene to enzyme function in one day. *Chem. Commun.* **2015,** *51* (1), 122-124.

24. Daugherty, A. B.; Horton, J. R.; Cheng, X.; Lutz, S., Structural and Functional Consequences of Circular Permutation on the Active Site of Old Yellow Enzyme. *ACS Catalysis* **2015,** *5* (2), 892-899.

25. Horita, S.; Kataoka, M.; Kitamura, N.; Nakagawa, T.; Miyakawa, T.; Ohtsuka, J.; Nagata, K.; Shimizu, S.; Tanokura, M., An engineered old yellow enzyme that enables efficient

synthesis of (4R,6R)-Actinol in a one-pot reduction system. *ChemBioChem* **2015,** *16* (3), 440-5.

26. Rockah-Shmuel, L.; Tóth-Petróczy, Á.; Sela, A.; Wurtzel, O.; Sorek, R.; Tawfik, D. S., Correlated Occurrence and Bypass of Frame-Shifting Insertion-Deletions (InDels) to Give Functional Proteins. *PLoS Genet.* **2013,** *9* (10), e1003882.

27. Arpino, James A. J.; Reddington, Samuel C.; Halliwell, Lisa M.; Rizkallah, Pierre J.; Jones, D. D., Random Single Amino Acid Deletion Sampling Unveils Structural Tolerance and the Benefits of Helical Registry Shift on GFP Folding and Structure. *Structure* **2014,** *22* (6), 889-898.

28. Fox, K. M.; Karplus, P. A., Old yellow enzyme at 2 A resolution: overall structure, ligand binding, and comparison with related flavoproteins. *Curr. Biol.* **1994,** *2* (11), 1089-1105.

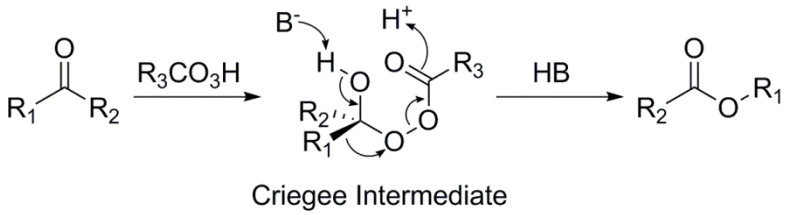
29. Brown, B. J.; Deng, Z.; Karplus, P. A.; Massey, V., On the Active Site of Old Yellow Enzyme: Role of Histidine 191 and Asparagine 194. *J. Biol. Chem.* **1998,** *273* (49), 32753-32762.

30. Park, J. T.; Gómez Ramos, L. M.; Bommarius, A. S., Engineering towards Nitroreductase Functionality in Ene-Reductase Scaffolds. *ChemBioChem* **2015,** *16* (5), 811-818.

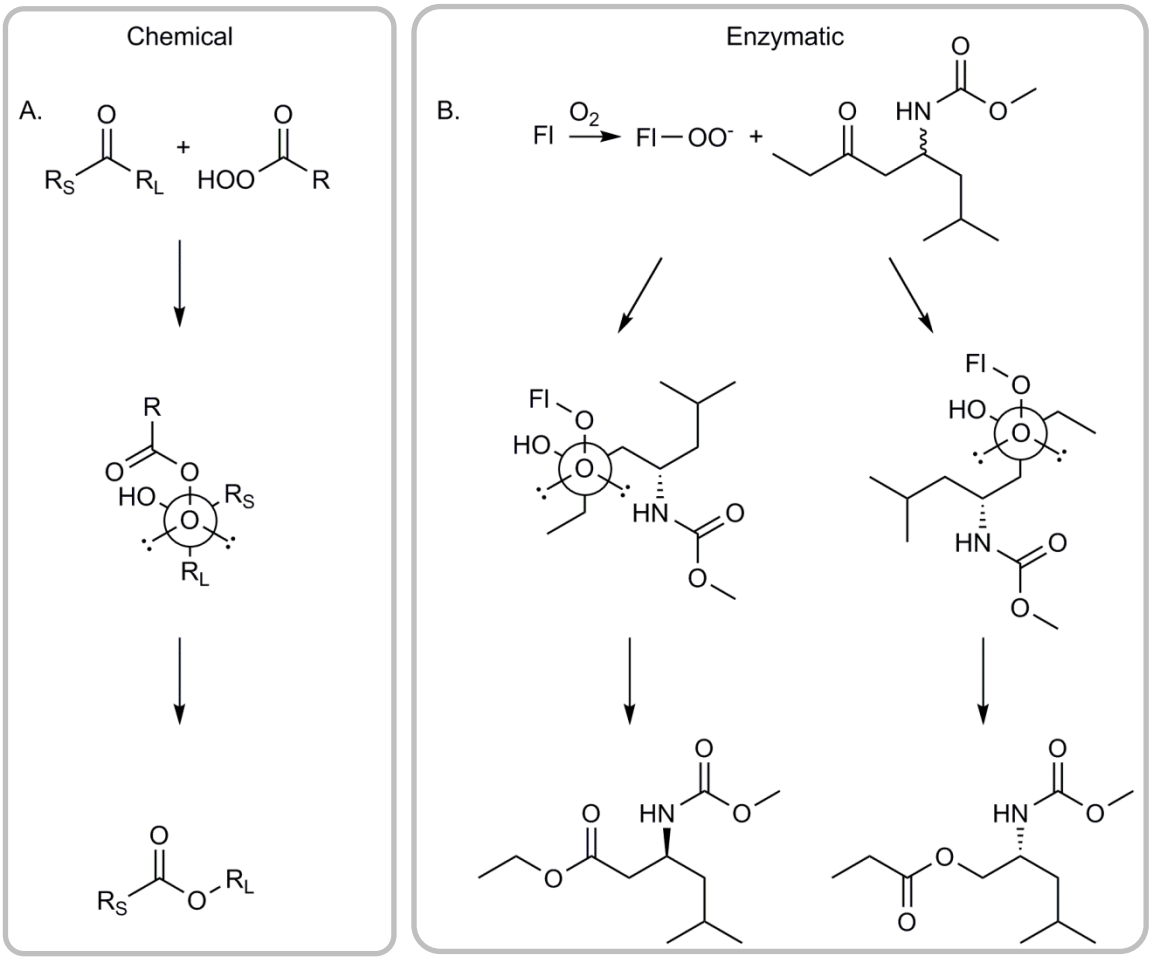
Chapter 4: Expanding the regioselectivity of Cyclododecanone Monooxygenase

Introduction

Transforming ketones into esters or lactones through the use of peracids via the Baeyer-Villiger oxidation has been a key organic synthetic reaction since its discovery over a century ago[1]. The Baeyer-Villiger oxidation results in the formation of either an ester or a lactone from a ketone or cyclic ketone, respectively. The reaction is carried out using peroxyacid or peroxide as the oxidant to insert an oxygen via the formation of a tetrahedral Criegee intermediate[2], shown in Scheme 4.1. One of the substituents on the ketone then migrates to the oxygen of the peroxide group, and the carboxylic acid leaves as a by-product[3]. The substituent that moves next to the oxygen to become part of the resulting ester is known as the migrating group. In the case of asymmetric ketones, two regioisomers can be produced as a result of a Baeyer-Villiger oxidation. Control over which of these regioisomers will be produced is highly sought after and is typically determined by the migratory aptitude of the substituents on the ketone[3]. One of the principles that determines the migratory effect is the primary stereoelectronic effect, which states that the oxygen-oxygen bond of the peroxide must be aligned anti-periplanar to the migrating group, as represented in Scheme 4.2A[4]. Since the migrating group must be able to stabilize a positive charge as it moves, the migratory aptitude of substituents are typically predictable and ranked in the following order: tertiary alkyl > cyclohexyl > secondary alkyl > benzyl > phenyl > primary alkyl > CH_3[5]. Additionally, the bulkier substituents will prefer to be anti-periplanar to the peroxide to minimize steric clashes, resulting in a greater likelihood for larger substituents to migrate[6].



Scheme 4.1. General chemical Baeyer-Villiger oxidation, with the Criegee intermediate shown.



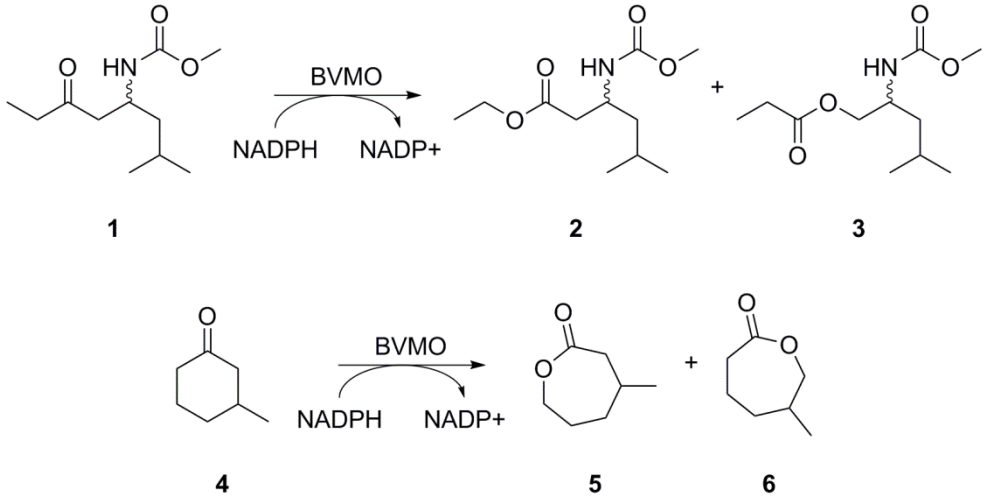
Scheme 4.2. A) Representation of the primary stereoelectronic effect of the chemical Baeyer-Villiger oxidation, demonstrating the necessity of the migratory group (R_L) to be anti-periplanar to the oxygen-oxygen bond of the peroxide. B) Representation within the enzymatic environment of the ability of **1** to form the Criegee intermediate while avoiding steric clashes with the flavin (Fl) by adopting two different conformations. (*S*)-**1** requires the smaller, ethyl group to be anti-periplanar to the oxygen-oxygen bond of the peroxide. (*R*)-**1** requires the larger group to be anti-periplanar to the oxygen-oxygen bond of the peroxide.

The focus of this study is primarily on controlling the regioselectivity of the oxidation of two substrates: N-protected β-amino ketones and 3-methylcyclohexanone (Scheme 4.3). The N-protected β-amino ketone, **1**, produces β-amino acid ester (**2**) and β-amino alkylacetate (**3**) as a result of a Baeyer-Villiger oxidation. These products can subsequently be hydrolyzed into a β-amino alcohol and a β-amino acid, respectively. Each of these compounds, individually, is of industrial interest. β-amino alcohols can be used as chiral auxiliaries in asymmetric syntheses[7] as well as in the syntheses of high value agrochemicals and pharmaceuticals[8], while β-amino acids have gained attention in recent

years due to their interesting pharmacological applications, such as their antiketogenic properties as well as their antibacterial and antifungal activities[9]. **1** is not commercially available; as a result, a novel synthetic route was developed for both the racemic and chiral β-amino ketone by Dr. Huanyu Zhao.

A Baeyer-Villiger oxidation of 3-methylcyclohexanone (**4**) produces both 3-methyl ε-caprolactone (**5**) and 5-methyl ε-caprolactone (**6**), which are useful precursors for polyester synthesis. The two regioisomers produced have different rates of polymerization to form polyesters in their downstream use[10]. Additionally, each regioisomer can be a mixture of enantiomers and previous work has shown that different enantiomers are preferred depending upon which lipase is employed for ring opening[11]. Having control over the chirality and regioisomer of the lactone produced is important for subsequent polymer applications[11].

While having good regio- and/or enantioselectivity for the Baeyer-Villiger oxidation of these compounds is desired, this is not a trivial task. The chemical Baeyer-Villiger oxidation of **1** and **4** using the traditional reaction method of *m*-CPBA with 30% H_2O_2 demonstrated poor regio- and enantioselectivity. **4** went to completion after three days, producing both **5** and **6** in equal amounts. In the reaction with **1**, the chemical Baeyer-



Scheme 4.3. Oxidations of interest for this study.

Villiger oxidation had less than 10% conversion after one week under standard conditions (see experimental). After subjecting **1** to much harsher conditions, creating trifluoroperacetic acid *in situ*, the reaction proceeded to 84% completion after one week. This resulted in both enantiomers being consumed equally, and the products regioisomers being produced in a 35:65 ratio of **2** to **3**.

The use of these strong organic peracids, which are unstable, explosive, and shock sensitive is one of the primary disadvantages of the chemical Baeyer-Villiger oxidation[12]. In contrast, Baeyer-Villiger Monooxygenases (BVMOs) are a class of flavoenzymes that perform the Baeyer-Villiger oxidation with typically good regio-, enantio- and/or chemo-selectivity by using molecular oxygen instead of harsh organic peracids[13-15]. Additionally, BVMOs are not limited to the chemically preferred, or 'normal', product and have been shown repeatedly to catalyze the formation of the 'abnormal' product[8, 16-20]. By utilizing BVMOs, the high regioselectivity of the enzymes can be exploited or, if necessary, engineered[20-26]. Mechanistically, BVMOs follow the classic example of monooxygenases and utilize a C4a-peroxyflavin to perform oxidation chemistry. Briefly, BVMOs generally utilize a reductive half reaction wherein NADPH reduces FAD, followed by an oxidative half reaction where molecular oxygen re-oxidizes flavin adenine dinucleotide (FAD) via a C4a-peroxyflavin intermediate[27]. Once the nicotinamide cofactor binds and reduces FAD, $NADP^+$ remains tightly bound to the enzyme for the rest of the catalytic cycle. Following substrate binding the C4a-peroxyflavin reacts quickly to form a Criegee intermediate[2, 27-28]. A subsequent conformational change then triggers product release and $NADP^+$ release[2, 27-31]. This presumed conformational shift has been inferred based upon biochemical studies performed by the Massey lab, as well as crystal structures of cyclohexanone monooxygenase ($CHMO_{Rhodo}$) from *Rhodococcus* sp. HI-31 resolved by the Berguis group[29-31]. In a comparison of these structures during different parts of the catalytic cycle, the NADPH

binding domain of the enzyme showed clear shifts relative to the rest of the protein (refer to Figure 1.9A)[29-31].

Due to the benefits in regard to regio- and enantio- selectivity that BVMOs can provide under benign conditions, we turned to them for the oxidation of **1** and **4** since previous work has demonstrated that the selectivity of BVMOs can be optimized or changed via protein engineering[20-25, 32-33]. Certain BVMOs have demonstrated the ability to catalyze the formation of the abnormal ester or lactone, as seen when cyclohexanone monooxygenase ($CHMO_{Arthro}$) from *Arthrobacter* sp. utilizes (+)-*trans*-dihydrocarvone as a substrate[20]. Work performed by the Bornscheuer group engineered $CHMO_{Arthro}$ to change the regioselectivity from the abnormal product to the normal product[20]. They were able to achieve this through the targeting of several phenylalanine and leucine residues, shown in Figure 4.1[20]. Specifically, they substituted valine or alanine for F299 and F330 and phenylalanine for L196 or L197, individually and in various combinations[20]. This work was the first to demonstrate a complete reversal of regioselectivity, favoring formation of the

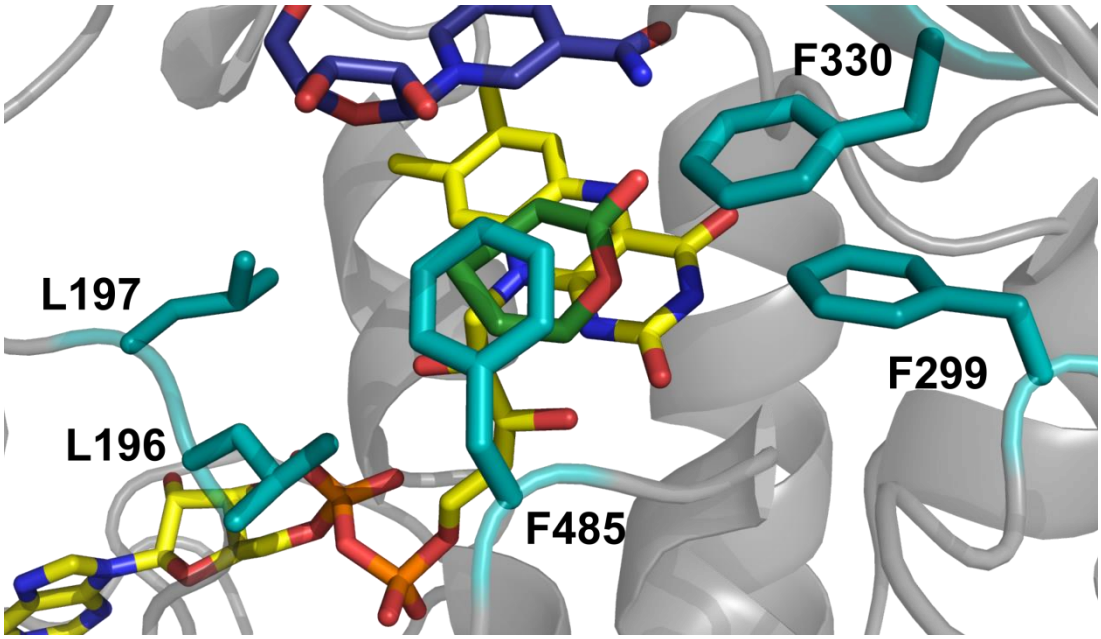


Figure 4.1. Close up of the active site of CHMO with the residues targeted for engineering highlighted in cyan. The FAD (yellow), NADP (blue) and ε-caprolactone (green) are shown in the active site. PDBID: 4RG4.[31]

normal, kinetically preferred product, over the natural, abnormal regioisomer.

More recent work by the Bornscheuer lab applied the same strategy of replacing large residues with small ones to change the regioselectivity of 2-oxo-Δ3-4,5,5-trimethylcyclopentenylacetyl-CoA monooxygenase (OTEMO) towards ((–)-(1*S*,5*R*)-cis-bicyclo[3.2.0]hept-2-en-6-one) away from a 52:48 normal:abnormal mixture[23]. This resulted in the creation of two variants which favored either regioisomer by at least 90%[23]. However, the amino acid substitution necessary for creation of the abnormal product (W501V) resulted in decreased activity of the enzyme[34]. The residues targeted for mutagenesis in this study were homologous to the phenylalanine residues targeted in the $CHMO_{Arthro}$ study[20, 23]. This work demonstrated the ability to change the regioselectivity of a BVMO towards whichever regioisomer is most desired.

Other researchers have reported regioselectivity changes to favor the abnormal product, including work by van Beek et al. on CHMO from *Acinetobacter calcoaceticus* to produce methyl propanoate[22]. However, this work was not as high yielding (increasing by only 14% to a final production of 40%) compared to the work performed by the Bornscheuer lab[22]. Another example of altering the regioselectivity towards the abnormal product comes from the Fraaije lab when engineering phenylacetone monooxygenase (PAMO). Throughout the process of mapping the binding site of PAMO, they discovered a variant (Q152F) that changed the regioselectivity towards bicycloheptenone to prefer the abnormal product 70:30[35]. This is a homologous position to L196 in $CHMO_{Arthro}$ targeted by the Bornscheuer group. However, while the substitution did show promise in altering the regioselectivity, it also caused a decrease in activity which became problematic when combined with certain other amino acid changes[35].

While there have been engineering studies performed on BVMOs, discussed above, the enzyme of interest for our work is cyclododecanone monooxygenase (CDMO) from

Rhodococcus ruber. CDMO is an FAD-containing Type I BVMO that contains key functional motifs, including the characteristic BVMO fingerprint[36]. Of particular interest for our work is that CDMO has a larger active site than most characterized BVMOs, making it more ideal for large substrates, such as **1**[37]. Work done by the Mihovilovic group presented a comparison of several BVMOs in regards to their performance in terms of both activity and selectivity[38]. This work found that CDMO was the most active towards N-protected β-amino ketones of the BVMOs analyzed, which allowed for a good starting point for our engineering efforts[39].

Results and Discussion

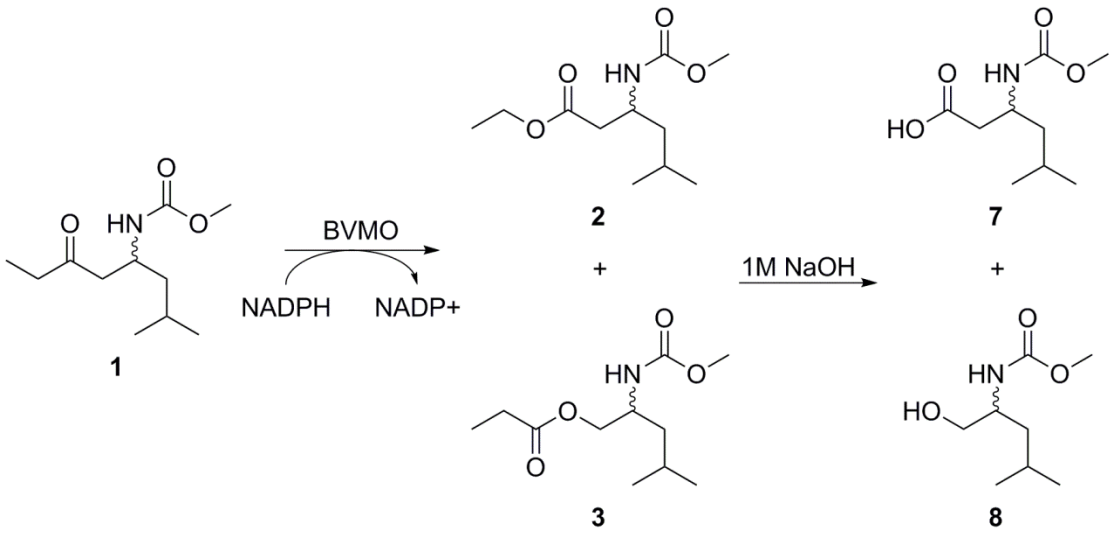
Initial characterization

The native CDMO was demonstrated to have a slight preference towards the formation of **2** when a racemic mixture of **1** was consumed, producing a ratio of 55:45 of **2** to **3** (Table 4.1). This mixture of **2** and **3** demonstrated that either regioisomer could be obtained, but would require engineering of CDMO in order shift the product profile to favor one regioisomer over the other. After an enzymatic reaction with CDMO, we were unable to separate the enantiomers of **2** and **3** via GCMS. As a result, **2** and **3** were hydrolyzed to their respective acid (**7**) and alcohol (**8**) (Scheme 4.4), which were able to be separated based

Temperature (°C)	Total % Conversion of **1**	Ratio of Products **2**:**3**	Total % Conversion of **4**	Ratio of Products **5**:**6**
22	10 ± 2%	55:45	23 ± 1%	70:30
28	17 ± 3%	58:42	33 ± 2%	68:32
34	26 ± 4%	60:40	52 ± 8%	65:35

Table 4.1. Conversion data for wtCDMO. Data for the conversion of **1** was obtained after 24 hours. Data for the conversion of **4** was obtained after 8 hours.

upon their enantiomers, and it became apparent that CDMO also has a clear enantiopreference. The products of the enzymatic reaction were exclusively (*S*)-**2** and (*R*)-**3**. This strict enantiopreference likely indicates two distinct binding orientations of **1**. (*R*)-



Scheme 4.4. Hydrolysis of the products of the BVMO reaction with **1**.

1 likely adopts a flipped orientation relative to (*S*)-**1**, resulting in the migratory bond switching from the large group (forming **3**) to the small (ethyl) group (forming **2**) (Scheme 4.2B). We also examined the initial activity of CDMO towards **4** as a secondary substrate against which to test activity and regioselectivity changes. When a racemic mixture of **4** was consumed, a ratio of 70:30 of **5**:**6** was produced, demonstrating a preference towards the production of the abnormal product by CDMO. Additionally, CDMO was more active towards **4** than **1**, resulting in the use of ten times more substrate, and only requiring eight hours of incubation instead of twenty four, for the activity assay **4**.

Creation of a homology model

Despite attempts at crystallization, there is currently no crystal structure of CDMO. In the absence of a detailed structure, we generated a homology model for CDMO to guide engineering efforts by identifying key active site residues and evaluating areas of potential structural differences. The homology model was developed using RaptorX[40], and most closely resembles the crystal structure of PockeMO, which has about 40% sequence identity with CDMO and an enlarged active site compared to other current crystal

structures of BVMOs[41]. The CDMO model contained FAD bound, as well as the ternary complex with FAD, $NADP^+$ and caprolactone, based upon superposition of the coordinates of $CHMO_{Rhodo}$. The model was refined using Rosetta Design Software[42]. After relaxation, the backbone resembles the PockeMO structure, with the exception of a loop region near the active site (Figure 4.2). The FAD remained in nearly the same position, but the nicotinamide moiety of $NADP^+$ shifted away from the active site, allowing for the catalytically relevant positioning of caprolactone relative to FAD, similar to the movement seen within the Berghuis structures[29-31].

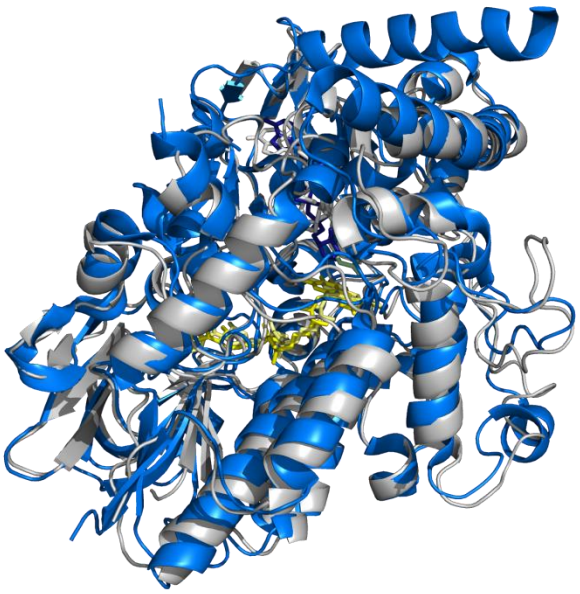


Figure 4.2. Overlay of the CDMO homology model (gray) and PockeMO (blue, PDBID: 5MQ6[43]). $NADP^+$ (blue/gray), FAD (yellow) and ε-caprolactone (green) are shown bound in the active site.

Selection of residues

Our engineering efforts towards CDMO were initially based upon the studies performed by the Bornscheuer lab[20]. In that work, several residues influencing the regioselectivity of $CHMO_{Arthro}$ were identified, which provided a starting point for our work on CDMO. An overlay of the CDMO homology model and the active site of $CHMO_{Rhodo}$, used by the Bornscheuer group for their model, showed similarities in two of the three regions of interest for engineering (Figure 4.3). The loop region of 497-501 (CDMO numbering) diverges between the two, since this loop region is longer in CDMO compared to CHMO. However, the other two regions, located at positions 190/191 and positions 299/325, are homologous. Where the CHMO active site has two sequential leucine residues, CDMO

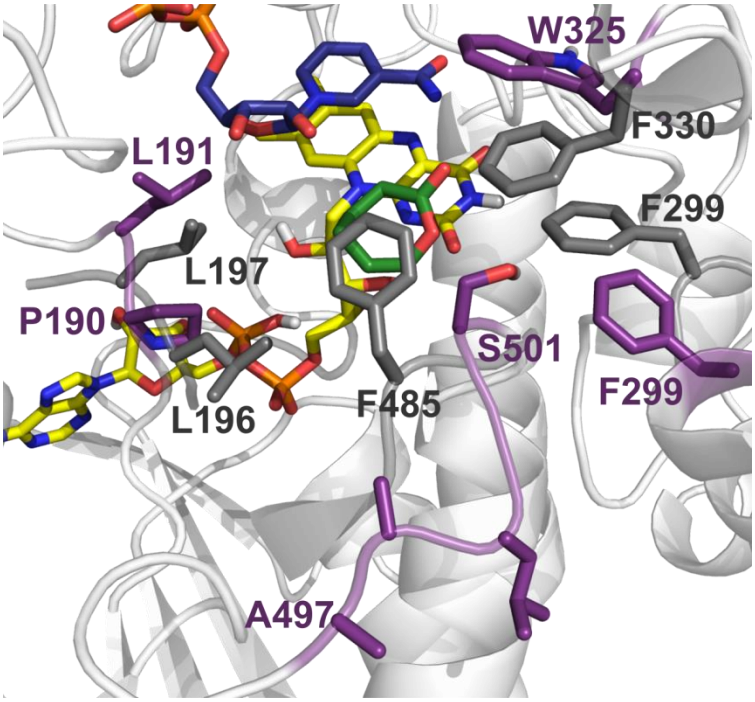


Figure 4.3. Active site overlay of the CDMO homology model (purple) and CHMO (gray, PDBID: 4RG4[31]) with the residues of interest for engineering highlighted. $NADP^+$ (blue) and ε-caprolactone (green) are shown bound in the active site.

contains a proline followed by a leucine. The 190/191 region is flanked by areas of high sequence homology, likely due to the FAD and NADPH binding domains found in this region. Positions 299/325 likewise has residues that are homologous to their CHMO counterparts, containing a phenylalanine and a tryptophan instead of two phenylalanine residues. However, this region does not have as high of sequence similarity in the regions surrounding it as the 190/191 region.

Substitutions at positions 299 and 325

Following the Bornscheuer work, F299 and W325 were initial targets for protein engineering to determine if the results observed with $CHMO_{Arthro}$ could be directly translated to CDMO. The selected residues were changed to either alanine or valine. The chosen substitutions indicated that the positions were not particularly amenable to amino acid changes, specifically when changed to valine. W325V retained some activity at room temperature (22 °C), but at elevated temperatures lost activity (Table 4.2). The alanine variants in each position showed diverging shifts in regioselectivity towards **1**. The opposite effects of these two alanine substitutions could is not entirely surprising, given the previous

work done on $CHMO_{Arthro}$. The previous work indicated that a single amino acid change at F330 had little effect on stereoselectivity, still favoring the abnormal product[20]. However a single amino acid substitution at F299 had a significant effect on stereoselectivity, favoring the normal product over the abnormal product[20]. Our work with CDMO supports these opposing effects, but activity was significantly impacted by substitutions at these positions. Since these residues are less amenable to substitution than originally anticipated, alternate amino acid substitutions at these positions were not further pursued.

	fold activity with **1** at			Ratio of Products	
Enzyme	22 °C	28 °C	34 °C	**2**:**3**	**5**:**6**
F299A	1.5	0.9	0.1	34:66	74:26
F299V	0.0	0.0	0.0	n.a.	n.a.
W325A	1.1	0.1	0.0	82:18	n.a.
W325V	0.4	0.1	0.0	87:13	n.a.
wt	1.0	1.0	1.0	55:45	70:30

Table 4.2. Activity of F299 and W325 variants after 24 hours at 22, 28, and 34 °C with **1**, relative to wtCDMO at each temperature. The regioselectivity is based on total product at the end of 24 hrs (**1**) or 8 hrs (**4**) at 22 °C. Regioselectivity of n.a. indicates no conversion.

Substitutions at positions 190 and 191

A panel of variants was created at Pro190 to explore steric effects at that position (Table 4.3). Position 190 was more amenable to substitution compared to positions 299 or 325, based upon relative activity of the variants tested. Most of the variants showed some gains in activity towards **1**, with P190L exhibiting an almost four-fold improvement compared to wild-type. Although the activity of P190L towards **1** was significantly higher than wtCDMO, the regioselectivity remained relatively unchanged (Figure 4.4). The regioselectivity of CDMO variants in regards to **1** differs widely depending upon which amino acid substitution was made, with the smallest and largest substitutions favoring the formation of **3**. Additionally, all variants converted both (*R*)-**1** and (*S*)-**1**, showing a slight preference for (*S*)-**1** (Figure 4.4). The exception to this was P190G, which showed a

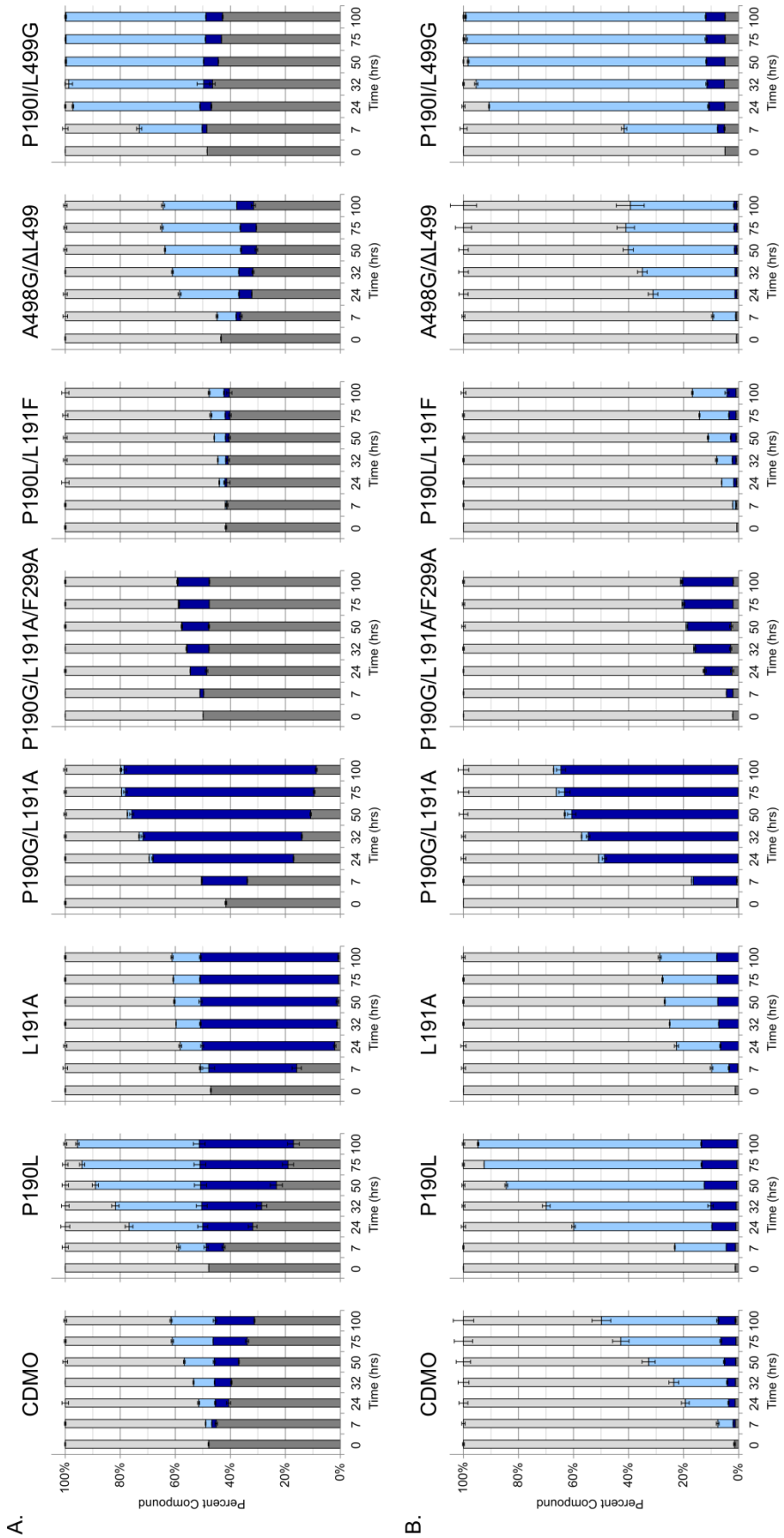


Figure 4.4. Time courses of wtCDMO and representative CDMO variants with A) racemic **1** and B) (*S*)-**1**. Remaining (*R*)-**1** (dark gray) and (*S*)-**1** (light gray) at each time point are represented along with amount of both **2** (dark blue) and **3** (light blue) as relative percentages of the total reaction components.

preference for (*R*)-**1**. One more variant to note in regards to its activity towards **1** is P190I, which changed the regioselectivity to produce 85% **2**, compared to wild-type which only produced 55%.

	fold activity with **1** at			Ratio of Products	
Enzyme	22 °C	28 °C	34 °C	**2**:**3**	**5**:**6**
P190G	2.8	1.9	1.4	14:86	67:33
P190A	2.3	1.7	1.3	17:83	79:21
P190V	0.8	0.8	0.9	77:23	53:47
P190I	0.8	0.9	1.1	85:15	58:42
P190L	3.9	3.0	1.9	61:39	57:43
P190M	2.3	1.9	1.6	55:45	84:16
P190F	0.6	0.4	0.0	72:28	41:59
P190Y	1.6	0.9	0.2	8:92	79:21
P190W	1.1	0.5	0.1	25:75	54:46
wt	1.0	1.0	1.0	55:45	70:30

Table 4.3. Activity of P190 variants after 24 hours at 22, 28, and 34 °C with **1**, relative to wtCDMO at each temperature. The regioselectivity is based on total product at the end of 24 hrs (**1**) or 8 hrs (**4**) at 22 °C.

Substitutions at P190 did not have as large of an impact on the regioselectivity of CDMO towards **4**, compared to the regioselectivity changes towards **1**. The most significant change was seen with P190F, which produced only 41% of **5** compared to 70% with wild-type. This variant is analogous to the Q152F variant made in PAMO by the Fraaije group which caused a change in regioselectivity towards bicycloheptenone, although with a decrease in activity[35]. Interestingly, our P190F variant causes a similar shift in regioselectivity towards both **1** and **4**, also with a decrease in activity. However, by expanding the substitutions beyond phenylalanine, we were able to find much more useful variants, such as P190L.

A similar panel of variants was created at Leu191 to continue the exploration of steric effects in this region (Table 4.4). Position 191 was less amenable to substitution than position P190, especially to larger residues. Presumably, this is due to steric clashes inhibiting the binding of **1** in the active site. This is further supported by the L191F, L191Y, and L191W variants having decreased activity towards **4**. Substitutions to smaller residues at L191 favored the formation of **3** dramatically. However, L191G proved to be an unstable variant, and at 28 °C and 34 °C had no activity towards **1** or **4** (Table 4.4). Interestingly,

L191A shows a 3.7 fold increase in activity relative to wtCDMO at 22 °C, and favors the formation of **3**, with a ratio of 8:92 (**2**:**3**). Additionally, L191A shows a clear preference for (*R*)-**1**, consuming most of (*R*)-**1** before consuming any of (*S*)-**1** (Figure 4.4). This indicates a significant switch in enantiopreference compared to the wild-type enzyme which consumes both enantiomers, at almost equal rates. The substitutions at L191 also showed a steric effect towards **4**, with larger residues significantly favoring the formation of **6**, even though they were inactive towards **1**. L191Y demonstrated the largest shift towards the formation of **6** of any variant, but was less active than L191F, which showed nearly the same regioselectivity gains. In combination, the data from both **1** and **4** indicate that L191 is an important residue in controlling the regioselectivity of CDMO.

	fold activity with **1** at			Ratio of Products	
Enzyme	22 °C	28 °C	34 °C	**2**:**3**	**5**:**6**
L191G	0.6	0.0	0.0	0:100	56:44
L191A	3.7	1.1	0.4	8:92	72:28
L191V	1.0	0.8	0.4	41:59	72:28
L191I	0.8	0.7	0.4	35:65	46:54
L191M	0.4	0.2	0.1	19:81	62:38
L191F	0.0	0.0	0.0	n.a.	7:93
L191Y	0.0	0.0	0.0	n.a.	3:97
L191W	0.0	0.0	0.0	n.a.	6:94
wt	1.0	1.0	1.0	55:45	70:30

Table 4.4. Activity of L191 variants after 24 hours at 22, 28, and 34 °C with **1**, relative to wtCDMO at each temperature. The regioselectivity is based on total product at the end of 24 hrs (**1**) or 8 hrs (**4**) at 22 °C. Regioselectivity of n.a. indicates no conversion.

Since the individual amino acid substitutions at positions 190 and 191 showed promise, an attempt was made to combine some of the more interesting variants in regard to activity and regioselectivity (Table 4.5). P190G/L191A aimed to optimize the regioselectivity towards **3** by combining the two lead variants at each position that still individually retained good activity. This resulted in a strongly regioselective variant that was nearly eight times as active as wild-type based upon end point conversion data. A surprising find was that P190G/L191A no longer had any enantioselectivity preference in regard to **1** by the enzyme, but produced almost exclusively **3** (Figure 4.4). This loss in

enantioselectivity could be attributed to changes in binding of the N-protecting group of **1** caused by dramatically opening up one end of the active site. Combining the P190L and L191A substitutions also produced a highly active variant, with a shift in regioselectivity favoring **3**, albeit not as strongly as P190G/L191A. Presumably, the shift is due to a slightly smaller active site in P190L/L191A than in P190G/L191A. Both P190G/L191A and P190L/L191A also favored the formation of **6**, which, like **3**, is the kinetically preferred, normal regioisomer. This further supports the hypothesis that these two variants are acting primarily by expanding the active site and allowing the more substituted group to become the migrating group, as is typical for the chemical Baeyer-Villiger oxidation.

	fold activity with **1** at			Ratio of Products	
Enzyme	22 °C	28 °C	34 °C	**2**:**3**	**5**:**6**
P190G/L191A	7.8	3.9	1.6	2:98	14:86
P190L/L191A	7.1	3.6	1.5	10:90	28:72
P190L/L191F	0.4	0.4	0.4	78:22	7:93
P190I/L191F	0.1	0.1	0.1	70:30	19:81
wt	1.0	1.0	1.0	55:45	70:30

Table 4.5. Activity of P190/L191 variants after 24 hours at 22, 28, and 34 °C with **1**, relative to wtCDMO at each temperature. The regioselectivity is based on total product at the end of 24 hrs (**1**) or 8 hrs (**4**) at 22 °C.

Combining the L191F substitution with amino acid changes at position P190 allowed for activity to be regained towards **1**. Both P190L/L191F and P190I/L191F favored the formation of **2**. P190L/L191F also kinetically preferred (*S*)-**1** over (*R*)-**1** (Figure 4.4). This data suggest that adding bulk to the P190/L191 side of the active site promotes the formation of **2** while removing bulk promotes the formation of **3**. However, since **1** is a fairly large substrate, it is relatively easy to lose activity by adding too much bulk to the active site, thus hindering the ability for **1** to bind. Taken collectively the P190/L191 region of the active site appears to gate regioselectivity. However, despite the regioselective tenability offered by this region, it remains much simpler to promote the formation of the normal product over the abnormal product by adjusting the P190 and L191 residues.

Substitutions in positions 497-501

A third set of residues chosen for engineering were located in the 497-501 loop. The deletion of positions 498 and 499 or the introduction of a proline residue at position 497 to make this region homologous to other BVMOs (Table 4.6) proved highly detrimental to activity, resulting in little to no activity at room temperature. However, shrinking the loop region gradually via single deletions or amino acid changes to glycine resulted in several favorable variants. Of particular note are L499G and A498G/ΔL499. Both of these variants

	fold activity with **1** at			Ratio of Products	
Enzyme	22 °C	28 °C	34 °C	**2**:**3**	**5**:**6**
A497P	0.0	0.0	0.0	n.a.	81:19
A497P/ΔA498L499	0.0	0.0	0.0	n.a.	n.a.
A498G	1.2	1.4	1.3	75:25	76:24
L499G	4.1	2.7	1.1	82:18	74:26
A498G/L499G	0.4	0.5	0.3	64:36	75:25
A498G/ΔL499	3.9	1.5	0.2	79:21	94:6
ΔA498L499	0.0	0.0	0.0	n.a.	n.a.
S501A	0.7	0.6	0.4	33:67	65:35
wt	1.0	1.0	1.0	55:45	70:30

Table 4.6. Activity of variants in the 497-501 loop after 24 hours at 22, 28, and 34 °C with **1**, relative to wtCDMO at each temperature. The regioselectivity is based on total product at the end of 24 hrs (**1**) or 8 hrs (**4**) at 22 °C. Regioselectivity of n.a. indicates no conversion.

demonstrated significant activity gains relative to wild-type and favored the formation of **3**. A498G/ΔL499 also showed the highest regioselectivity of any variant towards the formation of **5**. Time course data with A498G/ΔL499 indicated that the enantioselectivity preference of the variant is slightly shifted, as it consumes more (*R*)-**1** than **3** that it produces (Figure 4.4). This indicates that A498G/ΔL499 creates both (*R*)-**2** and (*S*)-**2**. The variants in this region indicated that decreasing the size of the 497-501 loop region can, to an extent, promote the formation of **2**, possibly through some stabilization of the

tetrahedral intermediate or by allowing the substrate to bind in a conformation that promotes the abnormal product.

Combination of regional variants

Finally, we explored the combination of beneficial substitutions from each region to create the most regioselective variants possible (Table 4.7). P190G/L191A/F299A produced almost exclusively **3**, with only trace amounts of **2** observable after several days (Figure 4.4). However, it did not retain the same activity gains of P190G/L191A. In general, the addition of F299A or W325A had a detrimental effect on activity when those

	fold activity with **1** at			Ratio of Products	
Enzyme	22 °C	28 °C	34 °C	**2**:**3**	**5**:**6**
P190G/L191A/F299A	0.7	0.2	0.0	0:100	n.a.
L191A/F299A	1.3	0.4	0.1	15:85	74:26
P190G/F299A	1.4	0.6	0.1	40:60	75:25
P190I/W325A	0.1	0.0	0.0	74:26	n.a.
P190I/A498G	5.6	3.4	2.2	88:12	60:40
P190I/L499G	5.7	3.1	1.7	92:8	65:35
P190I/A498G/ΔL499	3.0	0.8	0.2	97:3	88:12
wt	1.0	1.0	1.0	55:45	70:30

Table 4.7. Activity of combined regional variants after 24 hours at 22, 28, and 34 °C with **1**, relative to wtCDMO at each temperature. The regioselectivity is based on total product at the end of 24 hrs (**1**) or 8 hrs (**4**) at 22 °C. Regioselectivity of n.a. indicates no conversion.

substitutions were combined with other amino acid changes, even if they did improve the regioselectivity of the variants. However, the combination of P190I with substitutions at either A498 or L499 showed an increase in both activity and regioselectivity. In particular, P190I/A498G/ΔL499 was three times as active as wild-type and produced **2** as 97% of its total product at room temperature after 24 hours. P190I with the individual A498G or L499G substitutions both produced variants that were significantly more active than wild-type, with good regioselectivity towards **2** as well. P190I/L499G also showed a very strong

preference for (*S*)-**1**, consuming very little of (*R*)-**1** (Figure 4.4). The combinations of these substitutions indicate that the effects of the individual variants can be combined synergistically to form better variants for the conversion of **1**. Interestingly, the combination of substitutions did not provide the same effect towards **4**, and the variants either ended up as inactive or less regioselective than some of the single variants. This result is unsurprising however, since the combination of amino acid changes selected were targeted towards altering the regioselectivity of **1**, and **4** was used a secondary benchmark to determine if any regioselectivity changes could be more widely applied.

Conclusion

CDMO was successfully engineered to allow for the effective synthesis of either regioisomer of **1** with the most effective variants consisting of multiple amino acid substitutions. In particular, the P190G/L191A and P190G/L191A/F299A showed excellent conversion towards **3**. The latter variant showed no production of **2**, but lacked the activity of wtCDMO. However, P190G/L191A demonstrated significantly more activity than wtCDMO within 24 hours, and produced 98% **3**. These variants demonstrate that removing bulk from this region can be beneficial towards the formation of the normal product. Other researchers have found that substitutions in this region can have an impact on enantioselectivity or regioselectivity[20, 35], so it is likely that these changes can be more broadly applied to other BVMOs.

Of even more interest were the variants containing the P190I substitution along with substitutions in positions A498 or L499. All three of these variants showed enhanced conversion towards **2**, with P190I/A498G/ΔL499 producing 97% **2**. While P190I/A498G and P190I/L499G did not display activity towards **2** as high as the triple variant (88% and 92% respectively), both variants had more than five-fold activity gains compared to

wtCDMO at room temperature. With these variants we were able to produce the abnormal product with high regioselectivity and increase activity relative to the wild-type enzyme, whereas previous efforts to create the abnormal product generally generated variants with poorer activity than their wild-type counterparts[22-23, 35].

Although the regions homologous to the P190/L191 and 497-501 in other BVMOs have been the studied previously, we determined that by adding steric bulk to the P190/L191 end of the active site and removing it from the 497-501 loop, the formation of **2** could be favored in CDMO. Additionally, by removing bulk from the P190/L191 end of the active site, we were able to favor the formation of **3**. This likely alters the substrate binding orientation to allow for the formation of either product, depending on where bulk has been removed or added. Attempts to rationalize the substrate preferences and selectivities have been made using computational studies. However, in the absence of a full-length crystal structure of CDMO, the catalytic orientation of FAD and $NADP^+$ relative to the active site residues is unknown and is currently our prediction within a homology model, especially given the mobility of $NADP^+$ during the catalytic cycle. Therefore, any docking studies done within the homology model containing the cofactors would be operating off of several assumptions and would need significant verification. The regions that have been identified here offer starting points for future engineering studies of CDMO that may assess the regioselectivity of substrate oxidation for compounds other than those studied here. Additionally, current work is ongoing to determine the enantioselectivity of **2** and **3** for lead variants. This enantioselectivity information may provide more insight into the substrate binding mechanism in these variants, allowing for more complete knowledge of how the targeted residues are interacting with the substrate.

Experimental

General Information

All primers used were synthesized by Integrated DNA Technologies (Coralville, IA). All other reagents were purchased from Sigma Aldrich (St. Louis, MO) unless otherwise specified. GC spectra were obtained on an Agilent Technologies 6850 GC instrument equipped with a chiral CycloSil-B column (30m x 0,32mm/0.25 µm, Agilent, Santa Clara, CA) using hydrogen as a carrier gas (flow rate 1.8 mL/min) and an FID detector (detector temperature 200°C, split ratio 25:1). GC/MS spectra were obtained on a Shimadzu QP2010 SE instrument equipped with a chiral BGB-175 column (30 m x 0.25 mm/0.25 µm, BGB Analytik), an after-column splitter, an FID detector (detector temperature 200 °C, split ratio 1:1) and GC-MS detector, using helium as a carrier gas (column flow 3.69 mL/min). The GC-MS had an interface temperature 200 °C, MS mode, EI; detector voltage, 0.2 kV; mass range, 12-250 u; scan speed, 833 u s^{-1}. All GC/MS data were acquired by a GC/MS solution software (Shimadzu).

*Chemical Baeyer-Villiger Oxidation of **1** and **4***

m-Chloroperoxybenzoic acid (417 µmole) and $NaHCO_3$ (417 µmole) were suspended in dry CH_2Cl_2 (5.3 mL) under argon for 1 hr. Dry **4** (racemic/*R*, 138 µmole) was added dropwise. The reaction was stirred at r.t. for 3 days, then washed with 10% $NaHSO_3$, followed by a wash with $NaHCO_3$ and a wash with brine. The organic layer was dried over $MgSO_4$ and evaporated down. 13 mg was recovered, reaction went to completion. Purity and product were assessed by GC.

m-Chloroperoxybenzoic acid (417 µmole) and $NaHCO_3$ (417 µmole) were suspended in dry CH_2Cl_2 (5.3 mL) under argon for 1 hr. Dry **1** (138 µmole) was added dropwise. The reaction was stirred at r.t. for 7 days, then washed with 10% $NaHSO_3$, followed by a wash with $NaHCO_3$ and a wash with brine. The organic layer was dried over $MgSO_4$ and

evaporated down. 26 mg were recovered, less than 10% of starting material had been consumed based on GCMS trace.

The reaction was tried again on **1** using a smaller peroxide, H_2O_2. **1** (121 µmole) was dissolved in ethyl acetate (360 µL). 30% H_2O_2 (242 µmole) was slowly added in and the reaction was stirred at 50 °C for 24 hrs. The reaction was then washed with 10% $NaHSO_3$, followed by a wash with $NaHCO_3$ and a wash with brine. The organic layer was dried over $MgSO_4$ and evaporated down. 26 mg were recovered, less than 10% of starting material had been consumed based on GCMS trace.

The reaction was tried once more on **1** using a stronger peroxide made *in situ*. **1** (139 µmole) and Na_2HPO_4 (417 µmole) were suspended in CH_2Cl_2 (868 µL). At 0 °C, 30% H_2O_2 (347 µmole) was added to trifluoroacetic anyhydride (417 µmole); this was then added to the suspension of β-amino ketone at r.t. and allowed to react for 1 week. The reaction was then washed with 10% $NaHSO_3$, followed by a wash with $NaHCO_3$ and a wash with brine. The organic layer was dried over $MgSO_4$ and evaporated down. 25 mg were recovered, the reaction went to 84% completion based on GCMS trace.

Creation of the Mutant Library

The mutants were prepared via overlap extension PCR using pET28a-CDMO as a template. The forward fragment for each mutant was prepared with the T7 forward primer, and the appropriate reverse mutant primer (below). The reverse fragment for each mutant was prepared with the T7 terminator primer, and the appropriate forward mutant primer (below). The PCR reactions were run using iProof High Fidelity DNA Polymerase (BioRad, Hercules, CA) under standard conditions. The PCR products were cloned into pET28a via the restriction sites *XbaI* and *HindIII* (NEB, Ipswich, MA), and the products of ligation were

then transformed into *E. coli* DH5α cells and plated onto LB agar plates containing kanamycin (30 μg/mL).

Protein Expression and Purification

Plasmid DNA was co-transformed with pGro7 into *E. coli* BL21(DE3) cells for expression. Colonies were cultured overnight at 37 °C in 50 mL of LB medium containing kanamycin (50 μg/mL) and chloramphenicol (34 μg/mL). The overnight culture was used to inoculate TB medium containing kanamycin (50 μg/mL), chloramphenicol (34 μg/mL), and 1 mM riboflavin. The culture was grown until it reached an OD(600) 0.4. Overexpression of chaperones was induced with L-arabinose to a final concentration of 2 mg/L, and the culture was expressed at 20 °C for 20 hours. Overexpression of CDMO was then induced with IPTG to a final concentration of 0.1 mM, and the culture was expressed at 20 °C for 6.5 hours. The cultures were then centrifuged for 20 minutes at 3220 x g and 4 °C. The cell pellets were stored at -20 °C until purification. The cell pellets from the culture were resuspended in buffer A (25 mM Tris-HCl, pH 7.5, 300 mM NaCl, 10 mM imidazole, 1 mM β-mercaptoethanol), then 150 μL of protease inhibitor cocktail (Sigma), 15 μL of DNAseI (Calbiochem), and 20 μM flavin adenine dinucleotide (FAD) were added. The cells were then lysed via sonication (6x with 20 second pulses and 40 second pauses) and centrifuged for 60 minutes at 4 °C and 10000 x g. The lysate was purified via Ni-NTA chromatography using buffer A and elution buffer (25 mM Tris-HCl, pH 7.5, 300 mM NaCl, 250 mM imidazole) The product fractions were combined and SDS-PAGE of the final products showed >90% purity. The final protein fractions were dialyzed against storage buffer (25 mM Tris-HCl, pH 7.5, 300 mM NaCl) before final use.

Enzymatic Activity Assay

The activity for CDMO was measured at 22, 28, and 34 °C, and used glucose dehydrogenase (GDH) from *Thermoplasma acidophilum* for NADPH regeneration. The reaction mixture for the activity assay was composed of substrate(concentration below), 200 μM NADPH, 10 U/mL GDH, 100 mM glucose in 50 mM sodium phosphate buffer, pH 8. A final concentration of 2 μM of CDMO was added to the reaction mixture. The reactions were allowed to proceed for a given amount of time, depending upon the substrate, then were quenched by mixing thoroughly with an equal volume of ethyl acetate containing 0.5 mM methyl benzoate. A sample of the organic phase was collected and analyzed by either GC or GC/MS.

Methyl(2-methyl-6-oxooctan-4-yl)carbamate. Reaction time: 24 h; Final substrate concentration: 0.5 mM. GC/MS protocol A: 110 °C (15 min) to 140 °C (0.5 min) at 1 °C min^{-1}; injection volume 1 μL; linear velocity 71.8 cm/sec; injection temperature 220 °C. (retention times for (*R*)-**1**=39.1 min, (*S*)-**1**=39.5 min, **2**=37.4 min, **3**=42.6 min).

3-methylcyclohexanone. Reaction time: 8 h; Final substrate concentration: 5 mM. GC protocol: 115 °C (5 min) to 140 °C (2 min) at 10 °C min^{-1}; injection volume 1 μL; linear velocity 34 cm/sec; injection temperature 250 °C. (retention times for (*R*)-**4**=4.9 min, (*S*)-**4**=5.1 min, **5**=11.1 min, (*R*)-**6**=11.3 min, (*S*)-**6** =11.4 min).

Hydrolysis of Products

Hydrolysis of **2** and **3** was performed by first running a 1 mL enzymatic activity assay under the conditions described previously, followed by extraction in a glass vial with ethyl acetate (1 mL x2). The organic layer was dried in a glass vial under air then resuspended in 100 μL ethanol and 100 μL 2M NaOH. The hydrolysis was left to proceed at room temperature for 1 day while shaking. The reactions were then acidified using 5M HCl, extracted using diethyl ether (400 μL x3) and dried down under air. The aqueous layer was

then neutralized using 1 M sodium phosphate buffer, pH 8.0. The reaction was extracted again using ethyl acetate (400 µL x3), dried down under air and resuspended in 500 µL hexanes for analysis on GC/MS, protocol B.

7 *and* ***8***. GC/MS protocol B: 110 °C (15 min) to 140 °C (2 min) at 1 °C min^{-1} to 180 °C (20 min) at 5 °C min^{-1}; injection volume 1 µL; linear velocity 71.8 cm/sec; injection temperature 220 °C. (retention times for (*R*)-**1**=39.1 min, (*S*)-**1**=39.5 min, **2**=37.4 min, **3**=42.6 min), (*R*)-**7**=55.4 min, (*S*)-**7**=55.8 min, (*S*)-**8**=65.2 min, (*R*)-**8**=65.8 min.

List of Primers Used

Primer design: 299 or 325 position

F299Afor: CTGGACAGTGCCACCGCCATCTGG

F299Arev: CCAGATGGCGGTGGCACTGTCCAG

F299Vfor: CTGGACAGTGTCACCGCCATCTGG

F299Vrev: CCAGATGGCGGTGACACTGTCCAG

W325Afor: GACCTCGTCCAGGACGGGGCGACCGCGCTCGGCCAGAGGATG

W325Arev: GAGCGCGGTCGCCCCGTCCTGGACGAGGTC

W325Vfor: GACCTCGTCCAGGACGGGGTGACCGCGCTCGGCCAGAGGATG

W325Vrev: GAGCGCGGTCACCCCGTCCTGGACGAGGTC

Primer design: P190X

P190Gfor: GTATGGGCACCGGCGGTCTGCACGTGGCGC

P190Grev: GCGCCACGTGCAGACCGCCGGTGCCCATAC

P190Afor: GTATGGGCACCGGCGCTCTGCACGTGGCGC

P190Arev: GCGCCACGTGCAGAGCGCCGGTGCCCATAC

P190Vfor: GTATGGGCACCGGCGTTCTGCACGTGGCGC

P190Vrev: GCGCCACGTGCAGAACGCCGGTGCCCATAC

P190Ifor: GTATGGGCACCGGCATTCTGCACGTGGCGC

P190Irev: GCGCCACGTGCAGAATGCCGGTGCCCATAC

P190Lfor: GGGCACCGGCTTACTGCACGTGG

P190Lrev: CCACGTGCAGTAAGCCGGTGCCC

P190Mfor: GTATGGGCACCGGCATGCTGCACGTGGCGC

P190Mrev: GCGCCACGTGCAGCATGCCGGTGCCCATAC

P190Ffor: GGTATGGGCACCGGCTTTCTGCACGTGGCGCAGC

P190Frev: GCTGCGCCACGTGCAGAAAGCCGGTGCCCATACC

P190Yfor: GTATGGGCACCGGCTATCTGCACGTGGCGC

P190Yrev: GCGCCACGTGCAGATAGCCGGTGCCCATAC

P190Wfor: GTATGGGCACCGGCTGGCTGCACGTGGCGC

P190Wrev: GCGCCACGTGCAGCCAGCCGGTGCCCATAC

Primer design: L191X

L191Gfor: GGGCACCGGCCCTGGGCACGTGGCGCAGCTGC

L191Grev: GCAGCTGCGCCACGTGCCCAGGGCCGGTGCCC

L191Afor: GGGCACCGGCCCTGCGCACGTGGCGCAGCTGC

L191Arev: GCAGCTGCGCCACGTGCGCAGGGCCGGTGCCC

L191Vfor: GGGCACCGGCCCTGTGCACGTGGCGCAGCTGC

L191Vrev: GCAGCTGCGCCACGTGCACAGGGCCGGTGCCC

L191Ifor: GGGCACCGGCCCTATCCACGTGGCGCAGCTGC

L191Irev: GCAGCTGCGCCACGTGGATAGGGCCGGTGCCC

L191Mfor: GGGCACCGGCCCTATGCACGTGGCGCAGCTGC

L191Mrev: GCAGCTGCGCCACGTGCATAGGGCCGGTGCCC

L191Ffor: GCACCGGCCCTTTTCACGTGGCGC

L191Frev: GCGCCACGTGAAAAGGGCCGGTGC

L191Yfor: GGGCACCGGCCCTTACCACGTGGCGCAGCTGC

L191Yrev: GCAGCTGCGCCACGTGGTAAGGGCCGGTGCCC

L191Wfor: GGGCACCGGCCCTTGGCACGTGGCGCAGCTGC

L191Wrev: GCAGCTGCGCCACGTGCCAAGGGCCGGTGCCC

Primer design: P190X/L191X

P190G/L191Afor: GTATGGGCACCGGCGGTGCGCACGTGGCGCAGC

P190G/L191Arev: GCTGCGCCACGTGCGCACCGCCGGTGCCCATAC

P190I/L191Ffor: GTATGGGCACCGGCATCTTCCACGTGGCGCAGCTGC

P190I/L191Frev: GCAGCTGCGCCACGTGGAAGATGCCGGTGCCCATAC

P190L/L191Afor: GGGCACCGGCCTTGCGCACGTGGCGCAG

P190L/L191Arev: CTGCGCCACGTGCGCAAGGCCGGTGCCC

P190L/L191Ffor: GGGCACCGGCCTTTTCCACGTGGCGCAG

P190L/L191Frev: CTGCGCCACGTGGAAAAGGCCGGTGCCC

Primer design: 497-501 loop

A497Pfor: GATGCAGGGCCCAGCTCTCGGATC

A497Prev: GATCCGAGAGCTGGGCCCTGCATC

P497AΔALfor: CCAGTTGATGCAGGGCCCAGGATCGAACATTCCCC

P497AΔALrev: GGGGAATGTTCGATCCTGGGCCCTGCATCAACTGG

A498GΔ499for: GATGCAGGGCGCAGGCGGATCGAACATTCCCC

A498GΔ499rev: GGGGAATGTTCGATCCGCCTGCGCCCTGCATC

A498G/L499Gfor: GCAGGGCGCAGGTGGCGGATCGAACATTCCCC

A498G/L499Grev: GGGGAATGTTCGATCCGCCACCTGCGCCCTGC

A498Gfor: GCAGGGCGCAGGTCTCGGATCGAACATTCCCC

A498Grev: GGGGAATGTTCGATCCGAGACCTGCGCCCTGC

L499Gfor: GCAGGGCGCAGCTGGCGGATCGAACATTCCCC

L499Grev: GGGGAATGTTCGATCCGCCAGCTGCGCCCTGC

ΔA498L499for: CCAGTTGATGCAGGGCGCAGGATCGAACATTCCCCAC

ΔA498L499rev: GTGGGGAATGTTCGATCCTGCGCCCTGCATCAACTGG

S501Afor: GATGCAGGGCGCAGCTCTCGGAGCGAACATTCCCCAC

S501Arev: GTGGGGAATGTTCGCTCCGAGAGCTGCGCCCTGCATC

Bibliography

1. Baeyer, A.; Villiger, V., Einwirkung des Caro'schen Reagens auf Ketone. *Berichte der deutschen chemischen Gesellschaft* **1899,** *32* (3), 3625-3633.

2. Criegee, R.; Schnorrenberg, W.; Becke, J., Zur Konstitution von Ketonperoxyden. *Justus Liebigs Ann. Chem.* **1949,** *565* (1), 7-21.

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

4. Crudden, C. M.; Chen, A. C.; Calhoun, L. A., A Demonstration of the Primary Stereoelectronic Effect in the Baeyer–Villiger Oxidation of α-Fluorocyclohexanones. *Angewandte Chemie International Edition* **2000,** *39* (16), 2851-2855.

5. ten Brink, G. J.; Arends, I. W. C. E.; Sheldon, R. A., The Baeyer–Villiger Reaction: New Developments toward Greener Procedures. *Chem. Rev.* **2004,** *104* (9), 4105-4124.

6. Hawthorne, M. F.; Emmons, W. D.; McCallum, K. S., A Re-examination of the Peroxyacid Cleavage of Ketones. I. Relative Migratory Aptitudes. *J. Am. Chem. Soc.* **1958,** *80* (23), 6393-6398.

7. Ager, D. J.; Prakash, I.; Schaad, D. R., 1,2-Amino Alcohols and Their Heterocyclic Derivatives as Chiral Auxiliaries in Asymmetric Synthesis.

8. Rehdorf, J.; Mihovilovic, M. D. D.; Fraaije, M. W. W.; Bornscheuer, U. T. T., Enzymatic Synthesis of Enantiomerically Pure β-Amino Ketones, β-Amino Esters, and β-Amino Alcohols with Baeyer–Villiger Monooxygenases. *Chemistry – A European Journal* **2010,** *16* (31), 9525-9535.

9. Ashfaq, M., Enantioselective Synthesis of β-amino acids: A Review. *Med. Chem.* **2015,** *5* (7).

10. Peeters, J.; Palmans, A. R. A.; Veld, M.; Scheijen, F.; Heise, A.; Meijer, E. W., Cascade Synthesis of Chiral Block Copolymers Combining Lipase Catalyzed Ring Opening Polymerization and Atom Transfer Radical Polymerization. *Biomacromolecules* **2004,** *5* (5), 1862-1868.

11. Peeters, J. W.; van Leeuwen, O.; Palmans, A. R. A.; Meijer, E. W., Lipase-Catalyzed Ring-Opening Polymerizations of 4-Substituted ε-Caprolactones: Mechanistic Considerations. *Macromolecules* **2005,** *38* (13), 5587-5592.

12. Leisch, H.; Morley, K.; Lau, P. C. K., Baeyer-Villiger Monooxygenases: more than just green chemistry. *Chem. Rev.* **2011,** *111*.

13. Leisch, H.; Morley, K.; Lau, P. C., Baeyer-Villiger monooxygenases: more than just green chemistry. *Chem. Rev.* **2011,** *111* (7), 4165-222.

14. De Gonzalo, G.; Mihovilovic, M. D.; Fraaije, M. W., Recent Developments in the Application of Baeyer-Villiger Monooxygenases as Biocatalysts. *ChemBioChem* **2010**.

15. 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.

16. Rehdorf, J.; Mihovilovic, M. D.; Bornscheuer, U. T., Exploiting the regioselectivity of Baeyer-Villiger monooxygenases for the formation of beta-amino acids and beta-amino alcohols. *Angew. Chem. Int. Ed. Engl.* **2010,** *49* (26), 4506-4508.

17. Rehdorf, J.; Lengar, A.; Bornscheuer, U. T.; Mihovilovic, M. D., Kinetic resolution of aliphatic acyclic β-hydroxyketones by recombinant whole-cell Baeyer–Villiger monooxygenases—Formation of enantiocomplementary regioisomeric esters. *Bioorg. Med. Chem. Lett.* **2009,** *19* (14), 3739-3743.

18. Fink, M. J.; Snajdrova, R.; Winninger, A.; Mihovilovic, M. D., Regio- and stereoselective synthesis of chiral nitrilolactones using Baeyer–Villiger monooxygenases. *Tetrahedron* **2016,** *72* (46), 7241-7248.

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

20. Balke, K.; Schmidt, S.; Genz, M.; Bornscheuer, U. T., Switching the Regioselectivity of a Cyclohexanone Monooxygenase toward (+)-trans-Dihydrocarvone by Rational Protein Design. *ACS Chem. Biol.* **2016,** *11* (1), 38-43.

21. Zhang, Z. G.; Parra, L. P.; Reetz, M. T., Protein engineering of stereoselective Baeyer-Villiger monooxygenases. *Chemistry - A European Journal* **2012**.

22. van Beek, H. L.; Romero, E.; Fraaije, M. W., Engineering Cyclohexanone Monooxygenase for the Production of Methyl Propanoate. *ACS Chem. Biol.* **2016**.

23. Balke, K.; Bäumgen, M.; Bornscheuer, U., Controlling the Regioselectivity of Baeyer-Villiger Monooxygenases by Mutation of Active Site Residues. *ChemBioChem* **2017**.

24. Li, G.; Fürst, M. J. L. J.; Mansouri, H. R.; Ressmann, A. K.; Ilie, A.; Rudroff, F.; Mihovilovic, M. D.; Fraaije, M. W.; Reetz, M. T., Manipulating the stereoselectivity of the thermostable Baeyer–Villiger monooxygenase TmCHMO by directed evolution. *Org. Biomol. Chem.* **2017**.

25. Chen, K.; Wu, S.; Zhu, L.; Zhang, C.; Xiang, W.; Deng, Z.; Ikeda, H.; Cane, D. E.; Zhu, D., Substitution of a Single Amino Acid Reverses the Regiospecificity of the Baeyer-Villiger Monooxygenase PntE in the Biosynthesis of the Antibiotic Pentalenolactone. *Biochemistry* **2016**.

26. van Beek, H. L.; Gonzalo, G. d.; Fraaije, M. W., Blending Baeyer–Villiger monooxygenases: using a robust BVMO as a scaffold for creating chimeric enzymes with novel catalytic properties. *Chem. Commun.* **2012**.

27. 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.

28. 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.

29. Mirza, I. A.; Yachnin, B. J.; Wang, S.; Grosse, S.; Bergeron, H.; Imura, A.; Iwaki, H.; Hasegawa, Y.; Lau, P. C. K.; 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.

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

31. Yachnin, B. J.; Sprules, T.; McEvoy, M. B.; Lau, P. C. K.; 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.

32. Mihovilovic, M. D.; Rudroff, F.; Winninger, A.; Schneider, T.; Schulz, F.; Reetz, M. T., Microbial Baeyer–Villiger Oxidation: Stereopreference and Substrate Acceptance of

Cyclohexanone Monooxygenase Mutants Prepared by Directed Evolution. *Org. Lett.* **2006,** *8* (6), 1221-1224.

33. 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. *Advanced Synthesis and Catalysis* **2007**.

34. Balke, K.; Beier, A.; Bornscheuer, U. T., Hot spots for the protein engineering of Baeyer-Villiger monooxygenases. *Biotechnol. Adv.* **2018,** *36* (1), 247-263.

35. Dudek, H. M.; de Gonzalo, G.; Torres Pazmiño, D. E.; Stępniak, P.; Wyrwicz, L. S.; Rychlewski, L.; Fraaije, M. W., Mapping the Substrate Binding Site of Phenylacetone Monooxygenase from Thermobifida fusca by Mutational Analysis. *Appl. Environ. Microbiol.* **2011,** *77* (16), 5730-5738.

36. 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), 6478-86.

37. 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.

38. Fink, M. J.; Rial, D. V.; Kapitanova, P.; Lengar, A.; Rehdorf, J.; Cheng, Q.; Rudroff, F.; Mihovilovic, M. D., Quantitative Comparison of Chiral Catalysts Selectivity and Performance: A Generic Concept Illustrated with Cyclododecanone Monooxygenase as Baeyer-Villiger Biocatalyst. *Adv. Synth. Catal.* **2012,** *354* (18), 3491-3500.

39. Fink, M. J.; Rial, D. V.; Kapitanova, P.; Lengar, A.; Rehdorf, J.; Cheng, Q.; Rudroff, F.; Mihovilovic, M. D., Quantitative comparison of chiral catalysts selectivity and performance: A generic concept illustrated with cyclododecanone monooxygenase as baeyer-villiger biocatalyst. *Advanced Synthesis and Catalysis* **2012,** *354* (18), 3491-3500.

40. Källberg, M.; Wang, H.; Wang, S.; Peng, J.; Wang, Z.; Lu, H.; Xu, J., Template-based protein structure modeling using the RaptorX web server. *Nat. Protoc.* **2012,** *7*, 1511.

41. Fürst, M. J. L. 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**.

42. Liu, Y.; Kuhlman, B., RosettaDesign server for protein design. *Nuc. Acids Res.* **2006,** *34*, W235-W238.

Chapter 5: Conclusions and Future Work

General Conclusions

Researchers have sought to both understand enzymes' native mechanisms and tailor their functions to suit other purposes through the use of protein engineering[1-4]. One example of protein engineering for the purpose of understanding a native mechanism is the engineering of Old Yellow Enzyme 1 by the Massey group to elucidate the role of potential catalytic residues within the active site[5]. Additionally, enzymes are desired for industrial uses due to their high enantio-, stereo-, or regio- specificity. However, since enzymes have evolved over thousands of years for specific roles in living systems and for high specificity towards natural substrates, they are not always ideally suited for industrial use. This creates the need for protein engineering efforts towards a wide variety of enzymes for specific applications. In this dissertation, I engineered two flavoenzymes, Old Yellow Enzyme 1 (OYE1) and cyclododecanone monooxygenase (CDMO), and assessed the resultant activity changes, as well as changes in enantio-, stereo-, or regio- selectivity. In my OYE1 engineering studies I addressed the bottleneck of heterologous protein expression for screening efforts using RAPPER and subsequently used this system to demonstrate the benefits of substrate profiling and the systematic assessment of amino acid variations. Finally, I created a range of CDMO variants that produced either product regioisomer for the oxidation of both an N-protected β-amino ketone and 3-methylcyclohexanone. Throughout my work, I used a variety of techniques to accomplish the goal of engineering flavoproteins to have altered activity towards several useful substrates.

The continuing use of in vitro transcription/translation systems

Increasing the throughput of screening enzyme variants is necessary to accelerate protein engineering efforts. In chapter two, I describe the creation of a novel method to rapidly and semi-quantitatively assess enzyme variants. This method was further

expanded, as described in chapter three, to be used on multiple substrates simultaneously, further increasing its throughput. Since I initially published the protocol for RAPPER, several other publications have reported their own use of RAPPER or related variations on the technique. These range from identification of parasite proteins to significantly speeding up the directed evolution of transaminases *in vitro*[6-7]. In particular, the work performed by Diefenbach, et al. highlights the ability to utilize the PURE system in combination with mass spectrometry using microfluidic devices[7]. This enabled the researchers to screen the library directly with their compounds of interest rather than having to use a fluorescently labeled analog for a FACS screen. By screening directly with the compound of interest, the chance for a false positive due to analog specific activity changes is eliminated. These reports highlight the utility of RAPPER and how widely it can be applied beyond its use described in this dissertation. Current work ongoing in our lab is also using aspects of RAPPER to screen a large library of OYE members against a variety of different substrates. As the costs of *in vitro* transcription/translation systems continue to decrease, RAPPER and variations on it will become a more viable option for enzyme engineering efforts.

Engineering OYE1 using multiple substrates

In chapter three, I demonstrated the necessity of testing multiple substrates when tailoring an enzyme for general activity, or making claims that a single residue is responsible for specific activity within the protein. The initial aim of screening OYE1 against multiple substrates was to locate a variant that was a generally better enzyme, either in activity or in selectivity. However, in the course of this research, it became apparent that with OYE1, due to the complex nature of enzyme dynamics, each substrate displayed different and highly specific interactions with each variant. By using RAPPER to screen OYE1 variants against up to eight substrates, I discovered that the positive effect of

the amino acid substitution P295A, which appeared to be a lead candidate for increasing activity of OYE1, was highly substrate dependent. Additionally, W116, which was previously indicated to provide control over stereoselectivity, also showed significant substrate dependence. These studies highlight once more that in enzyme engineering, '*you get what you screen for*'[8] and that using only one model substrate can lead to misleading results. This tenet has been carried forward to current work ongoing in the lab; OYE libraries are being screened against a wide library of substrates to minimize the chances of missing interesting or novel activity.

Two of the more interesting variants found in the course of these studies that could be analyzed in more depth are P295A and the alanine variants at the H191 and N194 positions. The P295A variant showed catalytic gains towards carvone and ketoisophorone, exceeding even a previous circularly permuted variant (cp303) discovered in our lab[9-10]. These gains were highly specific to P295A and did not materialize from any other substitution at the P295 position. The hypothesis that the cause of this catalytic gain was due to increases in the flexibility of the P295-containing loop, which may be the rate limiting step for these substrates, can be further explored with rapid enzyme kinetic measurements. These measurements could be taken via stopped-flow to help determine the mechanistic steps for substrates such as carvone, as well as for substrates that did not show an increase in activity with amino acid changes at the P295 position, such as α-methyl-*trans*-cinnamaldehyde. By applying these measurements to several P295 variants and several substrates, the cause of this gain in catalytic activity could be determined. In regard to H191A and N194A, these demonstrated vastly different activities with (*S*)-carvone versus methyl-2-(hydroxymethyl)acrylate. The cause of this variation in activity is likely due to different binding orientations within the active site, which could be further explored

by rigorous *in silico* studies or by attempting to co-crystallize these substrates within the active site of OYE1.

Future studies on cyclododecanone monooxygenase

BVMOs offer an attractive alternative to the traditional Baeyer-Villiger reaction for several reasons, including their use of molecular oxygen as well as the ability to tune their regioselectivity and enantioselectivity through protein engineering efforts. I was able to successfully tune the regioselectivity of CDMO towards **1** (Figure 5.1) by engineering multiple regions of the protein. Generally, removing bulk from the P190/L191 region favored the formation of the β-amino alkylacetate, while removing bulk from or increasing the flexibility of the 497-501 loop region favored the formation of the β-amino acid ester.

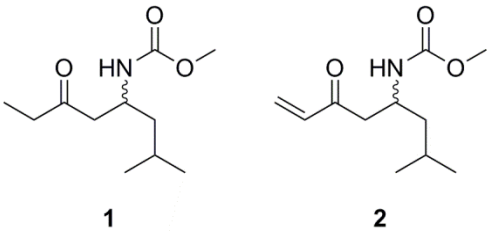


Figure 5.1 Structures of N-protected β-amino ketone and possible substrate analog.

The 497-501 loop region in CDMO corresponds to loops in other BVMOs that have been hypothesized to be involved in substrate specificity. The reason for this is that in solved crystal structures of other BVMOs, these loops occupy space adjacent to the arginine (391 in CDMO) that stabilizes the Criegee intermediate[11]. Therefore if bulk is added or removed from this loop, the binding properties of the substrate should change based on steric effects, influencing which group is in the migratory position. This argument relies heavily on steric bulk being the leading cause of the rearrangement of the substrate, not electronic effects. One way in which this hypothesis could be tested is through substrate engineering instead of protein engineering. In particular, a substrate analog for **1** could be utilized to determine whether the effects observed were due to sterics (as put forth in chapter four), or if electronic effects had a more significant role. The most useful analog to test this hypothesis is shown as compound

2 in Figure 5.1. The key difference between **1** and **2** is the presence of a vinyl group instead of an ethyl group. By substituting the ethyl substituent on the ketone with a vinyl group, the expected migratory aptitudes of the substituents are reversed due to the electronic properties of the vinyl substituent[12]. Therefore, if the regioselectivities remain the same between **1** and **2** for the variants discussed in chapter four, as well as for wtCDMO, the primary effect is likely sterics. However, if the regioselectivity is significantly altered, it is likely that electronics plays a more significant role in substrate orientation than currently hypothesized. Since **2** is not commercially available, it will need to be synthesized, which has proven to be a non-trivial task thus far.

With the recent discovery and crystallization of PockeMO, a thermostable BVMO with a large active site similar to CDMO[13], my work on CDMO can be expanded upon in several ways. One strategy would be to apply the substitutions found to be advantageous in CDMO to PockeMO to determine if these effects can be more widely applied to analogous enzymes within the BVMO family. If this proves successful, it would generate a more thermostable enzyme with desirable selectivities for industrial purposes. The primary advantage to this strategy would be that starting with a stable enzyme generally leads to more stable variants. However, while PockeMO and CDMO are fairly similar (40% sequence identity, 52% sequence similarity), they do have regions of non-homology, including within the 497-501 region (CDMO numbering). Therefore it is entirely possible that the mutations applied to CDMO will not be effective when applied to PockeMO. An alternate strategy might be to attempt directed evolution on CDMO to make it more thermostable, and thus more appealing for industrial use by making it more similar to PockeMO while retaining the beneficial substitutions outlined in chapter four. Another possible strategy would be to solve the crystal structure of CDMO, although efforts to date in our lab have failed in this attempt. A crystal structure would help elucidate what role the amino acid changes

discussed in chapter four had, as well as allowing for some *in silico* work in regards to docking studies with **1**.

Beyond efforts towards crystallizing CDMO or making it more thermostable, in depth mechanistic and kinetic studies could be performed on **1** with CDMO to further elucidate its interaction with the enzyme. These studies could be performed with the lead variants, as well as wild-type CDMO, to determine if there are any differences in the mechanisms between the variants that allow for the different regioisomers to be produced. This could also be explored with variants that change enantioselectivity to determine if the reaction has a new rate-limiting step when a different enantiomer is preferred or a different regioisomer is produced. This would help rationalize the effects of the substitutions made in chapter four of this thesis in a way that would provide guidance for future studies.

Bibliography

1. Farinas, E. T.; Bulter, T.; Arnold, F. H., Directed enzyme evolution. *Curr. Opin. Biotechnol.* **2001,** *12* (6), 545-551.

2. Brannigan, J. A.; Wilkinson, A. J., Protein engineering 20 years on. *Nat. Rev. Mol. Cell Biol.* **2002,** *3*, 964.

3. Renata, H.; Wang, Z. J.; Arnold, F. H., Expanding the Enzyme Universe: Accessing Non-Natural Reactions by Mechanism-Guided Directed Evolution. *Angewandte Chemie International Edition* **2015,** *54* (11), 3351-3367.

4. Porter, J. L.; Rusli, R. A.; Ollis, D. L., Directed Evolution of Enzymes for Industrial Biocatalysis. *ChemBioChem* **2016,** *17* (3), 197-203.

5. Kohli, R. M.; Massey, V., The Oxidative Half-reaction of Old Yellow Enzyme: The Role of Tyrosine 196. *J. Biol. Chem.* **1998,** *273* (49), 32763-32770.

6. Catherine, C.; Lee, S.-W.; Ju, J. W.; Kim, H.-C.; Shin, H.-I.; Kim, Y. J.; Kim, D.-M., Cell-Free Expression and In Situ Immobilization of Parasite Proteins from Clonorchis sinensis for Rapid Identification of Antigenic Candidates. *PLoS One* **2015,** *10* (11).

7. Diefenbach, X. W.; Farasat, I.; Guetschow, E. D.; Welch, C. J.; Kennedy, R. T.; Sun, S.; Moore, J. C., Enabling Biocatalysis by High-Throughput Protein Engineering Using Droplet Microfluidics Coupled to Mass Spectrometry. *ACS Omega* **2018,** *3* (2), 1498-1508.

8. You, L.; Arnold, F., Directed evolution of subtilisin E in Bacillus subtilis to enhance total activity in aqueous dimethylformamide. *Protein Eng. Des. Sel.* **1996,** *9* (1), 77-83.

9. Daugherty, A. B.; Govindarajan, S.; Lutz, S., Improved biocatalysts from a synthetic circular permutation library of the flavin-dependent oxidoreductase old yellow enzyme. *J. Am. Chem. Soc.* **2013,** *135* (38), 14425-32.

10. Quertinmont, L. T.; Orru, R.; Lutz, S., RApid Parallel Protein EvaluatoR (RAPPER), from gene to enzyme function in one day. *Chem. Commun.* **2015,** (51), 122-124.

11. Zhang, Z. G.; Parra, L. P.; Reetz, M. T., Protein engineering of stereoselective Baeyer-Villiger monooxygenases. *Chemistry - A European Journal* **2012**.

12. Krief, A.; Laboureur, J. L., Ring Expansion of Cyclenones: an Unusual Regioselectivity. *J. Chem. Soc., Chem. Commun* **1986**.

13. Fürst, M. J. L. 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**.