# New to Nature C–C Bond Forming Cyclases: Pushing the Boundaries of Ring Forming Reactions

Thesis by

Daniel Joseph Wackelin

In Partial Fulfillment of the Requirements for

the Degree of

Doctor of Philosophy

**Caltech**

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2023

(Defended August 31$^{st}$, 2023)

© 2023

Daniel Joseph Wackelin

ORCID: 0000-0001-8189-6985

# ACKNOWLEDGEMENTS

First, I would like to thank my advisor, Prof. Frances Arnold, for her help and guidance during my thesis work. Frances is always our harshest critic, but at the same time our strongest ally, and I would not be where I am today without her support. Thank you for challenging me to work on exciting projects and for allowing me the freedom to pursue them thoroughly. I am also thankful for my committee members, Prof. Harry Gray, Prof. Mikhail Shapiro, and Prof. Lu Wei for their questions and insights throughout the course of my doctoral studies.

I am grateful to the entirety of the Arnold lab, past and present, for the awesome environment they have cultivated. Our lab manager Dr. Sabine Brinkmann-Chen holds the lab together and is invaluable for research, writing, and life advice. Without her presence the lab would not be what it is today. I am indebted to Dr. Ruijie Zhang and Dr. Kai Chen for mentoring me during my rotation and beyond. Without their guidance, my time at Caltech would have been very different and I am so thankful they took the time to train me and work with me. I have also been fortunate enough to work with many brilliant scientists, including Dr. Runze Mao, Dr. Doris Taylor, Shilong Gao, Dr. Anuvab Das, Dr. Kathleen Sicinski, Dr. Cooper Jamieson, Dr. Xiongyi Huang, Dr. David Miller, Dr. Noah Dunham, Yueming Long, and Ravi Lal. You all, along with the entirety of the Arnold lab, have made graduate school a productive and fulfilling experience.

I would not be at Caltech without a tremendous amount of help during my undergraduate studies, and I am appreciative of everyone I worked with. I am especially grateful to my undergraduate advisor, Prof. Yongku Cho, who gave me a strong foundation in both technical skills and the research process that has proved invaluable as I progressed in my studies. Prof. Peter Tessier encouraged me to pursue a PhD and provided another perspective that helped me decide what to pursue in graduate school.

# ABSTRACT

Biocatalysts have shown themselves to be extremely powerful for the synthesis of pharmaceuticals, fragrances, and fine chemicals, providing products with high yields and selectivities. Recently, new-to-nature biocatalysis has received increased attention, allowing for the benefits of biocatalysis to be applied to reactions that were previously the sole domain of chemocatalysts. Engineers have begun to develop enzymes that catalyze new-to-nature C–C bond forming cyclisation reactions, which are quite powerful due to their ability to build the carbon skeleton of molecules. Despite this, this class of enzymes is limited in scope. This thesis details the expansion of C–C bond forming cyclases, including expanding the scope of cytochrome P411 cyclopropanation and an intramolecular C–H functionalization strategy for the synthesis of diverse rings. Chapter 1 introduces biocatalysis and its recent applications, especially as they apply to new-to-nature C–C bond forming cyclisation reactions. Chapter 2 shows the development of a cytochrome P411 that catalyzes the enantio- and diastero-specific synthesis of 1,2,3-polysubstituted cyclopropanes. Using directed evolution, this carbene transferase was evolved to react with internal alkenes and build two C–C bonds, expanding the scope and specificity of cyclopropanation reactions. Chapter 3 describes the expansion of this biocatalytic system toward the synthesis of stereoconvergent products, enabling more efficient synthesis from non-diasteropure starting materials. Chapter 4 details the evolution of a cytochrome P411 to perform an intramolecular C–H functionalization using diazo compounds, making a variety of differently sized rings with different molecular geometries. In summary, this work addresses the need for expansion of new-to-nature C–C bond forming cyclisation reactions and provides a guide for expanding new-to-nature reactions to their full potential.

# PUBLISHED CONTENT AND CONTRIBUTIONS

Mao, R.[†]; **Wackelin, D. J.**[†]; Jamieson, C. S.; Rogge, T.; Gao, S.; Das, A.; Taylor, D.M.; Houk, K. N.; Arnold, F. H. Enantio- and Diastereoenriched Enzymatic Synthesis of 1,2,3-Polysubstituted Cyclopropanes from (*Z/E*)-Trisubstituted Enol Acetates. *J. Am. Chem. Soc.* **2023**, *145*, *29*, 16176–16185. **DOI**: 10.1021/jacs.3c04870

R.M. conceived and designed the overall project with F.H.A. providing guidance. R.M. performed the initial screening of heme proteins and the directed evolution experiments. R.M. and D.J.W. studied the substrate scope. R.M., D.J.W., and S.G. synthesized the olefin substrates 1; R.M. and D.J.W. synthesized diazoacetonitrile. C.S.J. and T.R. carried out the computational studies with K.N.H providing guidance. A.D. conducted crystallization of 3j and 3k. D.M.T. performed sequencing of SSM libraries based on IC-G3. R.M., D.J.W., C.S.J., and F.H.A wrote the manuscript with input from all authors. R.M. and D.J.W. contributed equally to this work.

Mao, R.[†]; Taylor, D. M.[†]; **Wackelin, D. J.**[†]; Rogge, T.; Wu, S. J.; Sicinski, K. S.; Arnold, F. H. Biocatalytic, Stereoconvergent Alkylation of (*Z/E*)-Trisubstituted Silyl Enol Ethers. *Manuscript submitted for publication,* **2023**.

R.M. conceptualized and designed the overall project under the guidance of F.H.A.. R.M. and D.J.W. carried out the initial screening of heme proteins. D.M.T. performed the directed evolution experiments, with support from S.J.W.. R.M., and D.J.W. investigated the substrate scope and reaction mechanism. K.M.S. purified and obtained both (Z)-1a and (E)-1a. T.R. carried out the computational studies with K.N.H providing guidance. R.M. and F.H.A wrote the manuscript with input from all authors. R.M., D.M.T., and D.J.W. contributed equally to this work.

**Wackelin, D. J.**[†]; Mao, R. [†]; Sicinski, K. M.; Zhao, Y.; Das, A.; Chen, K.; Arnold, F. H.

Enzymatic Assembly of Diverse Lactone Structures: An Intramolecular C–H

Functionalization Strategy. *Manuscript submitted for publication,* **2023**.

D.J.W. and K.C. conceived and designed the overall project under F.H.A.'s guidance. D.J.W. and K.C. conducted the initial screening of heme proteins and executed the directed evolution experiments with R.M. and Y.Z.'s assistance. D.J.W., R.M., and K.C. explored and studied the substrate scope and applications, with assistance from K.M.S for product purification. A.D. and D.J.W. performed the crystallization of products. D.J.W., R.M., K.C., and F.H.A. wrote the manuscript, incorporating input from all contributors. D.J.W. and R.M. contributed equally to this work.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# NOMENCLATURE

| | |
|---|---|
| Ac | Acetate |
| ACN | Acetonitrile |
| ALA | 5-Aminolevulinic acid |
| CN | Cyano |
| e.e | Enantiomeric excess |
| Et | Ethyl |
| EtOH | Ethanol |
| GC | Gas chromatography |
| GC-FID | Gas chromatography with flame ionization detection |
| h | Hours |
| HB | Hyperbroth medium with 0.1 mg/mL ampicillin. |
| HPLC | High performance liquid chromatography |
| HRMS | High resolution mass spectrometry |
| IPTG | Isopropyl β-D-1-thiogalactopyranoside |
| LB | Luria-Bertani medium with 0.1 mg/mL ampicillin. |
| LC-MS | Liquid chromatography with mass spectrometry |
| M9-N | M9 minimal media without nitrogen source added |
| Me | Methyl |
| $OD_{600}$ | Optical density of a cell suspension measured at a wavelength of 600 nm |
| P450 | Cysteine-ligated cytochrome P450 monooxygenase |
| P411 | Serine-ligated cytochrome P450 monooxygenase |
| Piv | Pivalate |
| RT (or r.t.) | Room temperature |
| TON | Turnover number |
| TTN | Total turnover number |

*C h a p t e r   1*

# NON-NATURAL C–C BOND FORMING CYCLASES

**Abstract**

Biocatalytic processes have been used for the efficient and sustainable production of pharmaceuticals, fine chemicals, fragrances, and other products. They have shown themselves to be powerful complements to chemocatalysts, attaining high levels of selectivity and yield with mild reaction conditions. More recently, engineers have developed enzymes which perform new-to-nature reactions, which mimic synthetic catalysts while maintaining their high selectivity and yield and mild reaction conditions. C–C bond forming and cyclization reactions are critically important to modern synthesis, but new-to-nature biocatalytic C–C bond forming cyclization reactions are limited in scope. This chapter discusses examples of new-to-nature C–C bond forming cyclases, including the current state of biocatalytic cyclopropanation, radical cyclization, and pericyclization.

## 1.1 Introduction

Chemical synthesis is the backbone of modern life. Typically, synthesis involves harsh conditions, halogenated solvents, and poor atom economy, undesirable attributes that contribute to reduced yields of products and negative environmental impacts.[1,2] Many chemists have devoted their careers to discovering and developing novel small-molecule catalysts to improve the activity, selectivity, and yield of a broad range of different chemical reactions. Oftentimes, these catalysts rely on precious metals, such as platinum, rhodium, or ruthenium, to achieve the best selectivities and yields.[3,4] Unfortunately, the price of precious metal catalysts has been increasing, and their acquisition is becoming more environmentally costly, both during mining and during the processing of the crude ores.[5,6] As an alternative to chemocatalysis, biocatalysis has been proposed to solve many of these pressing issues. Biocatalysis typically employs much milder conditions, operates in aqueous solvents, and uses earth-abundant metals to perform highly selective and powerful transformations.[7,8] Biocatalysis has already proven itself as a powerful tool to complement small molecule catalysis in chemical synthesis, as seen most starkly in the pharmaceutical industry.[9] An early example of this synergy is in the manufacture of sitagliptin, an antidiabetic medication.[10] An evolved transaminase was developed to synthesize a broad array of chiral amines, including the aforementioned drug molecule. The biocatalytic process provides sitagliptin with a 10 to 13% increase in overall yield, a 53% increase in productivity (kg/l per day), a 19% reduction in total waste, the elimination of all heavy metals, and a reduction in total manufacturing cost when compared with the traditional chemocatalytic process. Even more impressive is the total synthesis of drug molecules through biocatalysis, completely replacing the traditional

synthetic route. An investigational HIV treatment, islatravir, was synthesized through a nine-enzyme, three-step biocatalytic cascade, which provided the drug molecule with an atom economy that far exceeded that of previous syntheses.[11] The product was attained in 51% overall yield with half the number of steps of previous syntheses under mild, aqueous conditions and without intermediate purification.
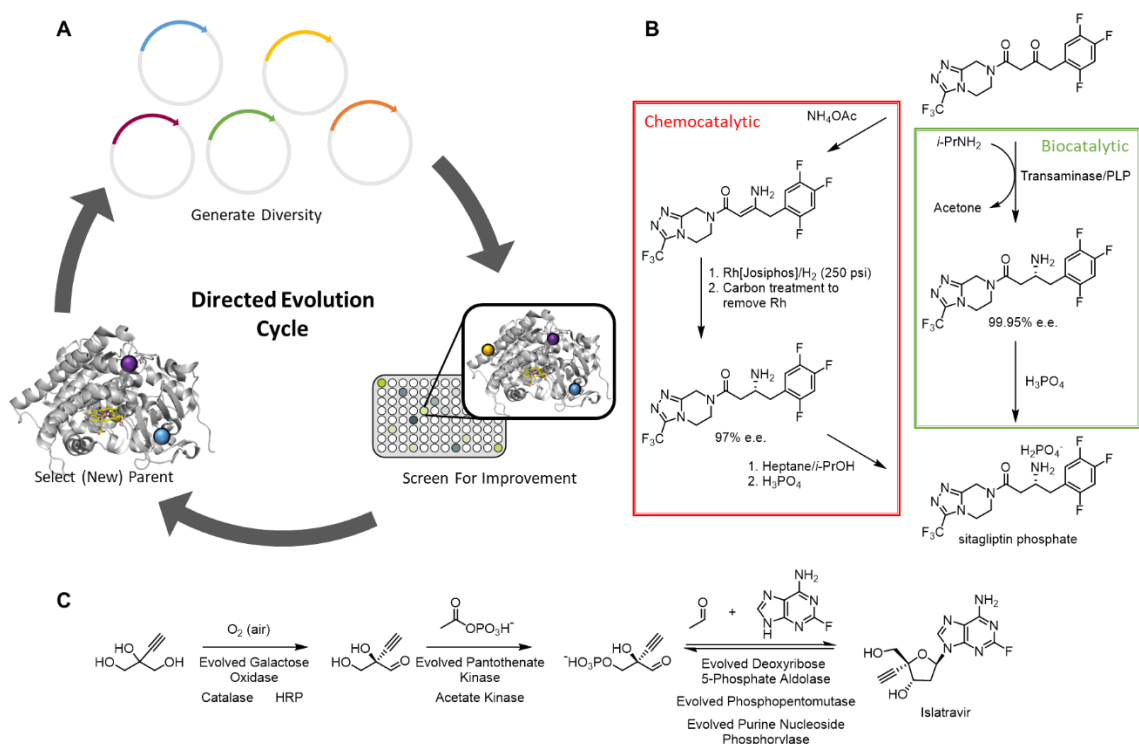


**Figure 1-1. Directed evolution cycle and application to industry.** (A) The directed evolution cycle. Starting with a parent enzyme, diversity is generated, and the mutated enzymes are screened. The best enzyme is selected to become the parent of the next round. (B) Industrial enzymatic synthesis of Sitagliptin. (C) Industrial enzymatic synthesis of Islatravir.

As seen in these examples, the development of evolutionary methods to evolve proteins offers a significant advantage to biocatalysis, allowing engineers to tailor their catalyst to the

requirements of a process, rather than tailoring the process to the requirements of the catalyst. In both cases, directed evolution was critical to the success of these biocatalytic syntheses because it improved the proteins for a desired function, such as yield, substrate loading, and selectivity of the desired reactions. Regardless of the specific function, be it yield, selectivity, or solvent tolerance, for example, directed evolution starts with some initial amount of the function, in the case of biocatalysis this is most often enzyme activity under a certain set of conditions, and through iterative rounds of mutagenesis and screening this function is improved.[12] Compared to natural evolution's timescale measured in the millions of years, directed evolution screens large libraries of enzyme variants in days to weeks, and thus directed evolution campaigns often span only a handful of months.[13] Enzymes have been evolved for higher activity and greater selectivity, as well as for other useful properties like higher solvent tolerance and substrate loading. Numerous examples have shown that with an initial activity, enzymes can be reliably evolved for whatever specific properties are needed in a particular system.[14,15] Recently, engineers have found enzymes that perform reactions well outside the scope of their native function.[16] These non-natural functions have similarly been shown to be highly evolvable, attaining high levels of activity and selectivity on reactions that nature either has not found useful or has not had an incentive to discover, but human chemists can harness for productive uses.[17,18]

One particularly striking example of this is the insertion of a carbene into a C–H bond and the formation of a C–C bond in its place.[19] C–C bond forming reactions are among the most useful in modern organic chemistry, critically important for their role in building the carbon skeleton of molecules. This method is limited in utility, however, in that, by design, it

appends a relatively small ethyl ester onto a larger molecule. This enzymatic C–H insertion method does not contribute toward building one of the more complicated aspects of difficult carbon skeletons, ring structures. Natural products and drug molecules often have rings, which make their syntheses significantly more difficult.[20] Synthetic chemists have developed a number of strategies for ring synthesis to mollify this difficulty, including radical cyclization, cross coupling, pericyclization, group transfer chemistry, and annulation reactions. Unfortunately, these methods suffer from the same issues with selectivity and sustainability as other chemocatalysed reactions.[21,22] Non-natural enzymatic methods to form rings through C–C disconnections would be extremely powerful, with high activity and selectivity allowing for more streamlined syntheses with simpler purification steps. Despite this, engineers have developed select few non-natural C–C bond forming cyclases. In this thesis, I will first outline previously developed new-to-nature carbene cyclopropanation, radical cyclization, and pericyclization reactions. I will then cover the efforts I have made to expand the scope of C–C bond forming cyclases.

## 1.2 Carbene cyclopropanation

Carbene cyclopropanation reactions utilize a high energy carbene intermediate to react with an olefin and generate a cyclopropane. Non-natural enzymatic cyclopropanation has been dominated by heme proteins, using diazo compounds to generate and transfer the carbene to a variety of olefins selectively and with high yields. The first example of this type of reaction in a non-natural enzyme was reported by Coelho *et al.*, who demonstrated that a cytochrome P450 and several other heme proteins can take ethyl diazoacetate to ethyl acetate iron carbene and transfer this carbene to styrene and styrene derivatives; engineered P450s did so with

high activity and selectivity.[23] While the wild-type P450 enzyme was not a good catalyst for the transformation, exhibiting only 1% yield and poor enantioselectivity, it proved to be quite evolvable for this reaction. A single active-site mutation afforded the trans product with 65% yield, 1:99 cis:trans, and -96% e.e. To attain the cis product, a variant fourteen mutations from obtained similar selectivity and yield (59% yield, 92:8 cis:trans, -97% e.e.). This demonstrated that the enzyme was amenable to the accumulation of mutations and could afford products that were thermodynamically unfavorable. Since this initial report, enzymatic cyclopropanation has been expanded significantly. Other non-P450 enzymes have been developed that similarly catalyze cyclopropanation reactions, including gas-binding hemoproteins that should be more amenable to evolution due to their greater thermostability.[24-27] Using P450s as well as these thermostable proteins, Knight *et al*. demonstrated that not only can enzymes transfer carbenes to unactivated alkenes, but also that, with enough effort and engineering, biocatalysis can afford all possible enantiomers of a given reaction selectively.[28] Using 1-octene and four evolved enzymes, they attained between 100 and 490 TTN, <1:99 and 89:11 d.r., and 96% and >99% e.e., demonstrating definitively that all possible products of a reaction can be made with high yields, high diastereoselectivity, and high enantioselectivity.

Enzymes have also been demonstrated to perform reactions that are highly challenging with small-molecule catalysis. Heteroatom substituted alkenes are a significant challenge for small-molecule catalysts, as heteroatoms are often good ligands for the metals used in the transformation, and carbene insertion into these heteroatoms often has a lower activation barrier compared to the cyclopropanation reactions. Despite these challenges, cytochromes

P450 were developed to access N-, O-, and S-substituted cyclopropanes with high diastereoselectivities and enantioselectivities with higher total turnovers than previously reported chemocatalysts.[29] These enzymes afforded both the *cis*- and *trans*-cyclopropane products of N-, O-, and S-substituted cyclopropanes highly selectively, underscoring the ability of enzymes to control the stereoselectivity even in cases where chemoselectivity is a major challenge, and also attained up to 40,000 TTN, placing it amongst the most active carbene transferases and small molecule carbene transfer catalysts.[30-32]
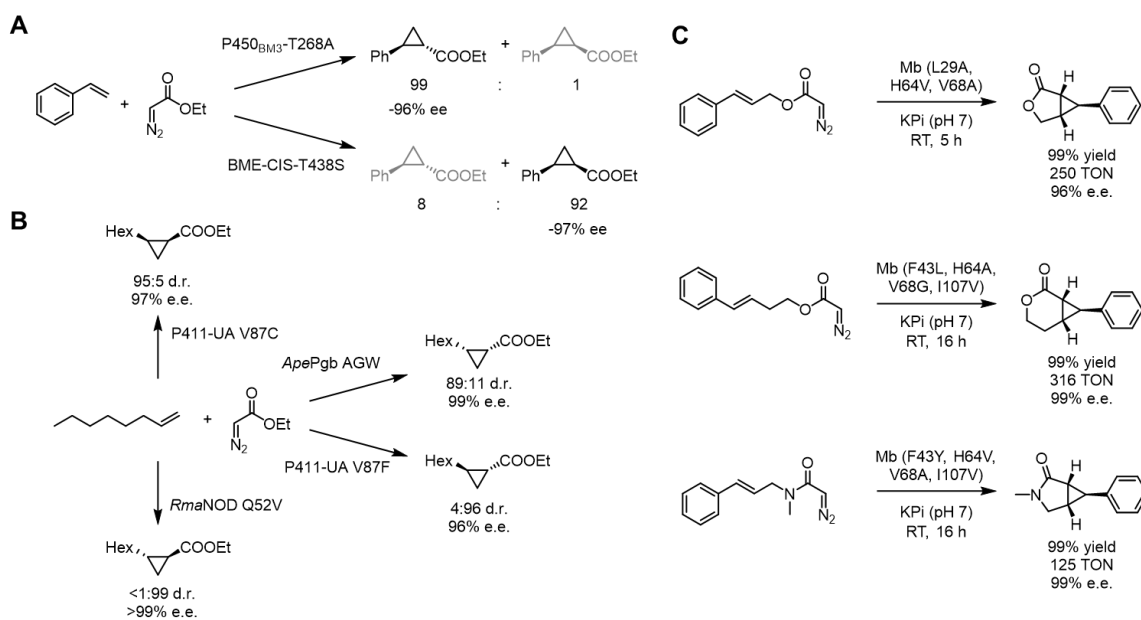


**Figure 1-2. Development of carbene cyclopropanases.** (A) Initial report of cyclopropanation with an iron-heme enzyme, which was evolved for two diastereomers of styrene cyclopropanation. (B) Cyclopropanation of unactivated alkenes, which was developed to form every possible product of the reaction using different evolved enzymes. (C) Intramolecular cyclopropanation using evolved enzymes simultaneously builds two rings.

These reactions could also be run intramolecularly, forming two rings simultaneously. These enzymatic intramolecular reactions could be quite challenging due to the limited active site space in many engineered heme proteins. Despite this, engineered myoglobins using diazo esters with aryl alkene tails were shown to perform the intramolecular cyclopropanation reaction to form 5-membered γ-lactones fused to the newly formed cyclopropanes.[33] Evolved variants attained high yields and enantioselectivites on aryl alkene substrates, though they struggled with selectivity on non-aryl alkenes. Further evolution using substrates with an alkyl chain extended by one methylene similarly allowed for formation of 6-membered cyclopropyl-δ-lactones.[34] They similarly attained moderate to high yields to form these products, with high enantioselectivity. Engineered myoglobins were not limited to forming lactones, and similar engineered variants utilized diazo amides intramolecularly to form cyclopropyl-γ-lactams with high yields and enantioselectivities.[35] This is notable as not only are all these enzymes forming two rings simultaneously, the lactam-forming enzyme can also use an unprotected amide to form an unprotected lactam, albeit with reduced yield.

## 1.3 Radical cyclization

Another powerful method to generate larger rings, of greater than four-members, is radical cyclization, which encompasses a broad set of reactions that all involve the formation of a radical and subsequent cyclisation. It was recently discovered that 'ene'-reductases, which naturally reduce alkenes using a proton and a hydride, upon irradiation with light form a carbon-centered radical. This carbon-centered radical can then cyclize with an alkene to form a new C–C bond and a new ring. Light is required, as flavin hydroquinone ($FMN_{hq}$), the

cofactor for these enzymes, is a modest single electron reductant ($E_{1/2}$ = -0.45 V versus saturated calomel electrode (SCE)), and as such electron transfer to the substrates, α-chloroamides, ($E_{p/2}^{red}$ = -1.65 V versus SCE), is thermodynamically unlikely. Using light, this cofactor enters an excited state with a much greater reduction potential ($E_{1/2}$* = -2.26 V versus SCE) that can affect the electron transfer. Once the radical is formed, the enzyme must control this high-energy intermediate and bring the alkene in close contact to allow for reaction, without first performing their natural function and reducing the alkene to an alkane. Using several different enzymes, methyl-protected lactams varying in size from 5–8 members were formed with good yields and selectivities.[36] This work was further extended to allow for the maintenance of the alkene functionality after the reaction[37]. By adding a labile trimethyl silyl (TMS) group adjacent to the alkene, upon addition of the carbon radical to the alkene the TMS group leaves and the alkene migrates by one carbon, terminating the radical by β-scission or polar crossover rather than hydrogen atom transfer. Through directed evolution, an 'ene'-reductase that performs this transformation with good yields and high levels of enantioselectivity was developed. While these examples were limited to lactam formation, it would be desirable to form all-carbon rings as well. Accordingly, unstabilized alkyl iodides were used to form the carbon centered radical, allowing for a cyclization onto a Michael acceptor.[38] While more limited in alkene scope, requiring an electron deficient alkene, these reactions demonstrated high yields. Unfortunately, these reactions also generally had lower levels of enantioselectivity due to the instability of the alkyl radical intermediate.
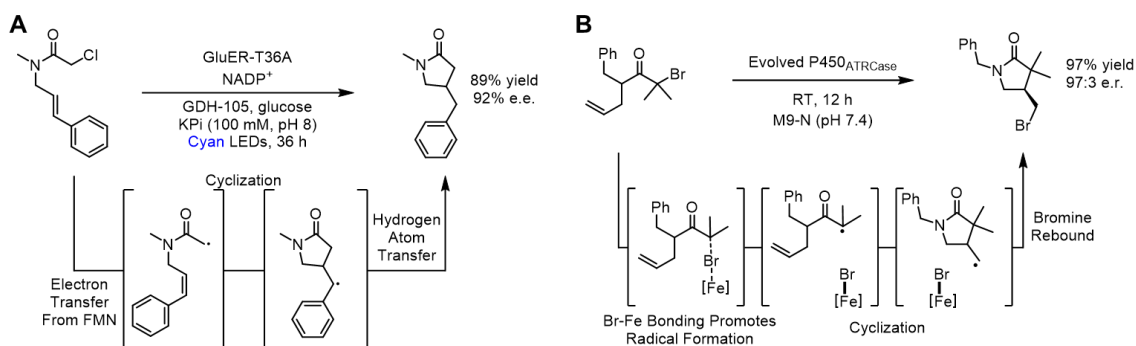
**Figure 1-3. Mechanisms of radical cyclase cyclization.** (A) Using 'ene'-reductases, light excites FMN and allows for the scission of the C–Cl bond and the formation of the carbon radical. Subsequent cyclization and hydrogen atom transfer leads to the product. (B) Using heme enzymes, interaction between the heme and the bromine atom weakens the C–Br bond, which allows for bond cleavage and carbon radical formation. Subsequent cyclization and bromine atom transfer leads to the product. [Fe] = heme.

Other systems have been developed to leverage a similar intermediate and control a radical for cyclization. One recent example utilized an engineered heme-containing cytochrome P450 to perform a similar reaction without the addition of light to drive the reaction forward.[39] Despite the lack of light, these enzymes are still able to achieve high levels of activity and selectivity. This is due to the ground state reactivity of the iron and the enzymes' ability to form the Fe–Br bond, lowering the activation energy, meaning light is not required to perform the C–Br bond scission, a significant advance over other methods. Interestingly, in general, these enzymes quench the radical through a complementary atom transfer mechanism; the bromine in the starting material is used to quench the radical after the cyclization and is present in the product post-reaction. This offers access to a different set of products despite going through similar reaction pathways and underscores the importance of

carefully choosing your biocatalyst and the strength of the diversity of enzymes; while either of these enzymes could likely be evolved to form the other's product, it is much easier to simply express the gene that leads to the desired product and avoid potentially months of evolution.

## 1.4 Pericyclization

Another class of non-natural C–C bond forming cyclases are pericyclases, enzymes that feature a concerted reaction with a cyclic flow of electrons in a single transition state. Pericyclases were one of the earliest examples of non-natural enzymes, with a catalytic antibody developed for the catalysis of the Diels-Alder reaction to form a cyclohexene first published in 1989.[40] This was considered theoretically feasible due to the high entropic barrier, typically -30 to -40 cal $K^{-1}$ $mol^{-1}$ in this reaction, allowing for the use of an antibody as an "entropic trap" to overcome this barrier. Due to the high degree of order of the transition state and the similarity of the transition state to the product, antibodies were matured against a product analogue, which afforded the desired product with greater than 50 total turnovers. Since this first report, it was immediately desirable to have some control over the reaction outcome and generate different possible stereoisomers of the product selectively. Using density functional theory calculations, the transition states of the *endo-* and *exo-*products were calculated, then transition state analogues were designed and antibodies were matured against them.[41] Two of twenty-two and four of twenty-five antibodies isolated from the mice were selective for *endo-* and *exo-*product formation, respectively. From these, the most selective from each set were selected for further analysis, and they were both able to deliver the products with >98% enantioselectivity.
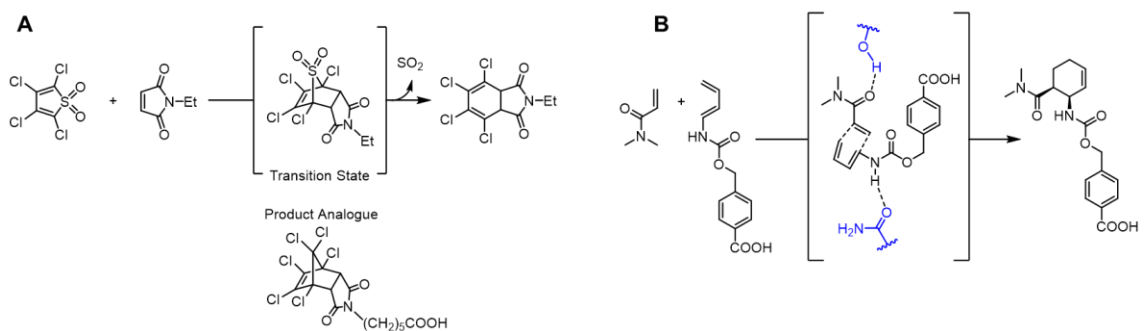
**Figure 1-4. New-to-nature pericyclases, evolved and predicted.** (A) Initial report of an antibody-catalyzed Diels-Alder reaction. (B) *In silico* designed transition state. Substrates are placed by a hydrogen bond donor and acceptor in the protein and supported by the protein scaffold.

Since these early reports, it was clear that transition state positioning was critical to catalysis. The Baker group endeavored to design enzymes *in silico* that position the substrates precisely and catalyze Diels-Alder reactions.[42] A minimal active site geometry was generated using the Rosetta methodology that stabilizes the transition state, then an outer shell was computationally selected to shield this transition state from the solvent and position the critical residues properly. Of 86 predicted and experimentally tested designs, two were found to have Diels-Alderase activity. These two enzymes were rationally mutated, targeting residues in direct contact with the transition state to residues predicted to maintain or improve transition state binding. While exciting, these enzymes had only very modest catalytic performances; the best performing variant achieved a $k_{cat}$ of only 2.1 hr$^{-1}$. Further computational improvements were found with FoldIt, a crowdsourced computer game centered on protein folding.[43] This resulted in the addition of a helix-containing lid that buried the active site more deeply into a hydrophobic pocket, which did not affect $k_{cat}$ but improved $k_{cat}/(K_{diene} \times K_{dieneophile})$ 14-fold. Follow-up work employed directed evolution to

improve the efficiency of this enzyme further.[44] Through twelve rounds of mutagenesis and screening, first on the originally designed variant and later on the Foldit-produced variant, the enzyme was improved significantly to obtain a final variant with a $k_{cat}$ of 10.8 hr$^{-1}$ and nearly 90-fold improvement in $k_{cat}/(K_{diene}$ x $K_{dieneophile})$ compared with the initial rationally mutated variant. This resulted in one of the most efficient and selective artificial Diels-Alderases to date, allowing for preparative-scale synthesis of the *endo*-product with more than 90% conversion.

## 1.5 Conclusion

Biocatalysis offers a promising pathway for improving the sustainability, yield, and selectivity of chemical synthesis. Engineers have developed select few new-to-nature enzymatic reactions that allow for the construction of rings through C–C bond disconnections, though those that have been developed have shown their utility through their high yields, chemoselectivities, and enantioselectivities. In the following chapters, I will demonstrate the work I have done to further develop new-to-nature C–C bond forming cyclases, including the development of cyclopropanation reactions and the development of novel intramolecular C–H functionalization methods of cyclization. Future advances in directed evolution, and the efforts of future protein engineers, will expand the ability of enzymes to build rings and catalyze new-to-nature C–C bond forming reactions.

## 1.6 References for Chapter 1

1. Fanourakis, A.; Docherty, P. J.; Chuentragool, P.; Phipps, R.J. Recent Developments in Enantioselective Transition Metal Catalysis Featuring Attractive Noncovalent Interactions between Ligand and Substrate. *ACS Catal.* **2020**, 10, 18, 10672–10714.

2. Davies, H. M. L.; Manning, J. R. Catalytic C–H Functionalization by Metal Carbenoid and Nitrenoid Insertion. *Nature.* **2008,** 451, 417–424.

3. Labinger, J. A. Platinum-Catalyzed C–H Functionalization. *Chem. Rev.* **2017**, 117, 13, 8483–8496.

4. González, J., van Dijk, L., Goetzke, F.W. et al. Highly Enantioselective Rhodium-Catalyzed Cross-Coupling of Boronic Acids and Racemic Allyl Halides. *Nat Protoc* **2019**, 14, 2972–2985.

5. Price pressures on metals. *Nat Catal* **2019**, 2, 735.

6. Mencho, B. B. Assessing the Effects of Gold Mining on Environment: A Case Study of Shekiso District, Guji Zone, Ethiopia. *Heliyon.* **2022,** 8, 12.

7. Adelraheem, E. M. M.; Busch, H.; Hanefeld, U.; Tonin, F. Biocatalysis Explained: From Pharmaceutical to Bulk Chemical Production. *React. Chem. Eng.* **2019,** 4, 1878–1894.

8. On Advances and Challenges in Biocatalysis. *Nat. Catal.* **2018**, 1, 635–636.

9. Srivastava, A. & Srivastava, M. Enzymes Market—Global Opportunity Analysis and Industry Forecast, 2017–2024. **2018**; https://www.alliedmarketresearch.com/enzymes-market

10. Savile, C. K. *et al*. Biocatalytic Asymmetric Synthesis of Chiral Amines from Ketones Applied to Sitagliptin Manufacture. *Science.* **2010,** 329, 305–309.

11. Huffman, M. A. *et al*. Design of an in vitro Biocatalytic Cascade for the Manufacture of Islatravir. *Science.* **2019,** 366, 1255–1259.

12. Bloom, J. D.; Labthavikul, S. T.; Otey, C. R.; Arnold, F. H. Protein Stability Promotes Evolvability. *Proc. Nat. Acad. Sci.* **2006,** 103, 5869–5874.

13. Pluchinsky, A. J.; Wackelin, D. J.; Huang, X.; Arnold, F. H.; Mrksich, M. High Throughput Screening with SAMDI Mass Spectrometry for Directed Evolution. *J. Am. Chem. Soc.* **2020,** 142, 19804–19808.

14. Martin-Diaz, J.; Molina-Espeja, P.; Hofrichter, M.; Hollman, F.; Alcalde, M. Directed Evolution of Unspecific Peroxygenase in Organic Solvents. *Biotechnol. Bioeng.* **2021,** 118, 3002–3014.

15. Wang, M.; Si, T.; Zhao, H. Biocatalyst Development by Directed Evolution. *Bioresour. Technol.* **2012,** 115, 117–125.

16. Chen, K.; Arnold, F. H. Engineering New Catalytic Activities in Enzymes. *Nat. Catal.* **2020,** 3, 203–213.

17. Kan, S. B. J.; Lewis, R. L.; Chen, K.; Arnold, F. H. Directed Evolution of Cytochrome C for Carbon–Silicon Bond Formation: Bringing Silicon to Life. *Science.* **2016,** 354, 1048–1051.

18. Kan, S. B. J.; Huang, X; Gumulya, Y.; Chen, K.; Arnold, F. H. Genetically Programmed Chiral Organoborane Synthesis. *Nature.* **2017,** 552, 132–136.

19. Zhang, R. K.; Chen, K.; Huang, X.; Wohlschlager, L.; Renata, H.; Arnold, F. H. Enzymatic Assembly of Carbon–Carbon Bonds via Iron-Catalysed $sp^3$ C–H Functionalization. *Nature.* **2019,** 565, 67–72.

20. Chen, Y.; Rosenkranz, C.; Hirte, S.; Kirchmair, J. Ring Systems in Natural Products: Structural Diversity, Physicochemical Properties, and Coverage by Synthetic Compounds. *Nat. Prod. Rep.* **2022,** 39, 1544–1556.

21. Mei, T. S.; Ma, S.; Engle, K. M.; Yu, J. Q. Heterocycle Formation via Palladium-Catalyzed C–H Functionalization. *Synthesis.* **2012,** 44, 1778–1791.

22. Zheng, L.; Hua, R. Recent Advances in Construction of Polycyclic Natural Product Scaffolds via One-Pot Reactions Involving Alkyne Annulation. *Front. Chem.* **2020,** 8:580355.

23. Coelho, P. S.; Brustad, E. M.; Kannan, A.; Arnold, F. H. Olefin Cyclopropanation via Carbene Transfer Catalyzed by Engineered Cytochrome P450 Enzymes. *Science.* **2012,** 339, 307–310.

24. Wittman, B. J. *et al.* Diversity-Oriented Enzymatic Synthesis of Cyclopropane Building Blocks. *ACS Catal.* **2020,** 10, 13, 7112–7116.

25. Sriboe, M. G.; Vargas, D. A.; Fasan, R. Dehaloperoxidase Catalyzed Stereoselective Synthesis of Cyclopropanol Esters. *J. Org. Chem.* **2022,** 88, 12, 7630–7640.

26. Fasan, R.; Siriboe, M. G. Engineered Myoglobin Catalysts for Asymmetric Intermolecular Cyclopropanation Reactions. *Bull. Jap. Soc. Coord. Chem.* **2022,** 80, 4–13.

27. Ren, X. *et al.* Highly Stereoselective and Enantiodivergent Synthesis of Cyclopropylphosphonates with Engineered Carbene Transferases. *Chem. Sci.* **2022,** 13, 8550–8556.

28. Knight, *et al.* Diverse Engineered Heme Proteins Enable Stereodivergent Cyclopropanation of Unactivated Alkenes. *ACS Cent. Sci.* **2018,** 4, 3, 372–377.

29. Brandenburg, O. F. *et al*. Stereoselective Enzymatic Synthesis of Heteroatom-Substituted Cyclopropanes. *ACS Catal.* **2018,** 8, 4, 2629–2634.

30. Bordeaux, M.; Tyagi, V.; Fasan, R. Highly Diastereoselective and Enantioselective Olefin Cyclopropanation Using Engineered Myoglobin-Based Catalysts. *Angew. Chemie.* **2014,** 54, 1744–1748.

31. Coelho, P. S. *et al.* A Serine-Substituted P450 Catalyzes Highly Efficient Carbene Transfer to Olefins *in vivo*. *Nat. Chem. Bio.* **2013,** 9, 485–487.

32. Renata, H. *et al.* Identification of Mechanism-Based Inactivation in P450-Catalyzed Cyclopropanation Facilitates Engineering of Improved Enzymes. *Synthesis.* **2016,** 138, 38, 12527–12533.

33. Chandgude, A. L.; Ren, X.; Fasan, R. Stereodivergent Intramolecular Cyclopropanation Enabled by Engineered Carbene Transferases. *J. Am. Chem. Soc.* **2019,** 141, 23, 9145–9150.

34. Ren, X.; Liu, N.; Chandgude, A. L.; Fasan, R. An Enzymatic Platform for the Highly Enantioselective and Stereodivergent Construction of Cyclopropyl-δ-lactones. *Angew. Chemie.* **2020,** 59, 48, 21634–21639.

35. Ren, X.; Chandgude, A. L.; Fasan, R. Highly Stereoselective Synthesis of Fused Cyclopropane-γ-Lactams via Biocatalytic Iron-Catalyzed Intramolecular Cyclopropanation. *ACS Catal.* **2020,** 10, 3, 2308–2313.

36. Biegasiewicz, K. F. *et al*. Photoexcitation of Flavoenzymes Enables a Stereoselective Radical Cyclization. *Science.* **2019,** 364, 1166–1169.

37. Laguerre, N. *et al.* Radical Termination via β-Scission Enables Photoenzymatic Allylic Alkylation Using "Ene"-Reductases. *ACS Catal.* **2022,** 12, 15, 9801–9805.

38. Clayman, P. D.; Hyster, T. K. Photoenzymatic Generation of Unstabilized Alkyl Radicals: An Asymmetric Reductive Cyclization. *J. Am. Chem. Soc.* **2020,** 142, 32, 15673–15677.

39. Zhou, Q.; Chin, M.; Liu, P.; Yang, Y. Stereodivergent Atom-Transfer Radical Cyclization by Engineered Cytochromes P450. *Science.* **2021,** 374, 1612–1616.

40. Hilvert, D.; Hill, K. W.; Nared, K. D.; Auditor, T. M. Antibody Catalysis of the Diels-Alder Reaction. *J. Am. Chem. Soc.* **1989,** 111, 26, 9261–9262.

41. Gouverneur, V. E. *et al*. Control of the exo and endo Pathways of the Diels-Alder Reaction by Antibody Catalysis. *Science.* **1993,** 262, 204–208.

42. Siegel, J. B. *et al*. Computational Design of an Enzyme Catalyst for a Stereoselective Bimolecular Diels-Alder Reaction. *Science.* **2010,** 329, 309–313.

43. Eiben, C. B. *et al*. Increased Diels-Alderase Activity Through Backbone Remodeling Guided by Foldit Players. *Nat. Biotech.* **2012,** 30, 190–192.

44. Preiswerk, N. *et al*. Impact of Scaffold Rigidity on the Design and Evolution of an Artificial Diels-Alderase. *Proc. Nat. Acad. Sci.* **2014,** 111, 8013–8018.

*C h a p t e r   2*

# ENANTIO- AND DIASTEREOENRICHED ENZYMATIC SYNTHESIS OF 1,2,3-POLYSUBSTITUTED CYCLOPROPANES FROM (*Z/E*)-TRISUBSTITUTED ENOL ACETATES

Content in this chapter is adapted from published work:

**Abstract**

In nature and synthetic chemistry, stereoselective [2+1] cyclopropanation is the most prevalent strategy for the synthesis of chiral cyclopropanes, a class of key pharmacophores in pharmaceuticals and bioactive natural products. One of the most extensively studied reactions in the organic chemist's arsenal, stereoselective [2+1] cyclopropanation, largely relies on the use of stereodefined olefins, which can require elaborate laboratory synthesis or tedious separation to ensure high stereoselectivity. Here we report engineered hemoproteins derived from a bacterial cytochrome P450 that catalyze the synthesis of chiral 1,2,3-polysubstituted cyclopropanes, regardless of the stereopurity of the olefin substrates used. Cytochrome P450BM3 variant P411-INC-5185 exclusively converts (Z)-enol acetates to enantio- and diastereoenriched cyclopropanes and in the model reaction delivers a leftover (E)-enol acetate with 98% stereopurity, using whole *Escherichia coli* cells. P411-INC-5185 was further engineered with a single mutation to enable the biotransformation of (E)-enol acetates to α-branched ketones with high levels of enantioselectivity while simultaneously catalyzing the cyclopropanation of (Z)-enol acetates with excellent activities and selectivities. We conducted docking studies and molecular dynamics simulations to understand how active-site residues distinguish between the substrate isomers and enable the enzyme to perform these distinct transformations with such high selectivities. Computational studies suggest the observed enantio- and diastereoselectivities are achieved through a stepwise pathway. These biotransformations streamline the synthesis of chiral 1,2,3-polysubstituted cyclopropanes from readily available mixtures of (*Z/E*)-olefins, adding a new dimension to classical cyclopropanation methods.

## 2.1 Introduction

Chiral 1,2,3-polysubstituted cyclopropanes are prevalent in natural products and bioactive compounds[1-4] and are versatile building blocks for subsequent downstream manipulation due to their innate ring strain and dense substitution.[5-6] Stereoselective [2+1] cyclopropanation starting from an olefin and a C1 motif represents a popular disconnection in the retrosynthetic analysis of chiral cyclopropanes; this synthetic approach is widely used in nature as well as in synthetic chemistry (Figures 2-1A and 2-1B).[1, 7-8] In nature, *S*-adenosyl methionine (SAM)-dependent methyltransferases transfer exogenous C1 units from the methyl group of SAM to olefins via polar or radical chemistry (Figure 2-1A).[9] While highly selective, this approach is inherently limited to 1,2-substituted cyclopropanes, as neither nature nor engineers have succeeded in transferring non-methyl groups with these enzymes. Synthetic chemists, on the other hand, achieve stereoselective [2+1] cyclopropanations via a wider variety of carbon units, such as metal carbenes, metal carbenoids, or sulfur/nitrogen ylides (Figure 2-1B),[1, 7, 10-14] complementing the scope of nature-synthesized cyclopropanes and allowing for access to 1,2,3-polysubstituted cyclopropanes.

Although nature and chemists employ different techniques, stereospecificity is a shared feature of these transformations (Figures 2-1A and 2-1B).[1-2, 7, 9, 12, 14] When an olefin containing stereochemical information such as a (*Z*)- or (*E*)-configuration is used for [2+1] cyclopropanation, this information is retained in the cyclopropane product. Thus, if an olefin substrate is stereopure, the stereoselectivity of the cyclopropane product can be controlled with relative ease or even predicted *a priori*. On the other hand, if an olefin is a mixture of

(Z/E)-isomers, the yield and/or selectivity of a desired cyclopropane product will be diminished. Therefore, the geometric purity of the olefin plays a decisive role in the level of stereoinduction and thus determines the stereoselectivity and utility of the method. Because the synthesis of geometrically pure olefins can be challenging,[15-16] difficult and time-consuming purification of the olefin is often required to ensure high stereopurity of the products.

We were curious whether a catalyst could catalyze conversion of mixtures of (Z/E)-olefins into highly enantio- and diastereoenriched cyclopropane products. This would obviate elaborate synthesis or time-consuming separation to construct predefined olefins (Figure 2-1C). Methods that enable enantio- and diastereoenriched cyclopropanation of (Z/E)-olefin mixtures, however, are rare. There are two notable difficulties: (i) The catalyst must recognize the stereochemical information inherent in the olefin substrates. A highly selective catalyst must act exclusively on one substrate while excluding the other. (ii) The catalyst must simultaneously achieve high diastereo- and enantioselectivities, each of which is difficult to achieve.[17] A method that streamlines the synthesis of chiral 1,2,3-polysubstituted cyclopropanes from mixtures of (Z/E)-olefins would be of significant synthetic utility.
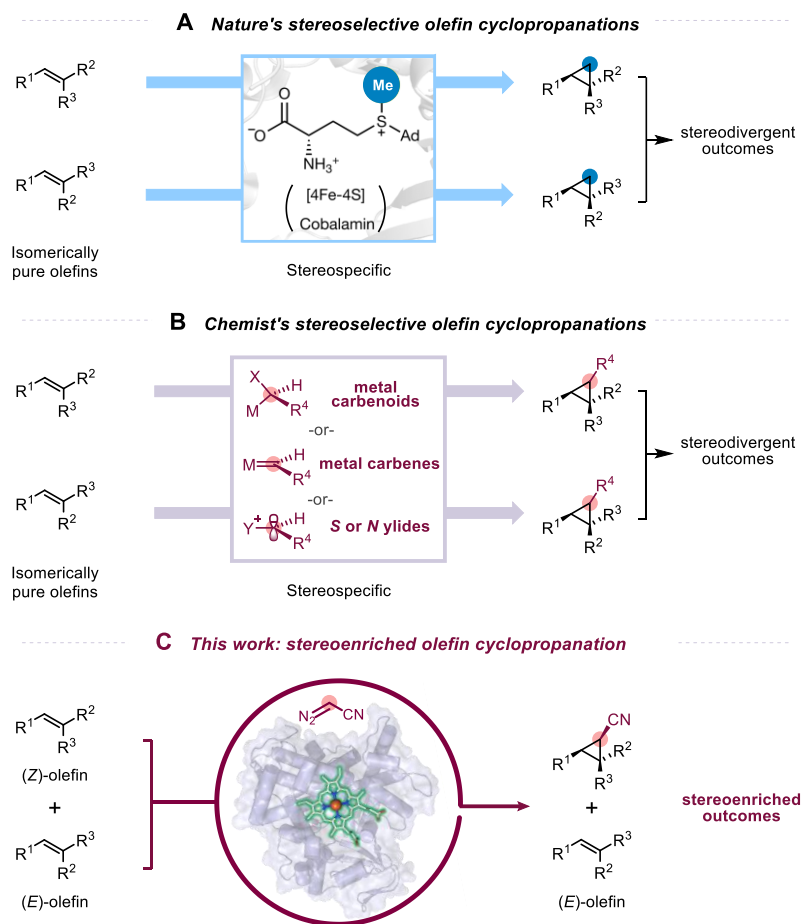
**Figure 2-1. Enzyme- and small molecule-catalyzed synthesis of cyclopropanes.** (A) Nature's stereoselective [2+1] cyclopropanation involving an exogenous C1 unit from SAM via a radical or a polar process. This approach is inherently limited to 1,2-substituted cyclopropanes. (B) Stereoselective olefin cyclopropanation methods invented by chemists: [2+1] cyclopropanation of olefins via high-energy intermediates such as metal carbenes, metal carbenoids, or sulfur/nitrogen ylides requires isomerically pure olefins for a stereopure product. (C) This work: enantio- and diastereoenriched olefin cyclopropanation does not require isomerically pure olefins and can form chiral 1,2,3-polysubstituted cyclopropanes and recover (*E*)-olefins. Structural illustrations are

adapted from Protein Data Bank (PDB) ID 5UL4 (radical SAM enzyme) and PDB 5UCW (cytochrome P450$_{BM3}$). Ad, adenosyl; R, organic groups; M, metal; X, leaving groups.

Enzymes are attractive candidates to meet these challenges, given their ability to exert exquisite control over both the substrates and the stereochemical outcome of chemical reactions.[18] There are reports highlighting new-to-nature transformations,[19] including olefin cyclopropanation, achieved by expanding the already enormous catalytic repertoire of the (iron)-heme-containing cytochrome P450 family.[20-21] Cytochromes P450 are excellent candidates for discovery of non-natural activities due to their structural flexibility and remarkable promiscuity.[22] A decade ago, we described repurposing a cytochrome P450 for non-natural cyclopropanation of styrenes with diazoesters, yielding cyclopropanes with high levels of selectivity via a putative carbene transfer process.[21] Since that first report, a plethora of hemoprotein-catalyzed carbene and nitrene transfer reactions have been developed.[20, 23-24] The activity and selectivity of these biocatalytic transformations are often complementary to the state-of-the-art processes based on small-molecule catalysts, making them a valuable addition to the synthetic chemist's toolbox. However, reports of biocatalytic cyclopropanation have largely been limited to terminal olefins and stereodefined internal olefins, sidestepping the issue of using geometrically different olefins (Figure 2-1C).[20, 23, 25-27] Based on these precedents, we saw opportunities to leverage cytochromes P450 for stereoenriched cyclopropanations from (Z/E)-mixed internal olefins.

## 2.2 Initial screening and directed evolution of alkyl transferase P411-INC-5185

We commenced the study by focusing on the cyclopropanation reaction between a 1:1 *Z/E* mixture of butyrophenone-derived enol acetate **1a** and diazoacetonitrile **2a**. These reagents were chosen for several reasons: (i) Enol acetates will be transformed into cyclopropyl acetates, which are synthetically and pharmaceutically relevant compounds.[28-29] (ii) The opposite polarities of the different groups on the alkene (polar acetyl and nonpolar phenyl) can help the enzyme distinguish different stereoisomers. (iii) Diazoacetonitrile **2a** was chosen for its strong electron-withdrawing character, small steric profile, and the valuable nature of the nitrile group. The nitrile substituent enhances the electrophilicity of the carbenoid and promotes reactions with the nucleophilic enol acetates. Additionally, the small steric profile of the nitrile is less likely to impede accommodation of diverse, sterically hindered substrates in the enzyme active site. Moreover, nitrile groups are valuable moieties that can be readily transformed into a variety of functional groups.

We screened a panel of 48 hemoproteins previously engineered for different carbene and nitrene transformations,[21, 30-32] including variants of cytochromes P411 (cytochromes P450 with an axial serine ligand), in intact *E. coli* cells. Compounds **1a** and **2a** were combined with the hemoproteins expressed in whole *E. coli* cells at room temperature under anaerobic conditions. The resulting reaction mixtures were analyzed after 20 hours for the formation of cyclopropane product **3a**. A truncated P411 variant lacking the FAD domain (named P411-INC-5182; see SI, Table S9, for more details) that was previously used for

intramolecular nitrene insertion into C($sp^3$)–H bonds[30] showed the best activity in this

initial screen, providing a total turnover number (TTN) of 68 and 9% yield (Figure 2-2A and

SI, Table 2-S1). The heme domain of this variant has seven mutations (A74G, V78L, L263Y,

T327I, T436L, L437Q, and S438T) with respect to the previously reported "**E10**" variant of

P450$_{\text{BM3}}$, which has a solved crystal structure (Figure 2-2B, PDB ID: 5UCW).[31] Control

experiments showed that formation of **3a** is neither catalyzed by the heme cofactor alone,

nor is it produced by the cellular background (SI, Table S1). When ethyl diazoacetate (EDA)

was used as a carbene precursor, no cyclopropane product was detected (SI, Table S1).
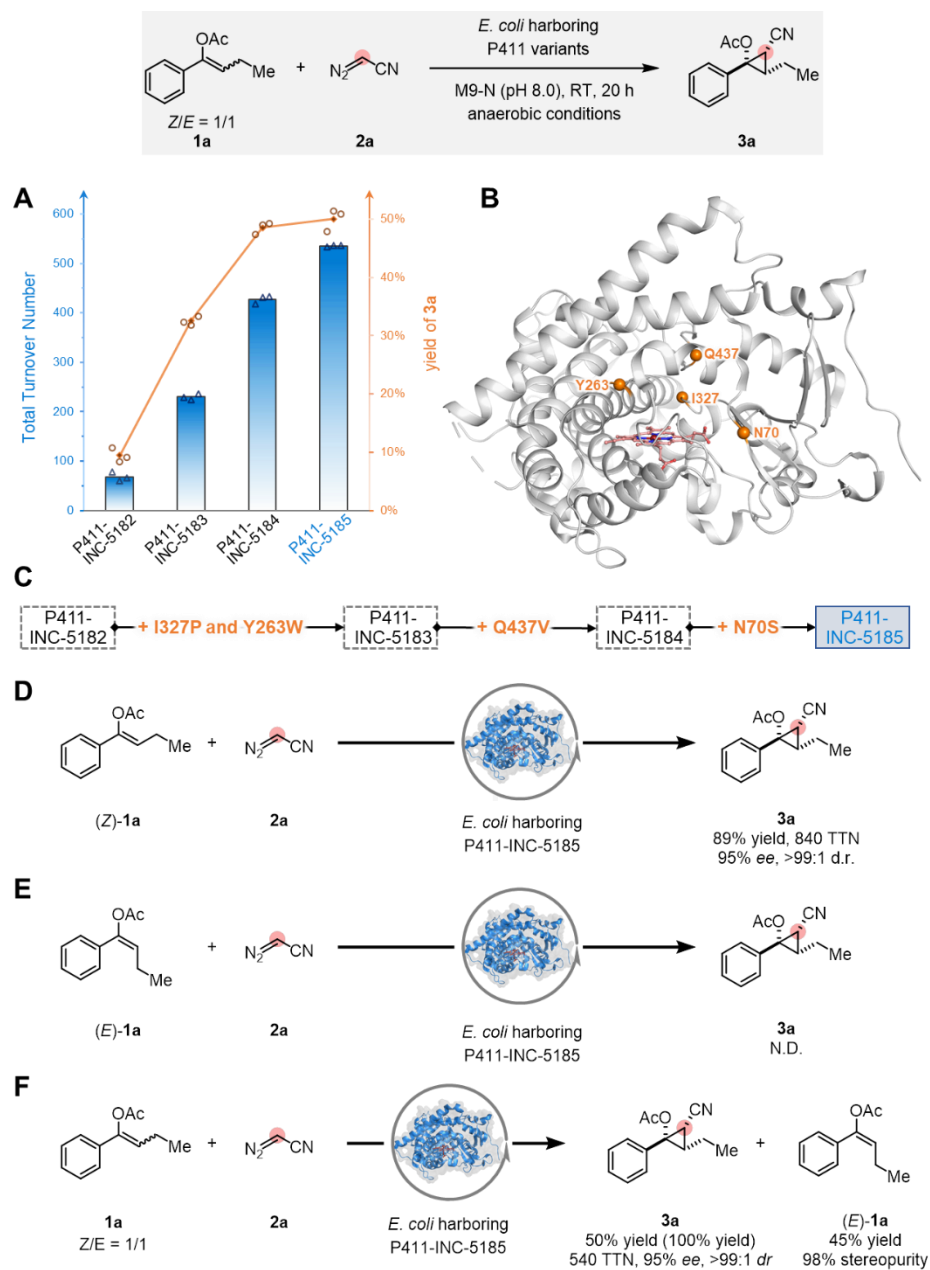
**Figure 2-2. Directed evolution for enantio- and diastereoenriched cyclopropanation.** Reaction conditions: 3 mM **1a**, 37 mM **2a**, *E. coli* whole cells harboring P411 variants ($OD_{600}$ = 30) in M9-N aqueous buffer (pH 8.0), 10% v/v EtOH (co-solvent), room temperature, anaerobic conditions, 20 h. (A) Directed evolution of enantio- and diastereoenriched alkyltransferase P411-INC-5185. Reactions were performed in triplicate (n = 3). Yields were calculated from high-performance liquid

chromatography (HPLC) calibration curves. Yields (blue bars) and TTNs (yellow line) reported are the means of three independent experiments. (B) The mutated residues (N70, Y263, I327, and Q437) conferring activity increases are highlighted in the enzyme active site structure (of closely related P411 variant **E10** (PDB ID: 5UCW)). (C) Evolutionary trajectory for enantio- and diastereoenriched cyclopropanation. (D) P411-INC-5185-catalyzed cyclopropanation reaction using (*Z*)-**1a**. (E) P411-INC-5185-catalyzed cyclopropanation reaction using (*E*)-**1a**. (F) P411-INC-5185-catalyzed cyclopropanation reaction using a mixture of (*Z/E*)-**1a**. Enantiomeric excesses of **3** were determined by converting cyclopropanated products **3** into *α*-branched ketones in the presence of 15 mol% LiOH at 4 °C (see SI, general procedure F, for more details). Ac, acetyl group; TTN, total turnover number.

Since P411-INC-5182 showed promising activity toward the desired cyclopropanation reaction, we decided to revisit other enzymes in the collection that are closely related to P411-INC-5182.[30, 32] Variant P411-INC-5183, with mutations I327P and Y263W relative to P411-INC-5182, showed higher activity than P411-INC-5182 (P411-INC-5183 can catalyze the formation of **3a** in 33% yield and 230 TTN, Figure 2-2A). We chose variant P411-INC-5183 as the parent for directed evolution via iterative rounds of site-saturation mutagenesis (SSM)[33-34] and screening, targeting amino acid residues close to the heme cofactor (Figure 2-2B). Mutation Q437V, which resides on the flexible loop above the heme (Figure 2-2B), increased the TTN to 430 (49% yield of **3a**, Figure 2-2B). Exploring more amino acid residues in the enzyme pocket using SSM identified the N70S mutation and increased the TTN to 540 (50% yield, 95% *ee*, and >99:1 *dr* of **3a**). We found it difficult to push the yield higher at this stage. Given the 1:1 ratio of (*Z/E*)-**1a**, we speculated this 50% yield ceiling meant the enzyme was reacting with only one isomer of the starting material.

To verify this and gain more insight into the biotransformation, stereoisomers (*Z*)-**1a** and

(*E*)-**1a** were used individually as the substrate for P411-INC-5185-catalyzed

cyclopropanation reactions with diazoacetonitrile **2a** under standard conditions (Figures 2-

2D and 2-2E). P411-INC-5185 converts pure (*Z*)-**1a** into **3a** with 89% yield and 840 TTN

(95% *ee* and >99:1 *dr*) but does not react with pure (*E*)-**1a** to form a cyclopropane product

(Figures 2-2D and 2-2E). Encouraged by this, we re-examined the reactions with a 1:1 *Z/E*

mixture of **1a** and found the remaining starting material was highly enriched in (*E*)-isomer

(*Z/E* = 2:98; Figure 2-2F). This is interesting, as pure (*E*)-olefins are more difficult to obtain

due to their lower thermodynamic stabilities than (*Z*)-olefins, and only a handful of methods

are reliable for the diastereoselective synthesis of these (*E*)-olefins.[35] In this regard, in

addition to the generation of high-value added chiral 1,2,3-polysubstituted cyclopropanes,

P411-INC-5185 also holds potential utility in delivering stereopure (*E*)-enol acetates.

## 2.3 Substrate scope study

We surveyed the activity of P411-INC-5185 on a series of trisubstituted enolate substrates

under the standard whole-cell reaction conditions (Figure 2-3). P411-INC-5185 catalyzes

cyclopropanation of *Z/E* mixtures of diverse *α*-aryl, *β*-alkyl-substituted enolates **1** with

diazoacetonitrile **2a**, delivering the desired products **3** in synthetically useful yields and

excellent diastereo- and enantioselectivities (up to >99:1 *dr* and >99% *ee*, Figure 2-3),

regardless of the stereopurity of the olefin substrates used. Substrates with diverse

substituents on the *α*-aryl group are compatible with this reaction. Electron-donating, -

neutral, and -withdrawing substituents on the aromatic rings were all compatible (**3a**–**3n**,

Figure 2-3), affording 1,2,3-polysubstituted cyclopropanes with uniformly high levels of diastereo- and enantioselectivities. *Para*- and *meta*-substituted α-aryl enolates both reacted well with diazoacetonitrile **2a** to give the corresponding cyclopropanes (**3b** and **3c**, Figure 2-3). Substrates bearing a halogen functional group, such as fluoro- (**1d** and **1e**), and bromo-
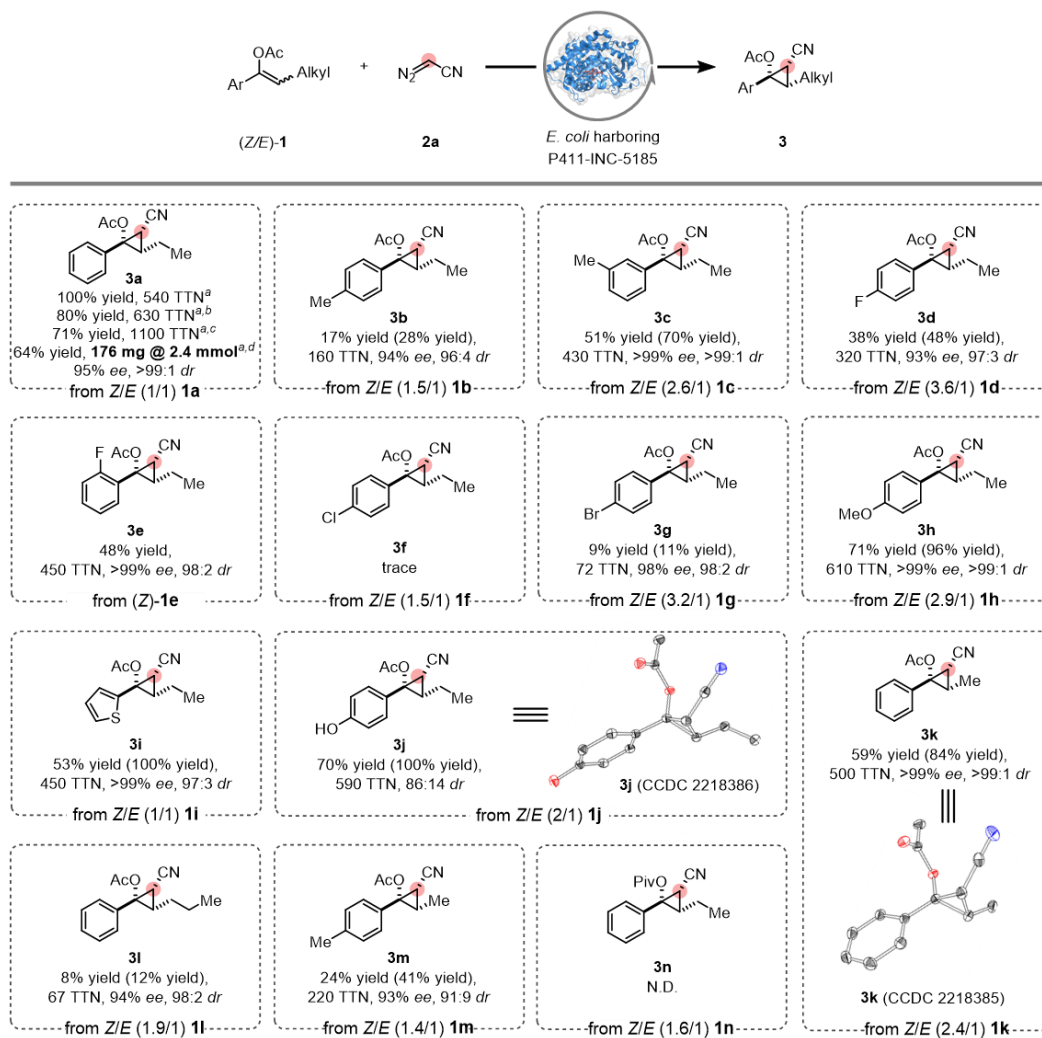


**Figure 2-3. Substrate scope study.** Reaction conditions: 3 mM **1**, 37 mM **2a**, *E. coli* whole cells harboring P411-INC-5185 (OD$_{600}$ = 30) in M9-N aqueous buffer (pH 8.0), 10% v/v EtOH (co-solvent), room temperature, anaerobic conditions, 20 h. Yields are based on mixed enol acetates.

Yields in parentheses are based on (*Z*)-isomer. Yields were calculated from HPLC calibration curves and the average of triplicate experiments. [a]Yields are based on (*Z*)-**1a**. [b]5 mM 1:1 *Z/E*-**1a** and 35 mM **2a** were used. [c]10 mM 1:1 *Z/E*-**1a** and 30 mM **2a** were used. [d]Preparative-scale reaction conditions: 2.4 mmol 10 mM 1:1 *Z/E*-**1a**, 7.2 mmol 30 mM **2a**, *E. coli* whole cells harboring P411-INC-5185 ($OD_{600}$ = 30) in M9-N aqueous buffer (pH 8.0), 10% v/v EtOH (co-solvent), room temperature, anaerobic conditions, 20 h; isolated yield. Enantiomeric excesses of **3** were determined by converting cyclopropanated products **3** into $\alpha$-branched ketones in the presence of 15 mol% LiOH at 4 °C (see SI, general procedure F, for more details). Ar, aryl groups; Alkyl, alkyl groups; Ac, acetyl group; Piv, pivaloyl group; N.D. = no product was detected; TTN, total turnover number.

(**1g**), were tolerated to generate cyclopropanated products (**3d**, **3e** and **3g**, Figure 2-3) with excellent selectivities. A chloro group (**1f**, Figure 2-3), however, is detrimental to the cyclopropanation reaction, giving a trace amount cyclopropane product **3f** (Figure 2-3). Introducing an electron-donating group can increase the yield of cyclopropanes (**3h** and **3j,** Figure 2-3). Notably, a phenolic hydroxyl group, which is generally incompatible with small-molecule carbene transfer reactions due to its nucleophilicity, [35] is well tolerated by the enzymatic system (**3j,** Figure 2-3), which highlights the catalyst's functional group tolerance. Structural perturbations, such as the substitution of the aryl ring by thiophene (**3i,** Figure 2-3)**,** are also well accepted. Furthermore, $\beta$-substituted enol acetates bearing various alkyl chains (**1k**–**1m**, Figure 2-3) could be transformed to the corresponding cyclopropyl acetates with high levels of diastereo- and enantioselectivities (up to >99:1 *dr* and >99% *ee*). We also attempted to replace the enol acetyl group with a pivaloyl group, but no cyclopropanated product was observed (**3n**, Figure 2-3), likely because the bulky pivaloyl group prevented recognition of **1n** by the enzyme. The absolute stereochemistry for enzymatic products **3j**

and **3k** was assigned as *S, S, S* through X-ray crystallography (see SI, **Section 12**). The other polysubstituted cyclopropyl acetates **3** were assigned by analogy.

To showcase the utility of P411-INC-5185, we challenged the biocatalyst with higher substrate loadings. Under standard conditions and an alkene concentration of 3 mM 1:1 *Z/E*-**1a**, the TTN of the template reaction was 540 (**3a**, Figure 2-3). Encouragingly, increased substrate loadings and decreased **2a** to **1a** ratios also resulted in synthetically useful yields, while increasing the TTNs to 630 (5 mM; **2a**/**1a** = 7/1) and 1100 (10 mM; **2a**/**1a** = 3/1) (**3a**, Figure 2-3 and SI, Table S4). P411-INC-5185 was also shown to be robust when catalyzing the biotransformation of mixed **1a** in a preparative scale, delivering 176 mg **3a** in 64% isolated yield (Figure 2-3 and see SI, **Section 3.9** for details).

## 2.4 Engineering diastereomer-differentiating carbene transferase P411-INC-5186

Re-examining site-saturation mutagenesis libraries based on P411-INC-5185, we found that mutations at residue 263 resulted in a significant reduction in remaining starting material and the appearance of a new product peak in some variants. Mutation of the tryptophan to smaller, more flexible methionine yielded the largest new product peak, which was identified as $\alpha$-branched ketone product **4a**. The specificity of this new variant, P411-INC-5186, was tested with pure (*Z*)-**1a** and (*E*)-**1a** separately (Figures 2-4A and 2-4B).

Interestingly, P411-INC-5186 not only converts (*Z*)-**1a** and (*E*)-**1a**, it converts them into different products. To be specific, (*Z*)-**1a** was converted to the cyclopropane product **3a** in

91% yield with high selectivities (98% *ee* and >99:1 *dr***,** Figure 2-4A**)**, but no **4a** was detected (>99% chemoselectivity of **3a** over **4a**; Figure 2-4A). Conversely, P411-INC-5186 catalyzes the transformation of (*E*)-**1a** into **4a** in 30% yield and with 97% *ee*, but **3a** was not detected (>99% chemoselectivity of **4a** over **3a**; Figure 2-4B). We also investigated the substrate scope of P411-INC-5186-catalyzed diastereomer-differentiating transformations (SI, Figure 2-S1). P411-INC-5186 accepts a variety of *Z/E* mixtures of *α*-aryl, *β*-alkyl-substituted enolates **1** with diazoacetonitrile **2a**, delivering the desired cyclopropane products **3** (14 examples, up to quantitative yield (yields calculated based on the (*Z*)-isomer), >99% chemoselectivity, >99:1 *dr* and >99% *ee*; SI, Figure 2-S1) and *α*-branched ketone products **4** (Figure 2-4B and SI, Figure 2-S1) in synthetically useful yields and excellent diastereo- and enantioselectivities.
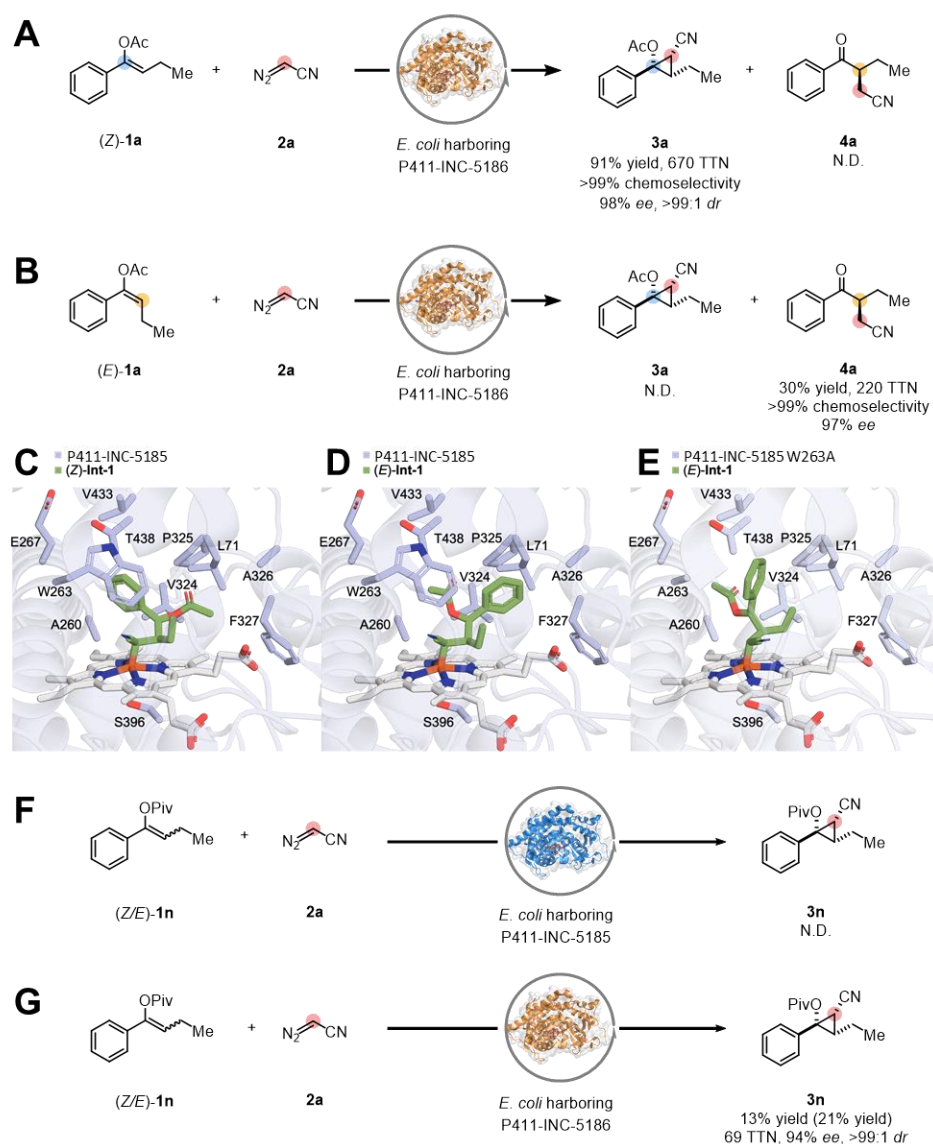
**Figure 2-4. Discovery of the diastereomer-differentiating alkyl transferase P411-INC-5186 and its activity toward divergent alkyl transfer reactions.** (A) P411-INC-5186-catalyzed alkyl transfer reaction using (*Z*)-**1a**. (B) P411-INC-5186-catalyzed alkyl transfer reaction using (*E*)-**1a**. Reaction conditions: 3 mM (*Z*)-**1a** or (*E*)-**1a**, 37 mM **2a**, *E. coli* whole cells harboring P411-INC-5186 (OD$_{600}$ = 30) in M9-N aqueous buffer (pH 8.0), 10% v/v EtOH (co-solvent), room temperature, anaerobic conditions, 20 h. Docking simulations of: (C) (*Z*)-**Int-1** with P411-INC-5185, (D) (*E*)-**Int-1** with

P411-INC-5185, and (E) (*E*)-**Int-1** with P411-INC-5185 A263. (F) P411-INC-5185-catalyzed

alkyl transfer reaction using a mixture of *Z/E* (1.6/1)-**1n**. (G) P411-INC-5186-catalyzed alkyl transfer

reaction using a mixture of *Z/E* (1.6/1)-**1n**. Reaction conditions: 3 mM (*Z/E*)-**1n**, 37 mM **2a**, *E. coli*

whole cells harboring P411-INC-5185 or P411-INC-5186($OD_{600}$ = 30). Yields were calculated from

HPLC calibration curves and the average of triplicate experiments. Enantiomeric excesses of **3** were

determined by converting cyclopropanated products **3** into *α*-branched ketones in the presence of 15

mol% LiOH at 4 °C (see SI, general procedure F, for more details). Ac, acetyl group; Piv, pivaloyl

group; N.D. = no product was detected; TTN, total turnover number.

As P411-INC-5185 and P411-INC-5186 differ only by a single amino acid, at residue 263,

we were curious to understand this site's role in substrate recognition. We first conducted

docking simulations to understand how P411-INC-5185 catalyzes the cyclopropanation of

(*Z*)-**1a** to form **3a** and not **4a**. To do this, we turned to quantum mechanics to calculate the

reaction pathway and docked in the heme-coordinated intermediate ((*Z*)-**Int-1** in Figure 2-

4C) that forms after nucleophilic attack by the *β*-carbon of (*Z*)-**1a** with the electrophilic

heme-carbene. Docking simulations predict that (*Z*)-**Int-1** fits tightly in the active site of

P411-INC-5185 (Figure 2-4C). The phenyl and acetyl handles of (*Z*)-**Int-1** form stabilizing

hydrophobic interactions with the side chains of residues V324 and W263. The conformation

of docked (*Z*)-**Int-1** is distorted toward a geometry that allows for direct ring closure to the

observed cyclopropane **3a** product. In the structure (Figure 2-4C), there is a polarity

mismatch: the phenyl group is situated adjacent to the hydrophilic side chains of T438 and

E267 and the acetyl group is forced into a hydrophobic pocket surrounded by V324 and

P325. By docking **Int-1** with free rotation around all single bonds, we discovered a second

pose (*E*)-**Int-1** that must form from (*E*)-**1a** based on the orientation of the phenyl and acetyl

groups (Figure 2-4D). In the (*E*)-**Int-1** pose, the phenyl group is wedged between W263 and P325 forming stabilizing hydrophobic interactions, and the acetyl group is pointed toward the hydrophilic active site region near T438 – here, the substituents have rotated to a conformation that allows the polarity of the substituents to match the active site residue polarity. However, the (*E*)-**Int-1** docking pose overall fits very tightly in the active site, which may result in a number of destabilizing steric interactions. Since experimental evidence shows that no products derived from (*E*)-**1a** are observed when P411-INC-5185 is the catalyst, it can be reasonably assumed that this (*E*)-**Int-1** pose is an artifact of the docking simulations that forces other probable poses when there are none. To gain insights into the altered reactivity with P411-INC-5186 as the catalyst, we performed an *in silico* mutation of W263 to alanine. This W263A mutation (Figure 2-4E) significantly enlarges the active site and allows rotation of the phenyl group into a conformation < 1.7 Å from where W263 would have been. Furthermore, rotation of the phenyl group effectively removes stabilization of the benzylic position and pushes this partial charge onto the oxygen, further facilitating the loss of the acetyl group and formation of the α-alkyl transfer product **4a** (Figures 2-4B and 2-4D). Perhaps from this (*E*)-**Int-1** conformation - when the acetyl group is near hydrophilic T438 - hydrolysis and loss of an acetic acid precludes cyclopropane formation and leads to α-alkyl transfer product **4a** instead (Figures 2-4B and 2-4D). Similarly, introducing the W263M mutation (P411-INC-5186) shows a docking position similar to the alanine mutant, implying the exclusion of (*E*)-**1a** in P411-INC-5185 is due to the smaller active site cavity compared to the other variants (see SI, **Section 10.4**, for more details). This is corroborated by the fact that using a bulkier pivaloyl protecting group on the oxygen shuts down cyclopropanation with P411-INC-5185 (**3n**, Figures 2-3 and 2-4F), whereas P411-INC-5186 was able to

catalyze the transformation of (*Z/E*)-**1n** and **1f** to **3n** and **3f** in 13% (69 TTN) and 23% (170 TTN) yields, respectively, with excellent selectivities (Figure 2-4G, and SI, Figure 2-S1). We attribute these differences to the enlarged active site of P411-INC-5186.

## 2.5 Computational modeling supports a stepwise pathway

We were curious how this enzyme can overcome potential severe steric constraints in order to cyclopropanate highly substituted olefins. Classical cyclopropanation reactions following a concerted reaction pathway encounter steric clashes which often make reactions with densely functionalized olefins challenging.[26, 36] The potential energy surface for cyclopropanation was calculated by means of density functional theory (DFT) calculations using a standard heme model at the B3LYP-D3(BJ)-CPCM(Et$_2$O)/def2-TZVP//B3LYP-D3(BJ)/def2-SVP level of theory (Figure 2-5A).[37-42] These model calculations indicate that the key cyclopropanation step takes place via a stepwise mechanism involving two distinct C–C bond formations on an open-shell singlet (OSS) and a triplet surface instead of a concerted process (Figure 2-5; see SI, **Section 10.1**, for more details).[43-45] The computational study suggests a radical-based stepwise mechanism reminiscent of metallocarbene radical reactivity in cobalt and iron-catalyzed olefin cyclopropanation reactions.[46-48] Upon rapid formation of the iron-carbene **Int-2** via transition state **TS-1** and subsequent coordination of **1a**, the nucleophilic $\beta$-carbon of the enol acetate attacks the electrophilic carbene via (*E*)-**TS-2** and (*Z*)-**TS-2** to form two conformations of the intermediate, (*E*)-**Int-4** and (*Z*)-**Int-4**, respectively (Figures 2-5A and 2-5C). This first bond-forming event with both (*E*)-**1a** and (*Z*)-**1a** results in the formation of the (*S*)-configured stereocenter (Figures 2-5A and 2-5C).

Transition states (*E*)-**TS-2** and (*Z*)-**TS-2** are nearly isoenergetic ($\Delta\Delta G^{\ddagger} = 0.2$ kcal·mol$^{-1}$)

which indicates an almost equal rate of formation of intermediates (*E*)-**Int-4** and (*Z*)-**Int-4**

(Figure 2-5A). Calculations revealed an enantiomeric transition state, (*Z*)-**TS-2a**, which is

preferable to (*Z*)-**TS-2** by 1.1 kcal·mol$^{-1}$. This result indicates that the enzyme active site is

configured in such a way to disfavor binding and formation of (*Z*)-**TS-2a**. We hypothesize

that this is due to the conformation of the iron-carbene intermediate in the enzyme active site.

Molecular dynamics (MD) simulations on this iron-carbene intermediate indicate that the N–

Fe–C–C dihedral has an average angle of 0° and the highest probabilities were found at +50°

and –60° (Figure 2-5B), thus showing a preferential orientation of the cyano substituent

toward the less crowded, outward-facing area of the active site. These results suggest that

formation of (*Z*)-**TS-2a** in the enzyme requires (*Z*)-**1a** to approach in such a way that both

ethyl and acetoxy substituents are oriented towards the sterically more congested back of the

active site, thereby likely resulting in a destabilization of (*Z*)-**TS-2a** compared to (*Z*)-**TS-2**.

Afterwards, ring-closing C–C bond formation from intermediates (*E*)-**Int-4** and (*Z*)-**Int-4**

takes place via transition states (*Z*)-**TS-3** and (*E*)-**TS-3**, respectively. (*Z*)-**TS-3** leads to the

observed cyclopropane product **3a**, whereas the activation free energy for (*E*)-**TS-3** is 2.2

kcal·mol$^{-1}$ higher, corresponding to a 100-fold slower reaction at room temperature. This

energetic value is corroborated by the experimental result that only the cyclopropane **3a** is

observed. Presumably, due to this larger energetic barrier, (*E*)-**Int-4** undergoes rapid

hydrolysis to form the *α*-alkyl transfer product **4a**. Based on this computational evidence,

this process seems more likely than formation of the cyclopropane via (*E*)-**TS-3** and

concomitant ring opening to form **4a**.

Both experimental and computational results indicate that the chemoselectivity of P411-INC-5186 towards isomeric olefin substrates occurs in a step following the formation of the first chiral center (Figure 2-5C). These results support a stepwise mechanism for P411-INC-5186-catalyzed alkyl transfer reactions (Figures 2-4 and 2-5). However, further studies are required to validate this proposal. A catalytic cycle based on these computational results is proposed in Figure 2-5C.
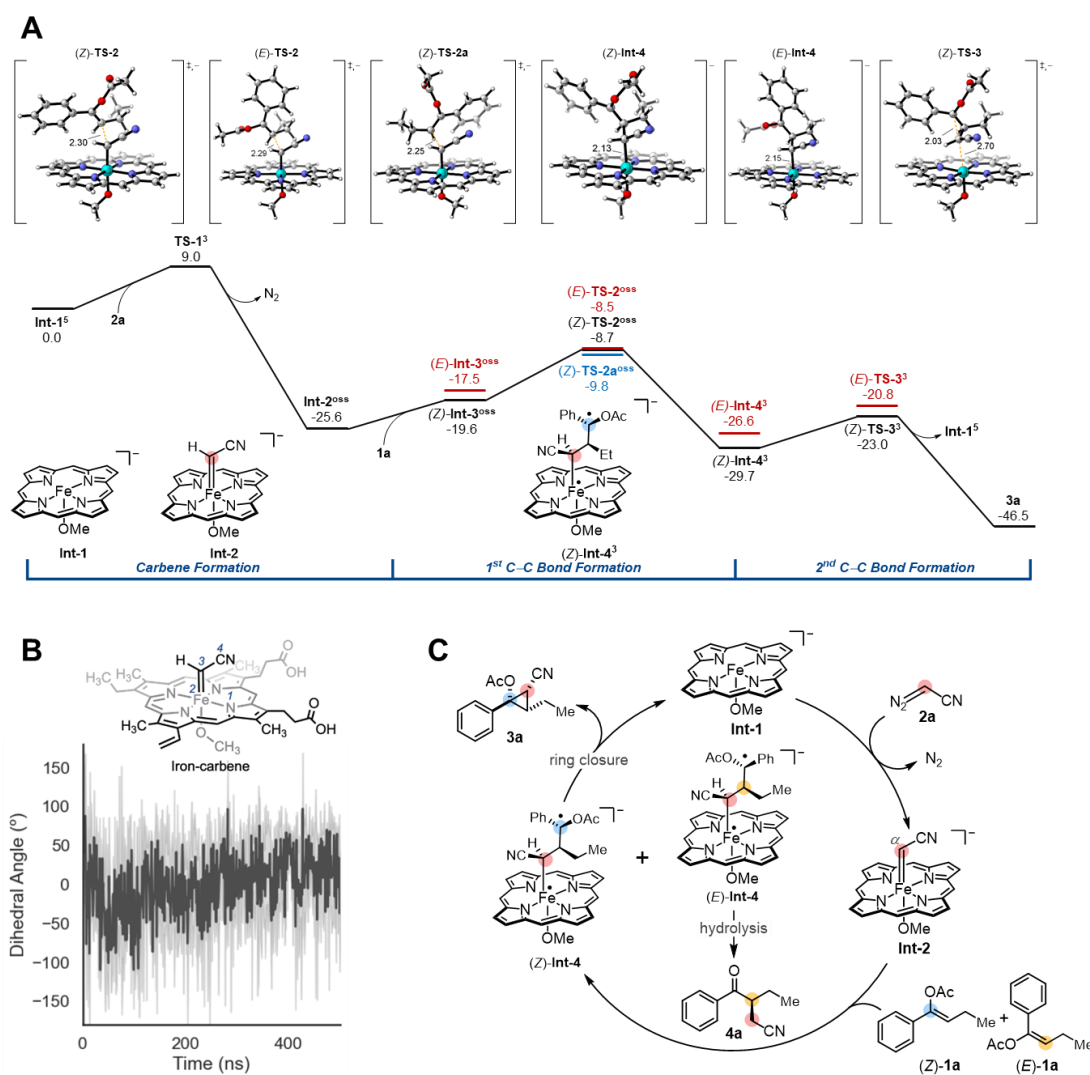
**Figure 2-5. Proposed stepwise pathway for enantio- and diastereoenriched enzymatic cyclopropanation.** (A) Calculated energy diagram for cyclopropanation (standard heme model at the B3LYP-D3(BJ)-CPCM(Et$_2$O)/def2-TZVP//B3LYP-D3(BJ)/def2-SVP level of theory; energies are Gibbs free energies in kcal mol$^{-1}$; superscripts correspond to the spin state; distances are given in Å). (B) Molecular dynamics simulations on the iron-carbene intermediate. (C) A proposed catalytic cycle of the biocatalytic cyclopropanation. TS, transition state; Int, intermediate; OSS, open-shell singlet; Ac, acetyl group.

## 2.6 Summary and conclusion

In summary, we have developed a biocatalytic platform for highly enantio- and diastereoselective cyclopropanation of mixed (*Z/E*)-trisubstituted enol acetates. These biocatalysts are fully genetically encoded, allowing for rapid tuning and reconfiguration via manipulation of the DNA sequence. Through directed evolution, we discovered two P411 variants, P411-INC-5185 and P411-INC-5186, where P411-INC-5185 exclusively catalyzes cyclopropanation and P411-INC-5186 enables diastereomer-differentiating transformations, both with excellent selectivities. Both DFT calculations and experimental results suggest a stepwise mechanism for these biotransformations, and docking simulations highlight the critical role of site 263 in controlling the active-site accommodation of the (*Z/E*)-olefinic isomers. This approach differs from traditional carbene-transfer cyclopropanations by converting hard-to-isolate olefinic mixtures into a single chiral cyclopropane product with exceptional selectivities. We anticipate that this biocatalytic platform will expedite the synthesis of chiral 1,2,3-polysubstituted cyclopropanes from readily available olefinic isomers.

## 2.7 Supplementary information for Chapter 2

**Tables**

**Table 2-S1**. Initial activity screening with engineered P411s. [a]
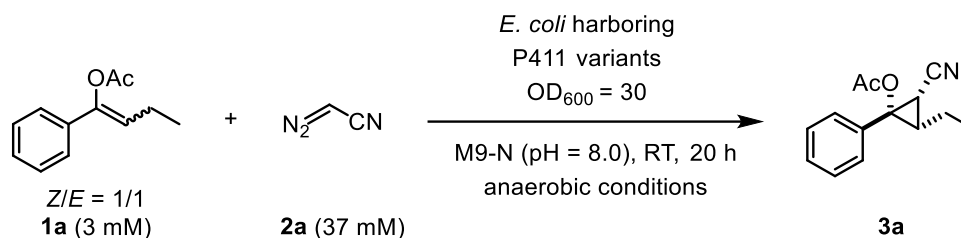


Forty-eight P450 and 411 variants from the Arnold lab culture collection were examined. A cytochrome P411 variant, P411$_{Diane1}$ (a truncated P411 variant lacking the FAD domain), previously used for intramolecular nitrene insertion into $C(sp^3)$–H bonds,[1] showed the best activity toward formation of cyclopropane product **3a** (9% yield).

| Entry | Variant / catalyst (+ reductant) | Yield of **3a** |
|---|---|---|
| 1 | P411-INC-5182 (P411$_{Diane1}$)[1] | 9% (19%) |
| 2 | hemin (20 μM) [b] | N.D. |
| 3 | hemin (20 μM) + Na$_2$S$_2$O$_4$ [c] | N.D. |
| 4 | hemin (20 μM) + Na$_2$S$_2$O$_4$ + BSA (20 μM) [c] | N.D. |
| 5 | cellular background [d] | N.D. |
| 6 | replace **2a** with ethyl diazoacetate (EDA) | N.D. [e] |

[a] Experiments were performed using whole *E. coli* cells according to the protocol described in **Sections 3.4** and **3.5**. The yields of **3a** were calculated based on comparing the signal

integration ratio of the products and an added internal standard compound (1,2,3-trimethoxylbenzene). The signal calibration curves between the products and the internal standard compound are described in **Section 8**. To enhance systematic management of different enzyme variants within the Arnold lab, we recently implemented a new nomenclature system. Variants are named as follows: family name-chemistry abbreviation-entry code. All enzyme variants in this study follow this nomenclature.

N.D. – no product was detected.

[b] Negative control experiments (with free hemin) were performed without the addition of $Na_2S_2O_4$. Under this condition, the resting oxidation state of hemin in aqueous buffer should be Fe(III).

[c] Negative control experiments using free hemin under reduced conditions (Fe(II)) were performed using an excess amount of $Na_2S_2O_4$ (20 mM).

[d] Cellular background control experiments were performed in whole-cell format, using *E. coli* (*E. cloni* BL21(DE3)) cells harboring an engineered tryptophan synthase *β*-subunit (Tm9D8*). The gene of this enzyme was also cloned into the pET22b(+) vector (Novagen) between restriction sites *Nde*I and *Xho*I. The protein expression protocol for this experiment followed the standard P450 and P411 expression conditions as described in **Sections 3.6**, **3.7**, and **3.8**.

[e] When using ethyl diazoacetate (EDA) as a carbene precursor, no corresponding cyclopropane product was detected.

**Table 2-S2**. Directed evolution of P411-INC-5185 for enantio- and diastereoenriched cyclopropanation. [a]
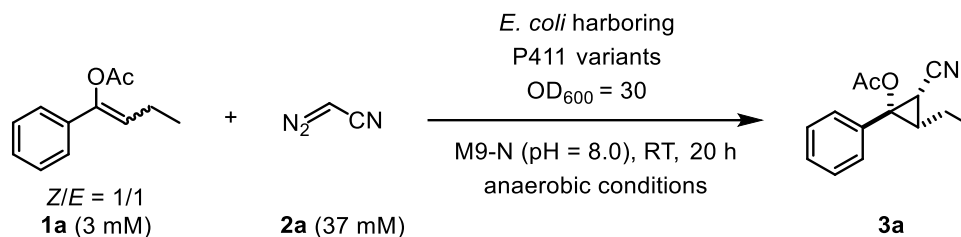


| Round # | Variant | Yield of 3a[a] | Standard deviation of yield | TTN of 3a | Standard deviation of TTN |
|---|---|---|---|---|---|
| 1 | P411-INC-5182[b] | 9% (19%) | 1.2% (2.5%) | 68 | 9 |
| 2 | P411-INC-5183 (P411-INC-5182 I327P, Y263W) | 33% (65%) | 0.8% (1.6%) | 230 | 6 |
| 3 | P411-INC-5184 (P411-INC-5183 Q437V) | 49% (97%) | 1.0% (2.0%) | 427 | 9 |
| 4 | P411-INC-5185 (P411-INC-5184 N70S)[b] | 50% (100%) | 0.2% (0.4%) | 536 | 2 |

[a] Experiments were performed using a suspension of *E. coli* cells harboring enzyme variants prepared according to the protocol described in **Sections 3.6**, **3.7**, and **3.8**. Reactions were performed in triplicate. Yields reported are the average of three experiments. Yields were calculated based on all **1a** added, and those in parentheses were calculated based on (*Z*)-**1a**.

[b] Final variant for enantio- and diastereoenriched cyclopropanation. (*E*)-**1a** was recovered in 45% yield and 98% stereopurity, see **Section 8** for details.

**Table 2-S3**. Optimization of pH values (M9-N minimal media) of P411-INC-5185-catalyzed cyclopropanation reactions. [a]
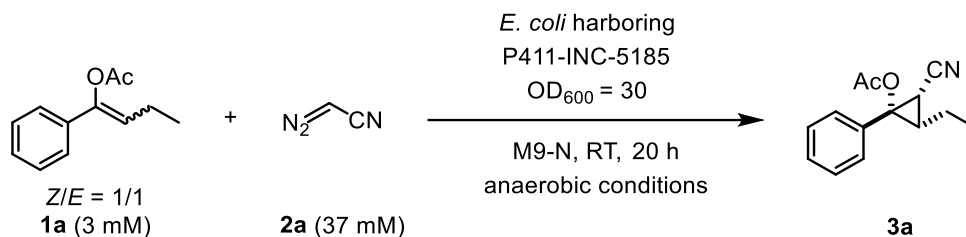


| Entry | Buffer condition | Yield of 3a [b] | Standard deviation of yield |
|-------|------------------|-----------------|------------------------------|
| 1 | M9-N (pH = 6.0) | 84% | 2.6% |
| 2 | M9-N (pH = 7.0) | 79% | 1.2% |
| 3 | M9-N (pH = 8.0) | 100% | 0.4% |
| 4 | M9-N (pH = 9.0) | 84% | 6.5% |

[a] Experiments were performed using a suspension of *E. coli* cells harboring P411-INC-5185 prepared according to the protocol described in **Sections 3.6**, **3.7**, and **3.8**. Reactions were performed in triplicate. Yields reported are the average of three experiments.

[b] Yields reported are the average of three experiments. Yields were calculated based on (*Z*)-**1a**.

**Table 2-S4**. Optimization of **1a**/**2a** ratios of P411-INC-5185-catalyzed cyclopropanation reactions. [a]
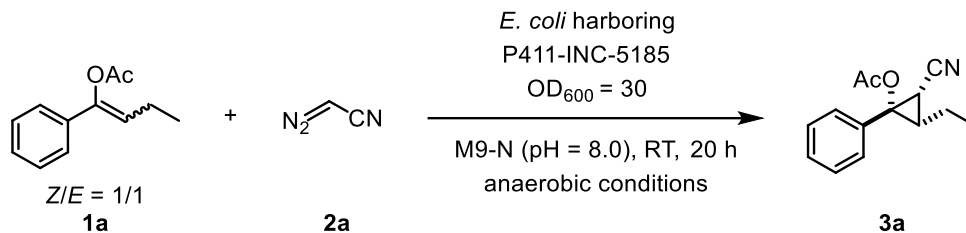


| Entry | Concentration of 1a | Concentration of 2a (equiv.) | Yield of 3a[b] | Standard deviation of yield | TTN of 3a | Standard deviation of TTN |
|-------|---------------------|------------------------------|----------------|------------------------------|-----------|----------------------------|
| 1 | 3 mM | 37 mM (12.3 equiv.) | 100% | 0.4% | 536 | 2 |
| 2 | 5 mM | 35 mM (7.0 equiv.) | 80% | 1.3% | 630 | 10 |
| 3 | 10 mM | 30 mM (3.0 equiv.) | 71% | 4.2% | 1122 | 66 |

[a] Experiments were performed using a suspension of *E. coli* cells harboring P411-INC-5185 prepared according to the protocol described in **Sections 3.6**, **3.7**, and **3.8**. Reactions were performed in triplicate. Yields reported are the average of three experiments.

[b] Yields reported are the average of three experiments. Yields were calculated based on (*Z*)-**1a**.

**Table 2-S5**. Optimization of *Z/E* ratios of P411-INC-5185-catalyzed cyclopropanation reactions. [a]



| Entry | *Z/E* (x/y) | Yield of 3a[b] | Standard deviation of yield | TTN of 3a | Standard deviation of TTN |
|-------|-------------|----------------|------------------------------|-----------|----------------------------|
| 1 | 1/1 | 100% | 0.4% | 536 | 2 |
| 2 | 0/1 | N.D. | - | N.D. | - |
| 3 | 1/0 | 89% | 14.6 | 838 | 138 |

[a] Experiments were performed using a suspension of *E. coli* cells harboring P411-INC-5185 prepared according to the protocol described in **Sections 3.6**, **3.7**, and **3.8**. Reactions were performed in triplicate. Yields reported are the average of three experiments. N.D. – no product was detected.

[b] Yields reported are the average of three experiments. Yields were calculated based on (*Z*)-**1a**.

**Table 2-S6**. Optimization of $OD_{600}$ of P411-INC-5185-catalyzed cyclopropanation reactions. [a]
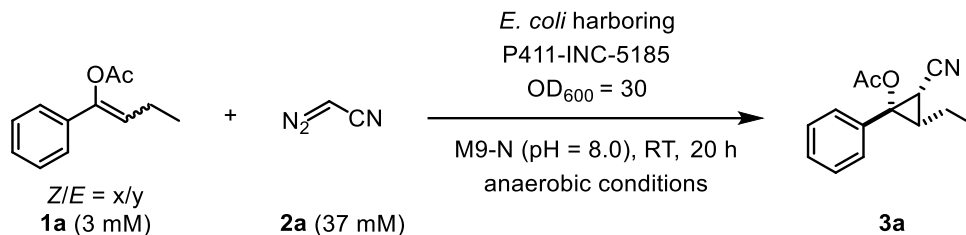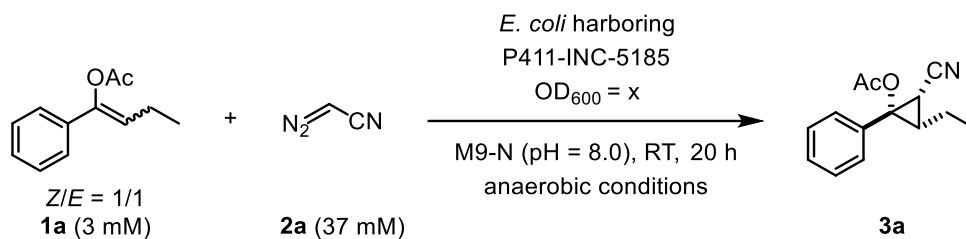


| Entry | $OD_{600}$ of cell suspension | Yield of 3a[b] | Standard deviation of yield | TTN of 3a | Standard deviation of TTN |
|-------|-------------------------------|----------------|------------------------------|-----------|----------------------------|
| 1 | 10 | 13% | 2.6% | 372 | 37 |
| 2 | 30 | 100% | 0.4% | 536 | 2 |

[a] Experiments were performed using a suspension of *E. coli* cells harboring P411-INC-5185 prepared according to the protocol described in **Sections 3.6**, **3.7**, and **3.8**. Reactions were performed in triplicate. Yields reported are the average of three experiments.

[b] Yields reported are the average of three experiments. Yields were calculated based on (*Z*)-**1a**.

**Table 2-S7**. Directed evolution of diastereomer-differentiating carbene transferase P411-INC-5186. [a]
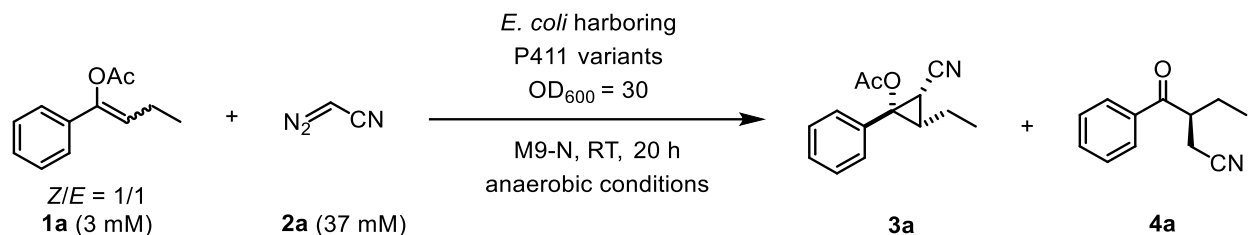


| Entry | Variant | Yield of 3a[b] | SD Yield of 3a | TTN of 3a | SD TTN of 3a | Yield of 4a[c] | SD Yield of 4a | TTN of 4a | SD TTN of 4a |
|---|---|---|---|---|---|---|---|---|---|
| 1 | P411-INC-5185 | 50% (100%) | 0.2% (0.4%) | 536 | 2 | N.D. | - | N.D. | - |
| 2 | P411-INC-5186 (P411-INC-5185 W263M) | 50% (100%) | 1.2% (2.4%) | 378 | 9 | 22% (44%) | 0.5% (1.1%) | 162 | 4 |

[a] Experiments were performed using a suspension of *E. coli* cells harboring variants prepared according to the protocol described in **Sections 3.6**, **3.7**, and **3.8**. Reactions were performed in triplicate. Yields reported are the average of three experiments.

[b] Yields were calculated based on all **1a** added, and those in parentheses were calculated based on (*Z*)-**1a** (toward the formation of **3a**).

[c] Yields were calculated based on all **1a**, and those in parentheses were calculated based on (*E*)-**1a** (toward the formation of **4a**).

**Table 2-S8**. Optimization of *Z/E* ratios of P411-INC-5186-catalyzed diastereomer-differentiating reactions. [a]
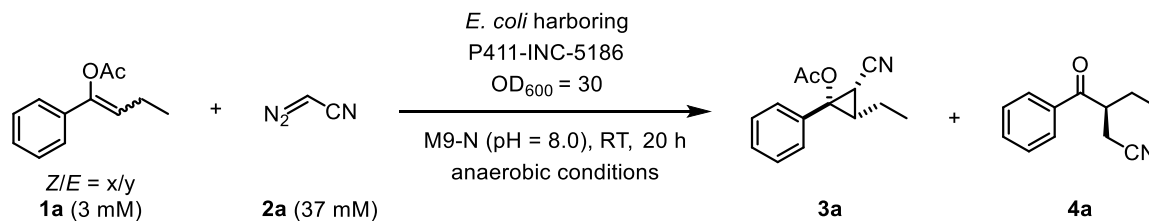


| Entry | x/y | Yield of 3a[b] | SD Yield of 3a | TTN of 3a | SD TTN of 3a | Yield of 4a[c] | SD Yield of 4a | TTN of 4a | SD TTN of 4a |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1/0 | 91% | 4.4% | 667 | 32 | N.D. | - | N.D. | - |
| 2[c] | 0/1 | N.D. | - | N.D. | - | 30 % | 1.6% | 224 | 12 |
| 3 | 1/1 | 100% | 2.4% | 378 | 9 | 44% | 1.1% | 162 | 4 |

[a] Experiments were performed using a suspension of *E. coli* cells harboring P411-INC-5186 prepared according to the protocol described in **Sections 3.6**, **3.7**, and **3.8**. Reactions were performed in triplicate. Yields reported are the average of three experiments. N.D. – no product was detected.

[b] Yields reported are the average of three experiments. Yields were calculated based on (*Z*)-**1a**.

[c] Yields reported are the average of three experiments. Yields were calculated based on (*E*)-**1a**.

# Figures


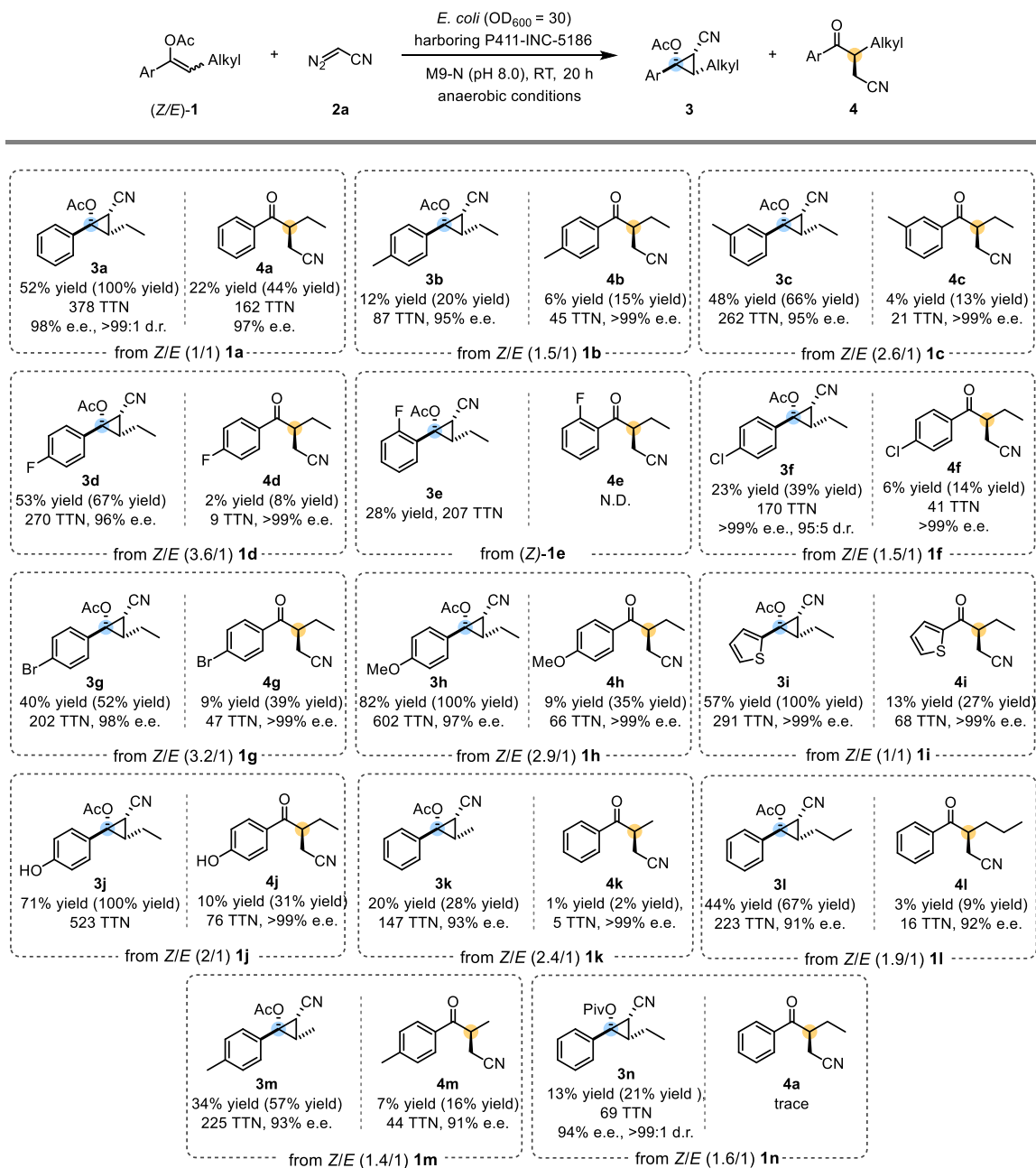
**Figure 2-S1. Substrate scope study for P411-INC-5186-catalyzed diastereomer-differentiating reactions.** Reaction conditions: 3 mM **1**, 37 mM **2a**, *E. coli* whole cells harboring P411-INC-5186 (OD$_{600}$ = 30) in M9-N aqueous buffer (pH 8.0), 10% v/v EtOH (co-solvent), room temperature, anaerobic conditions, 20 h. Yields are based on mixed enol acetates, yields in parenthesis are based

on (*Z*)-isomer for the formation of cyclopropanes **3** or based on (*E*)-isomer for the formation of *α*-branched ketones **4**. Ar, aryl groups; Alkyl, alkyl groups.

## Experimental Methods

### General

Unless otherwise noted, all chemicals and reagents were obtained from commercial suppliers (Sigma-Aldrich, VWR, Alfa Aesar, Acros, Combi Blocks, etc.) and used without further purification. Silica gel chromatography was carried out using AMD Silica Gel 60, 230–400 mesh. $^1$H and $^{13}$C NMR spectra were recorded on a Bruker Prodigy 400 MHz instrument (400 MHz for $^1$H and 101 MHz for $^{13}$C NMR) or a Varian 300 MHz Spectrometer (300 MHz for $^1$H NMR). Chemical shifts ($\delta$) are reported in ppm downfield from tetramethylsilane, using the solvent resonance as the internal standard ($^1$H NMR: $\delta = 7.26$, $^{13}$C NMR: $\delta = 77.36$ for CDCl$_3$). Data for $^1$H NMR are reported as follows: chemical shift ($\delta$ ppm), multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, p = pentet, sext = sextet, m = multiplet, dd = doublet of doublets, dt = doublet of triplets, ddd = doublet of doublet of doublets), coupling constant (Hz), integration. Sonication was performed using a Qsonica Q500 sonicator. High-resolution mass spectra were obtained at the California Institute of Technology Mass Spectral Facility. Samples were analyzed by field ionization (FI) using a JEOL AccuTOF GC-Alpha (JMS-T2000GC) mass spectrometer interfaced with an Agilent 8890 GC system. Ions detected by FI are radical cations.

*Escherichia coli* cells were grown using Luria-Bertani medium or HyperBroth (AthenaES) with 100 μg/mL ampicillin (LB$_{amp}$ or HB$_{amp}$). Primer sequences are available upon request. T5 exonuclease, Phusion polymerase, and *Taq* ligase were purchased from New England Biolabs (NEB, Ipswich, MA). M9-N minimal media (abbreviated as M9-N buffer; pH = 6.0, 7.0, 8.0, and 9.0) were used as buffering systems for whole cells, lysates, and purified proteins, unless otherwise specified. M9-N buffer was used without a carbon source; it contains 47.7 mM Na$_2$HPO$_4$, 22.0 mM KH$_2$PO$_4$, 8.6 mM NaCl, 2.0 mM MgSO$_4$, and 0.1 mM CaCl$_2$.

**Chromatography**

Chemical reactions were monitored using thin layer chromatography (Merck 60 silica gel plates) and a UV lamp for visualization. Reverse-phase high-performance liquid chromatography (HPLC) for analysis were carried out using Agilent 1200 series instruments, with an Agilent C18 column (InfinityLab Poroshell 120 EC-C18, 4.6 x 50 mm, 2.7 μm; Part Number: 699975-902T). Water and acetonitrile containing 0.1% acetic acid were used as eluents. Chiral analyses were conducted using either an Agilent 1200 series HPLC instrument with *n*-hexane and isopropanol as the mobile phase or an SFC using supercritical $CO_2$ and isopropanol as the mobile phase. Enantiomers were separated using one of the following chiral columns: Chiralpak IA (4.6 mm × 25 cm), Chiralpak IB (4.6 mm × 25 cm), Chiralpak IC (4.6 mm × 25 cm), Chiralpak AD-H, or Chiralpak OD-H.

**Cloning and site-saturation mutagenesis**

The genes encoding all enzymes described in this study were cloned using Gibson assembly[2] into vector pET22b(+) (Novagen) between restriction sites *Nde*I and *Xho*I in frame with a *C*-terminal 6×His-tag. Site-saturation mutagenesis was performed using the "22c-trick"[3] or "NNK" as degenerate codons. The PCR products were digested with *Dpn*I, gel purified, and ligated using Gibson Mix™.[2] Without further purification after the Gibson step, 1 μL of the Gibson product was used to transform 50 μL of electrocompetent *E. coli* BL21 *E. cloni*® (Lucigen) cells.

**Expression of P450 and P411 variants in 96-well plates**

Single colonies from $LB_{amp}$ agar plates were picked using sterile toothpicks and cultured in deep-well 96-well plates containing $LB_{amp}$ (400 μL/well) at 37 °C, 80% humidity, and 220 rpm shaking overnight. Subsequently, $HB_{amp}$ (1080 μL/well) in a deep-well plate was inoculated with an aliquot (120 μL/well) of these overnight cultures and allowed to shake for 3 hours at 37 °C, 80% humidity, and 220 rpm. The plates were then cooled on ice for 30 minutes, and the cultures were induced with 0.5 mM isopropyl *β*-D-1-thiogalactopyranoside (IPTG) and 1.0 mM 5-aminolevulinic acid (ALA) (final concentrations). Expression was then conducted at 20 °C, 220 rpm for 18–20 hours.

**Plate reaction screening in whole-cell format**

*E. coli* cells harboring P411 variants in deep-well 96-well plates were pelleted (3,500 g, 5 min, 4 °C) and resuspended in M9-N buffer (360 μL, pH 8.0) by gentle vortexing. The 96-well plates were then transferred to an anaerobic chamber. To deep-well plates of cell suspensions were added the enol acetate substrate **1a** (3 μL/well, 400 mM in EtOH) and diazoacetonitrile **2a** (37 μL/well, 400 mM in EtOH). During the addition, the stock solution of diazoacetonitrile **2a** was kept on an ice bath to minimize evaporation. The plates were sealed with aluminum sealing foil immediately after the addition and shaken in the anaerobic chamber at room temperature and 600 rpm. After 20 hours, the seals were removed, and ethanol (400 μL/well) was added. The plates were tightly sealed with silicone mats, vigorously vortexed, and centrifuged (4,000 x g, 5 min) to precipitate proteins and cell debris. The supernatant (200 μL/well) was filtered through an AcroPrep 96-well filter plate (0.2 μm) into a shallow 96-well plate for HPLC analysis to determine the yield.

**Flask expression of P411 variants**

*E. coli* (*E. cloni* BL21(DE3)) cells carrying plasmid encoding the appropriate P411 variant were grown overnight in 5 mL Luria-Bertani medium supplemented with 0.1 mg/mL ampicillin (LB$_{amp}$). The preculture (1 mL) was used to inoculate 50 mL of Hyperbroth medium supplemented with 0.1 mg/mL ampicillin (HB$_{amp}$) in a 125-mL Erlenmeyer flask. This culture was incubated at 37 °C and 230 rpm for 2.5 hours. The culture was then cooled on ice for 30 min and induced with 0.5 mM IPTG and 1.0 mM ALA (final concentrations). Expression was conducted at 20 °C, 150 rpm for 16–18 hours. Subsequently, the *E. coli* cells were pelleted by centrifugation (4,000 g, 4 min, 4 °C). Media were removed, and the pellets were resuspended to an optical density at 600 nm (OD$_{600}$) of 30 in M9-N minimal medium with pH adjusted to 8.0. Aliquots of the cell suspension (3–4 mL) were used to determine protein concentration after lysis by sonication.

**Hemochrome assay for the determination of hemoprotein concentration**

The concentration of hemoprotein in the clarified lysate was determined by the hemochrome assay.[4] Briefly, 500 μL of the lysate were added to a cuvette and mixed with

500 µL of solution I (0.2 M NaOH, 40% (v/v) pyridine, 500 µM potassium ferricyanide). The UV-Vis spectrum (380–650 nm) of the oxidized $Fe^{III}$ state was recorded immediately. Sodium dithionite (10 µL of 0.5 M solution in 0.5 M NaOH) was added, and the UV-Vis spectrum of the reduced $Fe^{II}$ state was recorded immediately. The pyridine hemochromagen concentration was determined using its Q bands, with the extinction coefficient of 34.7 $mM^{-1}$ $cm^{-1}$ at 557 nm.

**Biotransformation using whole *E. coli* cells**

Suspension of *E. coli* (*E. cloni* BL21(DE3)) cells expressing the appropriate hemoprotein variant in M9-N (pH = 8.0) buffer (typically $OD_{600} = 30$) were transferred to a reaction vial. The headspace of the reaction vial was purged with a stream of argon for at least 15 minutes. Enzymatic reactions were then set up in an anaerobic chamber. To a 2-mL vial were added degassed suspension of *E. coli* expressing P411 variant (typically $OD_{600} = 30$, 360 µL), enol acetate substrates (typically 3 µL of 400 mM stock solution in EtOH), and the diazoacetonitrile solution (typically 37 µL of 400 mM stock solution in EtOH). The concentration of diazo solution was measured by $^1$H-NMR and was kept in an ice bath and added last. The final volume of the biotransformation was set to be 400 µL, with 10% vol EtOH. The reaction vials were then capped and shaken in the anaerobic chamber at room temperature and 600 rpm for 20 hours. After the completion of the reaction, 400 µL ethanol containing 2 mM 1,2,3-trimethoxybenzene internal standard were added to the vial. The resulting mixture was transferred to a 1.7-mL microcentrifuge tube, vigorously vortexed, and centrifuged (14,000 × g, 5 min, 4 °C). A sample of the supernatant (0.2 mL) was transferred to a vial with an insert for HPLC analysis.

To determine the enantiomeric excess (e.e.) of the α-branched ketones **4**, the remaining supernatants of three parallel analytical reactions were combined and transferred to a 2-dram vial and the solvent was removed by blowing air. The α-branched ketones **4** were extracted from the remaining residues with a solution of hexane and EtOAc mixture (1:1) and subjected to normal-phase HPLC to determine the enantiomeric excess (e.e).

The enantiomeric excesses (e.e.) of the polysubstituted cyclopropyl acetates **3** were determined according to the General procedure F:.

**Enzymatic preparative synthesis**

The *E. coli* cell suspension harboring P411-INC-5186 was prepared as described in **Sections 3.6**, **3.7**, and **3.8**. The cell suspension (*E. coli* whole cells harboring P411 variants, $OD_{600}$ = 30, suspended in M9-N aqueous buffer, pH = 8.0) was placed in a sealed flask on ice, and the headspace was purged with a stream of argon for at least 15 minutes. The reaction flask was then transferred to an anaerobic chamber. To the reaction flask with cell suspension ($OD_{600}$ = 30), enol acetate substrate **1** (400 mM EtOH stock solution, final concentration = 10 mM, 1.0 equiv.) and diazoacetonitrile **2a** (400 mM EtOH stock solution, final concentration = 30 mM, 3.0 equiv.) were added to make 10% v/v EtOH (co-solvent). The reaction vial was immediately capped and sealed with parafilm and shaken in the anaerobic chamber at room temperature at 200 rpm for 20 hours. The reaction solution was then extracted with 100 mL 1:1 hexane/ethyl acetate three times. The combined organic layer was then washed with brine and dried over anhydrous $MgSO_4$. The solvent was removed *in vacuo*, and the product was purified by flash chromatography.

**Nucleotide and Amino Acid Sequences**

The genes encoding the heme proteins shown below were cloned using Gibson assembly[2] into vector pET-22b(+) (Novagen) between restriction sites *Nde*I and *Xho*I in frame with a *C*-terminal 6×His-tag.

**Table 2-S9**. Mutations of the P411 variants in this study.

| Name | Mutations relative to the wild type P450<sub>BM3</sub> | Mutations relative to P411-INC-5182 |
|---|---|---|
| P411-INC-5182 | A74G, V78L, A82L, F87A, P142S, T175I, M177L, A184V, S226R, H236Q, E252G, I263Y, T268G, A290V, | - |

| | T327I, A328V, L353V, I366V, C400S, T436L, L437Q, E442K, ΔFAD domain | |
|---|---|---|
| P411-INC-5183 | A74G, V78L, A82L, F87A, P142S, T175I, M177L, A184V, S226R, H236Q, E252G, I263W, T268G, A290V, T327P, A328V, L353V, I366V, C400S, T436L, L437Q, E442K, ΔFAD domain | Y263W, I327P |
| P411-INC-5184 | A74G, V78L, A82L, F87A, P142S, T175I, M177L, A184V, S226R, H236Q, E252G, I263W, T268G, A290V, T327P, A328V, L353V, I366V, C400S, T436L, L437V, E442K, ΔFAD domain | Y263W, I327P, Q437V |
| P411-INC-5185 | N70S, A74G, V78L, A82L, F87A, P142S, T175I, M177L, A184V, S226R, H236Q, E252G, I263W, T268G, A290V, T327P, A328V, L353V, I366V, C400S, T436L, L437V, E442K, ΔFAD domain | N70S, Y263W, I327P, Q437V |
| P411-INC-5186 | N70S, A74G, V78L, A82L, F87A, P142S, T175I, M177L, A184V, S226R, H236Q, E252G, I263M, T268G, A290V, T327P, A328V, L353V, I366V, C400S, T436L, L437V, E442K, ΔFAD domain | N70S, Y263M, I327P, Q437V |

DNA and amino acid sequences of P411-INC-5182:

ATGACAATTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATTT
ACCGTTATTAAACACAGATAAACCGGTTCAAGCTTTGATGAAAATTGCGGATG
AATTAGGAGAAATCTTTAAATTCGAGGCGCCTGGTCGTGTAACGCGCTACTTA
TCAAGTCAGCGTCTAATTAAAGAAGCATGCGATGAATCACGCTTTGATAAAAA
CTTAAGTCAAGGTCTGAAATTTCTGCGTGATTTTCTTGGAGACGGGTTAGCCAC
AAGCTGGACGCATGAAAAAATTGGAAAAAGCGCATAATATCTTACTTCCA
AGCTTTAGTCAGCAGGCAATGAAAGGCTATCATGCGATGATGGTCGATATCGC
CGTGCAGCTTGTTCAAAAGTGGGAGCGTCTAAATGCAGATGAGCATATTGAAG
TATCGGAAGACATGACACGTTTAACGCTTGATACAATTGGTCTTTGCGGCTTTA
ACTATCGCTTTAACAGCTTTTACCGAGATCAGCCTCATCCATTTATTATAAGTC

TGGTCCGTGCACTGGATGAAGTAATGAACAAGCTGCAGCGAGCAAATCCAGA
CGACCCAGCTTATGATGAAAACAAGCGCCAGTTTCAAGAAGATATCAAGGTG
ATGAACGACCTAGTAGATAAAATTATTGCAGATCGCAAAGCAAGGGGTGAAC
AAAGCGATGATTTATTAACGCAGATGCTAAACGGAAAAGATCCAGAAACGGG
TGAGCCGCTTGATGACGGGAACATTCGCTATCAAATTATTACATTCTTATATGC
GGGACACGAAGGTACAAGTGGTCTTTTATCATTTGCGCTGTATTTCTTAGTGAA
AAATCCACATGTATTACAAAAGTAGCAGAAGAAGCAGCACGAGTTCTAGTA
GATCCTGTTCCAAGCTACAAACAAGTCAAACAGCTTAAATATGTCGGCATGGT
CTTAAACGAAGCGCTGCGCTTATGGCCAATTGTTCCTGCGTTTTCCCTATATGC
AAAAGAAGATACGGTGCTTGGAGGAGAATATCCTTTAGAAAAAGGCGACGAA
GTAATGGTTCTGATTCCTCAGCTTCACCGTGATAAACAGTTTGGGGAGACGA
TGTGGAGGAGTTCCGTCCAGAGCGTTTTGAAAATCCAAGTGCGATTCCGCAGC
ATGCGTTTAAACCGTTTGGAAACGGTCAGCGTGCGTCTATCGGTCAGCAGTTC
GCTCTTCATGAAGCAACGCTGGTACTTGGTATGATGCTAAAACACTTTGACTTT
GAAGATCATACAAACTACGAGCTCGATATTAAAGAACTGCAGACGTTAAAAC
CTAAAGGCTTTGTGGTAAAAGCAAAATCGAAAAAAATTCCGCTTGGCGGTATT
CCTTCACCTAGCACTGAACAGTCTGCTAAAAAAGTACGCAAAAAGGCAGAAA
ACGCTCATAATACGCCGCTGCTTGTGCTATACGGTTCAAATATGGGTACCGCT
GAAGGAACGGCGCGTGATTTAGCAGATATTGCAATGAGCAAAGGATTTGCAC
CGCAGGTCGCAACGCTTGATTCACACGCCGGAAATCTTCCGCGCGAAGGAGCT
GTATTAATTGTAACGGCGTCTTATAACGGTCATCCGCCTGATAACGCAAAGCA
ATTTGTCGACTGGTTAGACCAAGCGTCTGCTGATGAAGTAAAAGGCGTTCGCT
ACTCCGTATTTGGATGCGGCGATAAAAACTGGGCTACTACGTATCAAAAGTG
CCTGCTTTTATCGATGAAACGCTTGCCGCTAAAGGGGCAGAAACATCGCTGA
CCGCGGTGAAGCAGATGCAAGCGACGACTTTGAAGGCACATATGAAGAATGG
CGTGAACATATGTGGAGTGACGTAGCAGCCTACTTTAACCTCGACATTGAAAA
CAGTGAAGATAATAAATCTACTCTTTCACTTCAATTTGTCGACAGCGCCGCGG
ATATGCCGCTTGCGAAAATGCACGGTGCGTTTTCAACGCTCGAGCACCACCAC
CACCACCACTGA

MTIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQ

RLIKEACDESRFDKNLSQGLKFLRDFLGDGLATSWTHEKNWKKAHNILLPSFSQQ

AMKGYHAMMVDIAVQLVQKWERLNADEHIEVSEDMTRLTLDTIGLCGFNYRFN

SFYRDQPHPFIISLVRALDEVMNKLQRANPDDPAYDENKRQFQEDIKVMNDLVDK

IIADRKARGEQSDDLLTQMLNGKDPETGEPLDDGNIRYQIITFLYAGHEGTSGLLSF

ALYFLVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRLWPIV

PAFSLYAKEDTVLGGEYPLEKGDEVMVLIPQLHRDKTVWGDDVEEFRPERFENPS

AIPQHAFKPFGNGQRASIGQQFALHEATLVLGMMLKHFDFEDHTNYELDIKELQT

LKPKGFVVKAKSKKIPLGGIPSPSTEQSAKKVRKKAENAHNTPLLVLYGSNMGTA

EGTARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNGHPPDNAKQF

VDWLDQASADEVKGVRYSVFGCGDKNWATTYQKVPAFIDETLAAKGAENIADR

GEADASDDFEGTYEEWREHMWSDVAAYFNLDIENSEDNKSTLSLQFVDSAADMP

LAKMHGAFSTLEHHHHHH*


DNA and amino acid sequences of P411-INC-5186:

ATGACAATTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATTT

ACCGTTATTAAACACAGATAAACCGGTTCAAGCTTTGATGAAAATTGCGGATG

AATTAGGAGAAATCTTTAAATTCGAGGCGCCTGGTCGTGTAACGCGCTACTTA

TCAAGTCAGCGTCTAATTAAAGAAGCATGCGATGAATCACGCTTTGATAAATC

TTTAAGTCAAGGTCTGAAATTTCTGCGTGATTTTCTTGGAGACGGGTTAGCCAC

AAGCTGGACGCATGAAAAAATTGGAAAAAGCGCATAATATCTTACTTCCA

AGCTTTAGTCAGCAGGCAATGAAAGGCTATCATGCGATGATGGTCGATATCGC

CGTGCAGCTTGTTCAAAAGTGGGAGCGTCTAAATGCAGATGAGCATATTGAAG

TATCGGAAGACATGACACGTTTAACGCTTGATACAATTGGTCTTTGCGGCTTTA

ACTATCGCTTTAACAGCTTTTACCGAGATCAGCCTCATCCATTTATTATAAGTC

TGGTCCGTGCACTGGATGAAGTAATGAACAAGCTGCAGCGAGCAAATCCAGA

CGACCCAGCTTATGATGAAAACAAGCGCCAGTTTCAAGAAGATATCAAGGTG

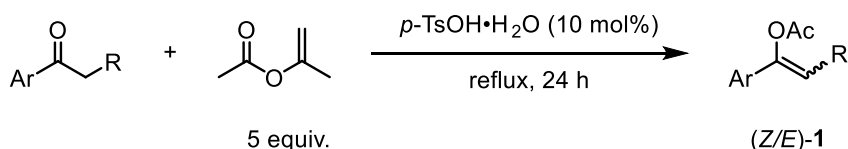ATGAACGACCTAGTAGATAAAATTATTGCAGATCGCAAAGCAAGGGGGTGAAC

AAAGCGATGATTTATTAACGCAGATGCTAAACGGAAAAGATCCAGAAACGGG
TGAGCCGCTTGATGACGGGAACATTCGCTATCAAATTATTACATTCTTAATGG
CGGGACACGAAGGTACAAGTGGTCTTTTATCATTTGCGCTGTATTTCTTAGTGA
AAAATCCACATGTATTACAAAAAGTAGCAGAAGAAGCAGCACGAGTTCTAGT
AGATCCTGTTCCAAGCTACAAACAAGTCAAACAGCTTAAATATGTCGGCATGG
TCTTAAACGAAGCGCTGCGCTTATGGCCACCGGTTCCTGCGTTTTCCCTATATG
CAAAAGAAGATACGGTGCTTGGAGGAGAATATCCTTTAGAAAAAGGCGACGA
AGTAATGGTTCTGATTCCTCAGCTTCACCGTGATAAACAGTTTGGGGAGACG
ATGTGGAGGAGTTCCGTCCAGAGCGTTTTGAAAATCCAAGTGCGATTCCGCAG
CATGCGTTTAAACCGTTTGGAAACGGTCAGCGTGCGTCTATCGGTCAGCAGTT
CGCTCTTCATGAAGCAACGCTGGTACTTGGTATGATGCTAAAACACTTTGACTT
TGAAGATCATACAAACTACGAGCTCGATATTAAAGAACTGGTGACGTTAAAA
CCTAAAGGCTTTGTGGTAAAAGCAAAATCGAAAAAAATTCCGCTTGGCGGTAT
TCCTTCACCTAGCACTGAACAGTCTGCTAAAAAAGTACGCAAAAAGGCAGAA
AACGCTCATAATACGCCGCTGCTTGTGCTATACGGTTCAAATATGGGTACCGC
TGAAGGAACGGCGCGTGATTTAGCAGATATTGCAATGAGCAAAGGATTTGCA
CCGCAGGTCGCAACGCTTGATTCACACGCCGGAAATCTTCCGCGCGAAGGAGC
TGTATTAATTGTAACGGCGTCTTATAACGGTCATCCGCCTGATAACGCAAAGC
AATTTGTCGACTGGTTAGACCAAGCGTCTGCTGATGAAGTAAAAGGCGTTCGC
TACTCCGTATTTGGATGCGGCGATAAAAACTGGGCTACTACGTATCAAAAAGT
GCCTGCTTTTATCGATGAAACGCTTGCCGCTAAAGGGGCAGAAAACATCGCTG
ACCGCGGTGAAGCAGATGCAAGCGACGACTTTGAAGGCACATATGAAGAATG
GCGTGAACATATGTGGAGTGACGTAGCAGCCTACTTTAACCTCGACATTGAAA
ACAGTGAAGATAATAAATCTACTCTTTCACTTCAATTTGTCGACAGCGCCGCG
GATATGCCGCTTGCGAAAATGCACGGTGCGTTTTCAACGCTCGAGCACCACCA
CCACCACCACTGA

MTIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQ
RLIKEACDESRFDKSLSQGLKFLRDFLGDGLATSWTHEKNWKKAHNILLPSFSQQ
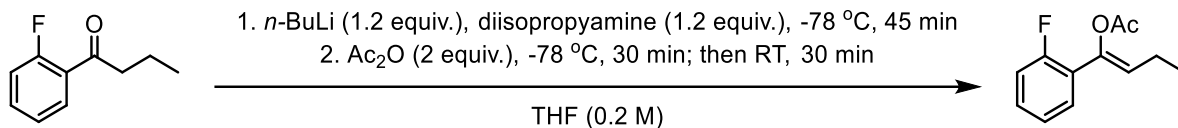AMKGYHAMMVDIAVQLVQKWERLNADEHIEVSEDMTRLTLDTIGLCGFNYRFN

SFYRDQPHPFIISLVRALDEVMNKLQRANPDDPAYDENKRQFQEDIKVMNDLVDK
IIADRKARGEQSDDLLTQMLNGKDPETGEPLDDGNIRYQIITFLMAGHEGTSGLLS
FALYFLVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRLWPP
VPAFSLYAKEDTVLGGEYPLEKGDEVMVLIPQLHRDKTVWGDDVEEFRPERFENP
SAIPQHAFKPFGNGQRASIGQQFALHEATLVLGMMLKHFDFEDHTNYELDIKELV
TLKPKGFVVKAKSKKIPLGGIPSPSTEQSAKKVRKKAENAHNTPLLVLYGSNMGT
AEGTARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNGHPPDNAKQ
FVDWLDQASADEVKGVRYSVFGCGDKNWATTYQKVPAFIDETLAAKGAENIAD
RGEADASDDFEGTYEEWREHMWSDVAAYFNLDIENSEDNKSTLSLQFVDSAADM
PLAKMHGAFSTLEHHHHHH*

## Substrate Syntheses and Characterizations

### General procedure A:  preparation of mixtures of enol acetates 1



$$Ar\text{-}C(=O)\text{-}CH_2R + \text{isopropenyl acetate} \xrightarrow[\text{reflux, 24 h}]{p\text{-TsOH}\cdot H_2O\ (10\ \text{mol\%})} Ar\text{-}C(OAc)=CHR$$

5 equiv.                    (Z/E)-**1**

Enol acetates were prepared using a procedure reported by Hu et al.[5] Ketones and isopropenyl acetate were obtained from commercial sources and were used as received. To a 25-ml Schlenk flask with a stir bar were added ketone (5 mmol), isopropenyl acetate (25 mmol) and p-TsOH•H₂O (0.5 mmol) under nitrogen. The reaction mixture was refluxed for 24 h, then cooled to room temperature. Next, the reaction mixture was diluted with Et₂O (30 mL) and water (10 mL). The aqueous layer was extracted with Et₂O (3 ×30 mL). The combined organic layers were dried over Na₂SO₄, and the solvents were removed under reduced pressure. The resulting crude mixture was purified by column chromatography with hexanes/ethyl acetate (100:1 to 20:1) as eluent to give the desired product **1** as a Z/E mixture. Z/E ratios were measured and determined by ¹H NMR (see 0 for details). Spectral data were in accordance with literature values.[5-6]

**General procedure B:  preparation of 1e**



1. *n*-BuLi (1.2 equiv.), diisopropyamine (1.2 equiv.), -78 °C, 45 min
2. Ac$_2$O (2 equiv.), -78 °C, 30 min; then RT,  30 min

THF (0.2 M)

Synthesis of **1e** was performed using a known procedure[7] with slight modification as follows. (*Z*)-Enol acetate **1e** were prepared using a known procedure[6] but with slight modification as follows. *n*-Butyllithium (9.6 mmol) was added to a solution of diisopropylamine (9.6 mmol) in THF (40 mL) in a flame-dried round-bottom flask under an argon atmosphere at −78 °C. The mixture was stirred for 30 min, and then 1-(2-fluorophenyl)butan-1-one (8 mmol) was added. The resulting mixture was stirred for 45 min, and then acetic anhydride (16 mmol) was added. The reaction was stirred for 30 min at −78 °C and another 30 min at room temperature. The mixture was poured into saturated NaHCO$_3$ (100 mL) and extracted three times with EtOAc (60 mL). The combined organic layers were washed with brine and dried over Na$_2$SO$_4$, and the solvent was removed under reduced atmosphere. The crude product was purified by normal phase chromatography to give the desired product **1e**. Spectral data were in accordance with literature values.[5-6]
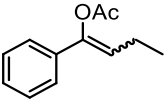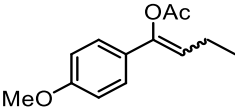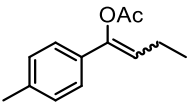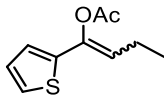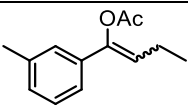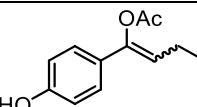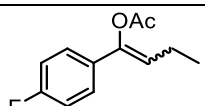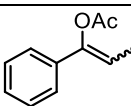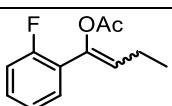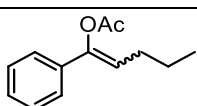
## General procedure C: preparation of (*Z/E*)-1n
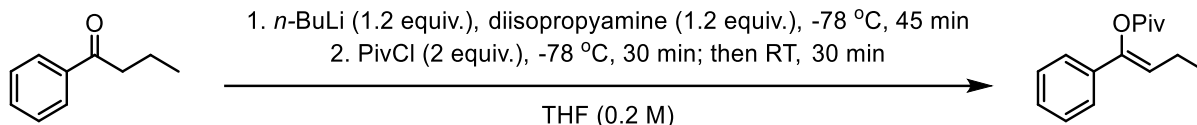


(1) Synthesis of (*Z*)-**1n** was performed using a known procedure[7] with slight modification as follows. n-Butyllithium (3.9 mmol) was added to a solution of diisopropylamine (3.9 mmol) in THF (17 mL) in a flame-dried round-bottom flask under an argon atmosphere at −78 °C. The mixture was stirred for 30 min, and then butyrophenone (3.25 mmol) was added. The resulting mixture was stirred for 45 min, and pivaloyl chloride (6.5 mmol) was added. The reaction was stirred for 30 min at −78 °C and another 30 min at room temperature. The mixture was poured into saturated $NaHCO_3$ (50 mL) and extracted three times with EtOAc (60 mL). The combined organic layers were washed with brine and dried over $Na_2SO_4$, and the solvent was removed under reduced atmosphere. Th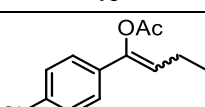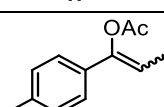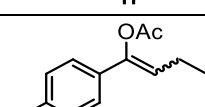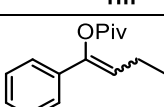e crude product was purified by normal phase chromatography. Spectral data were in accordance with literature values.[5-6]
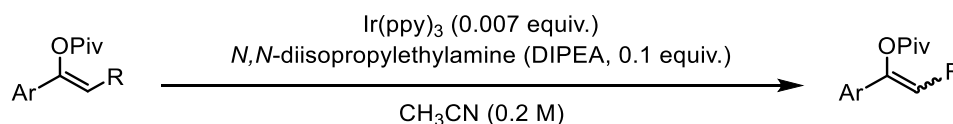


(2) Isomerization of (*Z*)-**1n** was performed using a known procedure[7] with slight modifications as follows. A vial fitted with a rubber septum was charged with Ir(ppy)3 (0.007 equiv., 1.88 mM stock solution of catalyst in $CH_3CN$), (*Z*)-**1n** (1 equiv.), diisopropylethylamine (DIPEA, 0.1 equiv.), and the reaction mixture was degassed via Argon (Ar) gas bubbling for 5–10 min and then left under positive Ar pressure by removing the exit needle. The vial was placed in a light bath (Blue LEDs were purchased from Kessil Co., Ltd. (40 W max., product No. A160WE) for 5 hours. The crude product was purified by normal phase chromatography. The Z/E ratio of **1n** was measured and determined by ^{1}H NMR (see 0 for details).

**Table 2-S10**. *Z/E* ratios of enol acetates **1**. [a]

| 1 | Z/E | 1 | Z/E |
|---|---|---|---|
|  **1a** | 1/1 [b] |  **1h** | 2.9/1 |
|  **1b** | 1.5/1 |  **1i** | 1/1 [b] |
|  **1c** | 2.6/1 |  **1j** | 2.0/1 |
|  **1d** | 3.6/1 |  **1k** | 2.4//1 |
|  **1e** | 1/0 [c] |  **1l** | 1.9/1 |
|  **1f** | 1.5/1 |  **1m** | 1.4/1 |
|  **1g** | 3.2/1 |  **1n** | 1.6/1 [d] |

[a] Enol acetates **1** were prepared according to the protocol described in General procedure A:. *Z/E*- ratios were determined by ^{1}H NMR. Representative examples of NMR spectra were provided below, and all spectral data were in accordance with literature values.[5-6]

[b] Pure (*Z/E*)-isomers were obtained using preparative HPLC and readjusted to mixtures with a ratio of (*Z/E*) = 1:1. Chiral chromatographic separations of the diastereomers of **1a** and **1i** were conducted using a Thar analytical SFC system using a Mettler-Toledo column compartment and an Agilent 1200 series G1315B diode-array detector using isopropanol/$CO_2$ and two 4.6 x 250 mm Chiralcel® AD-H columns.

[c] **1e** was prepared according to the protocol described in General procedure B:

[a] **1n** was prepared according to the protocol described in General procedure C:

## (*Z*)-1-phenylbut-1-en-1-yl acetate ((*Z*)-1a)



(*Z*)-**1a**
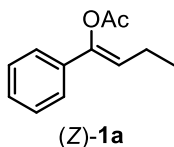
**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.42 – 7.38 (m, 2H), 7.35 – 7.29 (m, 2H), 7.28 – 7.25 (m, 1H), 5.81 (t, *J* = 7.3 Hz, 1H), 2.29 (s, 3H), 2.14 (p, *J* = 7.5 Hz, 2H), 1.07 (t, *J* = 7.6 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 168.96 , 145.70 , 135.14 , 128.61 , 128.18 , 124.46 , 119.94 , 20.80 , 19.83 , 13.65 .

Spectral data were in accordance with literature values.[5-6]

## (*E*)-1-phenylbut-1-en-1-yl acetate ((*E*)-1a)



(*E*)-**1a**
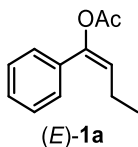
**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.40 – 7.28 (m, 5H), 5.44 (t, *J* = 7.8 Hz, 1H), 2.22 (p, *J* = 7.5 Hz, 2H), 2.14 (s, 3H), 1.04 (t, *J* = 7.5 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 145.88 , 134.72 , 128.44 , 128.29 , 128.23 , 122.13 , 21.19 , 20.74 , 14.61

Spectral data were in accordance with literature values.[5-6]

## 1-(2-fluorophenyl)but-1-en-1-yl acetate (1e)
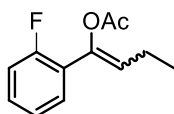


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.32 (td, *J* = 7.8, 1.8 Hz, 1H), 7.26 – 7.19 (m, 1H), 7.10 (td, *J* = 7.6, 1.3 Hz, 1H), 7.04 (ddd, *J* = 11.6, 8.2, 1.2 Hz, 1H), 5.86 (t, *J* = 7.3 Hz, 1H), 2.25 (s, 3H), 2.21 – 2.07 (m, 2H), 1.07 (t, *J* = 7.5 Hz, 3H).

Spectral data were in accordance with literature values.[5-6]

**1-(4-methoxyphenyl)but-1-en-1-yl acetate (1h)**
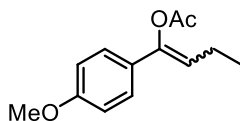


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.35 – 7.26 (m, 2H), 6.86 (dd, *J* = 13.2, 8.8 Hz, 2H), 5.52 (dt, *J* = 123.6, 7.5 Hz, 1H), 3.81 (d, *J* = 4.8 Hz, 3H), 2.45 – 1.99 (m, 5H), 1.05 (dt, *J* = 8.8, 7.5 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 169.02 , 159.62 , 145.48 , 129.59 , 127.88 , 125.85 , 121.15 , 118.06 , 114.01 , 113.70 , 77.36 , 55.45 , 55.40 , 21.25 , 20.84 , 20.78 , 19.77 , 14.65 , 13.76 .

Spectral data were in accordance with literature values.[5-6]

**1-(thiophen-2-yl)but-1-en-1-yl acetate**



**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.16 (dd, *J* = 4.9, 1.3 Hz, 1H), 6.99 – 6.89 (m, 2H), 5.73 (t, *J* = 7.4 Hz, 1H), 2.28 (s, 3H), 2.10 (p, *J* = 7.5 Hz, 2H), 1.05 (t, *J* = 7.5 Hz, 3H).

(*Z*)-**1i**, 98% stereopurity: **¹H NMR** (400 MHz, Chloroform-*d*) δ 5.73 (t, *J* = 7.4 Hz, 98%), 5.42 (t, *J* = 7.6 Hz, 2%); and (*E*)-**1i**, 92% stereopurity: **¹H NMR** (400 MHz, Chloroform-*d*) δ 5.73 (t, *J* = 7.4 Hz, 8%), 5.42 (t, *J* = 7.6 Hz, 92%).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 168.67 , 139.23 , 127.48 , 124.66 , 123.26 , 119.52 , 20.71 , 19.72 , 13.54 .
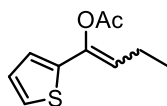
Spectral data were in accordance with literature values.[5-6]

**1-phenylbut-1-en-1-yl pivalate (1n)**
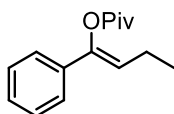


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.33 – 7.29 (m, 2H), 7.26 – 7.21 (m, 2H), 7.20 – 7.17 (m, 1H), 5.71 (t, *J* = 7.3 Hz, 1H), 2.04 (p, *J* = 7.5 Hz, 2H), 1.30 (d, *J* = 0.6 Hz, 9H), 0.99 (t,

$J = 7.5$ Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 145.88 , 128.57 , 128.05 , 124.45 , 119.73 , 39.32 , 27.45 , 19.56 , 13.69 .

Spectral data were in accordance with literature values.[5-6]

**General procedure D:  preparation of diazoacetonitrile 2a:**
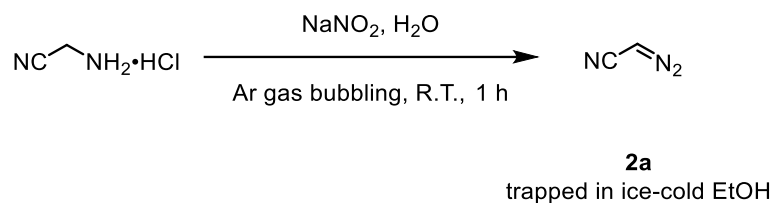


The stock solution of diazo compound **2** was prepared using a modified procedure reported by Hock et al.[8] To a vigorously stirring solution of 1.3 g (14 mmol) of aminoacetonitrile hydrochloride in 1 mL of water, a solution of 0.96 g NaNO₂ in 2 mL of water was added slowly (over 1 hour) at room temperature using a syringe pump. The rapidly evolving yellow gas was carefully bubbled (carrier argon gas was used) via PTFE tubing into 7 mL of ethanol placed in a 7.5-mL sealed vial which is chilled on an ice bath. Upon the completion of addition, the aqueous mixture was further bubbled for 20 min to maximize the stock solution concentration. The vial containing the diazo ethanol solution was carefully removed from the ice bath and stored at –20 °C. The substrate may deactivate/polymerize within two days to yield a black, insoluble mixture, therefore it needed to be prepared freshly. The concentration of the stock solution was measured by ¹H NMR with non-deuterated DMSO as internal standard. The diazo solution concentration was then adjusted to 400 mM for biotransformations unless noted.

**CAUTION:** Diazo compounds are toxic and potentially explosive and should be handled with care in a well-ventilated hood. Do not attempt to concentrate or isolate this compound.

**Characterizations of Standard Products**

**(*1S,2S,3S*)-2-cyano-3-ethyl-1-phenylcyclopropyl acetate (3a)**
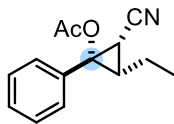


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.39 – 7.28 (m, 5H), 2.19 (d, *J* = 9.3 Hz, 1H), 2.13 (s, 3H), 1.93 – 1.84 (m, 1H), 1.80 – 1.74 (m, 1H), 1.73 – 1.63 (m, 1H), 1.20 (t, *J* = 7.3 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 169.72 , 137.14 , 128.71 , 128.68 , 127.39 , 116.25 , 65.11 , 31.57 , 20.95 , 18.46 , 15.39 , 12.79 .

**Physical State**: white solid.

**HRMS** (FI): calcd for $C_{14}H_{15}NO_2$ [M]$^{+•}$ 229.10887; found 229.10973.

**(*1S,2S,3S*)-2-cyano-3-ethyl-1-(p-tolyl)cyclopropyl acetate (3b)**
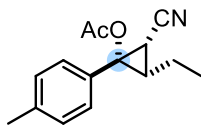


**¹H NMR** (400 MHz, Chloroform-*d*): (400 MHz, Chloroform-*d*) δ 7.30 – 7.26 (m, 2H), 7.21 – 7.06 (m, 2H), 2.33 (s, 3H), 2.14 (d, *J* = 9.3 Hz, 1H), 2.11 (s, 3H), 1.92 – 1.83 (m, 1H), 1.77 – 1.72 (m, 1H), 1.71 – 1.58 (m, 1H), 1.19 (t, *J* = 7.3 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 139.16 , 134.62 , 129.71 , 127.98 , 77.50 , 65.50 , 31.85 , 21.62 , 21.38 , 18.86 , 15.65 , 13.21 .

**Physical State**: colorless oil.

**HRMS** (FI): calcd for $C_{15}H_{17}NO_2$ [M]$^{+•}$ 243.12455; found 243.12538.

**(*1S,2S,3S*)-2-cyano-3-ethyl-1-(m-tolyl)cyclopropyl acetate (3c)**
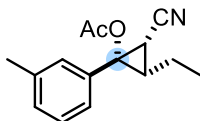


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.28 – 7.19 (m, 2H), 7.21 – 7.13 (m, 2H), 7.20 – 7.06 (m, 2H), 2.35 (s, 4H), 2.17 (d, *J* = 9.3 Hz, 1H), 2.13 (s, 3H), 1.93 – 1.83 (m, 1H), 1.80 – 1.73 (m, 1H), 1.73 – 1.62 (m, 1H), 1.20 (t, *J* = 7.3 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 169.79 , 138.47 , 137.12 , 129.52 , 128.61 , 127.92 , 124.46 , 116.33 , 65.18 , 31.59 , 21.52 , 21.02 , 18.51 , 15.40 , 12.84 .

**Physical State**: colorless oil.

**HRMS** (FI): calcd for $C_{15}H_{17}NO_2$ [M]⁺• 243.12432; found 243.12538.

**(*1S,2S,3S*)-2-cyano-3-ethyl-1-(4-fluorophenyl)cyclopropyl acetate (3d)**
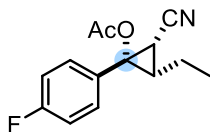


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.46 – 7.36 (m, 2H), 7.09 – 6.97 (m, 2H), 2.14 (d, *J* = 5.2 Hz, 1H), 2.11 (s, 2H), 1.93 – 1.81 (m, 1H), 1.76 – 1.62 (m, 2H), 1.20 (t, *J* = 7.2 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 169.84 , 164.03 , 161.55 , 133.07 (d, *J* = 3.3 Hz), 130.25 (d, *J* = 8.5 Hz), 116.10 (d, *J* = 8.5 Hz), 115.72 (d, *J* = 21.9 Hz), 64.72 , 31.44 , 21.01 , 18.49 , 15.30 , 12.87 .

**¹⁹F NMR** (377 MHz, Chloroform-*d*) δ -112.15.

**Physical State**: colorless oil.

**HRMS** (FI): calcd for $C_{14}H_{14}NO_2F$ [M]⁺• 247.09952; found 247.10031.

**(*1S,2S,3S*)-2-cyano-3-ethyl-1-(2-fluorophenyl)cyclopropyl acetate (3e)**
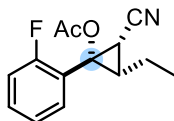


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.58 (td, *J* = 7.5, 1.8 Hz, 1H), 7.33 (dddd, *J* = 8.3, 7.4, 5.2, 1.8 Hz, 1H), 7.13 (td, *J* = 7.6, 1.2 Hz, 1H), 7.05 (ddd, *J* = 10.4, 8.2, 1.2 Hz, 1H), 2.23 (d, *J* = 9.2 Hz, 1H), 2.09 (s, 3H), 1.96 – 1.89 (m, 1H), 1.80 – 1.71 (m, 1H), 1.71 – 1.66 (m, 1H), 1.21 (t, *J* = 7.1 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 169.84 , 164.03 , 161.55 , 133.07 (d, *J* = 3.3 Hz), 130.25 (d, *J* = 8.5 Hz), 116.10 (d, *J* = 8.5 Hz), 115.72 (d, *J* = 21.9 Hz), 64.72 , 31.44 , 21.01 , 18.49 , 15.30 , 12.87 . ¹³C NMR (101 MHz, Chloroform-*d*) δ 170.07 , 162.91 , 160.41 , 132.06 (d, *J* = 2.8 Hz), 131.30 (d, *J* = 8.5 Hz), 124.11 , 124.00 (d, *J* = 3.7 Hz), 116.24 , 116.05 (d, *J* = 21.0 Hz), 60.89 , 30.82 (d, *J* = 1.2 Hz), 20.96 , 18.39 , 14.58 (d, *J* = 2.1 Hz), 12.64 .

**¹⁹F NMR** (377 MHz, Chloroform-*d*) δ -114.47.

**Physical State**: colorless oil.

**HRMS** (FI): calcd for $C_{14}H_{14}NO_2F$ [M]$^{+\bullet}$ 247.09949; found 247.10031.

**(*1S,2S,3S*)-2-cyano-3-ethyl-1-(4-chlorophenyl)cyclopropyl acetate (3f)**
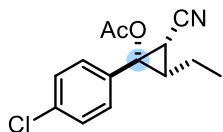


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.65 – 7.58 (m, 2H), 7.44 – 7.33 (m, 2H), 2.19 – 2.13 (m, 1H), 1.93 (s, 3H), 1.92 – 1.80 (m, 3H), 1.12 (t, *J* = 2.1 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 169.82 , 136.23 , 131.96 , 129.46 , 123.03 , 115.97 , 64.66 , 31.55 , 29.85 , 20.98 , 18.46 , 15.44 , 12.84 .

**Physical State**: colorless oil.

**HRMS** (FI): calcd for $C_{14}H_{14}NO_2Cl$ [M]$^{+\bullet}$ 263.07131; found 263.07143.

**(*1S,2S,3S*)-2-cyano-3-ethyl-1-(4-bromophenyl)cyclopropyl acetate (3g)**
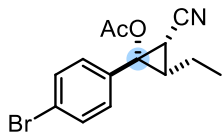


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.54 – 7.41 (m, 2H), 7.27 (dd, *J* = 6.5, 2.1 Hz, 2H), 2.15 (d, *J* = 9.3 Hz, 1H), 2.13 (s, 3H), 1.88 (dddd, *J* = 12.9, 7.5, 5.0, 2.1 Hz, 1H), 1.76 – 1.72 (m, 1H), 1.70 – 1.63 (m, 1H), 1.19 (t, *J* = 7.2 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 169.68 , 131.82 , 129.33 , 122.90 , 115.83 , 99.99, 64.52 , 20.84 , 18.32 , 15.31 .

**Physical State**: colorless oil.

**HRMS (**FI): calcd for C₁₄H₁₄NO₂Br [M]⁺˙ 307.01999; found 307.02024.

**(*1S,2S,3S*)-2-cyano-3-ethyl-1-(4-methoxyphenyl)cyclopropyl acetate (3h)**
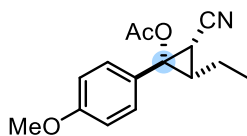


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.41 – 7.31 (m, 2H), 6.90 – 6.81 (m, 2H), 3.79 (s, 3H), 2.11 (s, 1H), 2.09 (s, 3H), 1.95 – 1.81 (m, 1H), 1.78 – 1.71 (m, 1H), 1.71 – 1.61 (m, 1H), 1.20 (t, *J* = 7.3 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 169.85 , 159.90 , 129.83 , 129.27 , 116.48 , 113.93 , 65.11 , 55.43 , 31.32 , 21.04 , 18.52 , 15.10 , 12.87 .

**Physical State**: colorless oil.

**HRMS (**FI): calcd for C₁₅H₁₇NO₃ [M]⁺˙ 259.11850; found 259.12029.

**(*1S,2S,3S*)-2-cyano-3-ethyl-1-(thiophen-2-yl)cyclopropyl acetate (3i)**
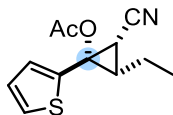


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.28 (dd, *J* = 5.1, 1.2 Hz, 1H), 7.11 (dd, *J* = 3.6, 1.3 Hz, 1H), 6.93 (dd, *J* = 5.1, 3.6 Hz, 1H), 2.22 (d, *J* = 9.5 Hz, 1H), 2.14 (s, 3H), 1.91 – 1.84 (m, 2H), 1.68 – 1.64 (m, 1H), 1.19 (t, *J* = 7.3 Hz, 4H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 169.78 , 140.74 , 127.84 , 126.98 , 126.72 , 115.91 , 60.53 , 32.75 , 29.85 , 21.01 , 18.54 , 17.18 , 12.72 .

**Physical State**: off-white solid.

**HRMS** (FI): calcd for $C_{12}H_{13}NO_2S$ [M]$^{+\bullet}$ 235.06513; found 235.06615.

**(*1S,2S,3S*)-2-cyano-3-ethyl-1-(4-hydroxyphenyl)cyclopropyl acetate (3j)**
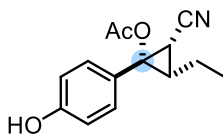


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.35 – 7.29 (m, 2H), 6.79 – 6.74 (m, 2H), 2.10 (s, 3H), 2.09 (d, *J* = 6.1 Hz, 4H), 1.90 – 1.84 (m, 1H), 1.74 – 1.64 (m, 2H), 1.20 (t, *J* = 7.3 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 170.07 , 156.09 , 130.15 , 129.44 , 116.48 , 115.47 , 65.16 , 31.36 , 21.12 , 18.55 , 15.14 .

**Physical State**: white solid.

**HRMS** (FI): calcd for $C_{14}H_{15}NO_3$ [M]$^{+\bullet}$ 245.10519; found 245.10427.

**(*1S,2S,3S*)-2-cyano-3-methyl-1-phenylcyclopropyl acetate (3k)**
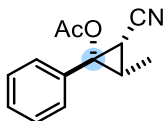


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.40 – 7.26 (m, 5H), 2.14 (s, 3H), 2.09 (d, *J* = 9.6 Hz, 1H), 2.06 – 1.97 (m, 1H), 1.43 (d, *J* = 6.3 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 169.79 , 137.12 , 128.81 , 128.75 , 127.34 , 116.16 , 64.97 , 24.39 , 20.94 , 16.31 , 9.68 .

**Physical State**: white solid.

**HRMS** (FI): calcd for $C_{13}H_{13}NO_2$ [M]$^{+•}$ 215.09332; found 215.09408.

**(*1S,2S,3S*)-2-cyano-1-phenyl-3-propylcyclopropyl acetate (3l)**
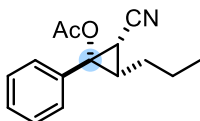


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.40 – 7.27 (m, 5H), 2.21 (d, *J* = 9.5 Hz, 1H), 2.14 (s, 3H), 1.89 – 1.76 (m, 2H), 1.72 – 1.57 (m, 3H), 1.04 (t, *J* = 7.2 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 169.80 , 137.23 , 128.76 , 128.74 , 127.35 , 116.38 , 64.99 , 29.96 , 26.90 , 21.71 , 21.03 , 15.57 , 14.04 .

**Physical State**: colorless oil.

**HRMS** (FI): calcd for $C_{15}H_{17}NO_2$ [M]$^{+•}$ 243.12491; found 243.12538.

**(*1S,2S,3S*)-2-cyano-3-methyl-1-(p-tolyl)cyclopropyl acetate (3m)**
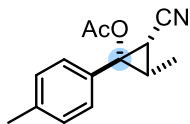


**$^1$H NMR** (400 MHz, Chloroform-*d*) δ 7.30 – 7.25 (m, 2H), 7.19 – 7.10 (m, 2H), 2.33 (s, 3H), 2.13 (s, 3H), 2.09 – 1.93 (m, 2H), 1.42 (d, *J* = 6.2 Hz, 3H).

**$^{13}$C NMR** (101 MHz, Chloroform-*d*) δ 169.83 , 138.87 , 134.23 , 129.39 , 127.56 , 116.30 , 64.99 , 24.27 , 21.29 , 20.99 , 16.17 , 9.72 .

**Physical State**: colorless oil.

**HRMS** (FI): calcd for $C_{14}H_{15}NO_2$ [M]$^{+\bullet}$ 229.11028; found 229.10988.

**(*1S,2S,3S*)-2-cyano-3-ethyl-1-phenylcyclopropyl pivalate (3n)**
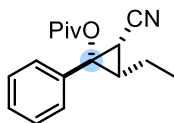


**$^1$H NMR** (400 MHz, Chloroform-*d*) δ 7.35 – 7.28 (m, 5H), 2.24 – 2.18 (m, 1H), 1.90 – 1.81 (m, 1H), 1.80 – 1.74 (m, 1H), 1.74 – 1.64 (m, 1H), 1.24 (s, 9H), 1.22 – 1.18 (m, 3H).

**$^{13}$C NMR** (101 MHz, Chloroform-*d*) δ 177.04 , 137.40 , 128.78 , 128.63 , 126.96 , 116.19 , 77.36 , 64.51 , 39.03 , 31.70 , 29.86 , 27.01 , 18.46 , 15.47 , 12.88 .

**Physical State**: white solid.

**HRMS** (FI): calcd for $C_{17}H_{21}NO_2$ [M]$^{+\bullet}$ 271.15670; found 271.15723.

**Synthesis of Reference Compounds**

**General procedure E:  preparation of $\alpha$-branched ketones 4:**
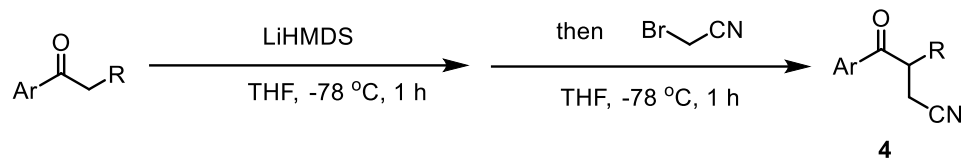


$\alpha$-Branched ketones were prepared using a modified procedure reported by Kleij et al.[9] A solution of lithium bis(trimethylsilyl)amide (LiHMDS, 1.0 M in THF, 12.0 mmol, 1.2 equiv.) in THF (10 mL) was cooled to −78 ℃. A solution of ketone (10 mmol, 1.0 equiv.) in THF (10 mL) was added dropwise, and the resultant solution was stirred at -78 ℃ for 1 h. After that, a solution of bromoacetonitrile (12.0 mmol, 1.2 equiv) in THF (10 mL) was added dropwise and the reaction mixture was stirred at -78 ℃ for 1 h, quenched with hydrochloric acid (1 N) and extracted with Et₂O (3 × 20 mL). The combined organic layers were dried over anhydrous Na₂SO₄, filtered and evaporated under reduced pressure. The crude product was then purified by flash chromatography affording product $\alpha$-branched ketones **4**.

### 3-benzoylpentanenitrile (4a)
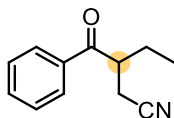


**¹H NMR** (400 MHz, Chloroform-*d*) δ 8.03 – 7.87 (m, 2H), 7.67 – 7.55 (m, 1H), 7.55 – 7.44 (m, 2H), 3.77 (ddd, *J* = 13.0, 6.9, 6.0 Hz, 1H), 2.74 (dd, *J* = 16.9, 6.8 Hz, 1H), 2.64 (dd, *J* = 16.9, 7.2 Hz, 1H), 1.99 – 1.88 (m, 1H), 1.85 – 1.70 (m, 1H), 0.91 (t, *J* = 7.5 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 199.91 , 135.63 , 133.86 , 129.04 , 128.48 , 118.75 , 43.95 , 25.25 , 18.00 , 10.49 .

**Physical State**: colorless oil.

**HRMS** (ESI): calcd for $C_{12}H_{14}NO$ [M+H]$^+$ 188.1075; found 188.1080.

### 3-(4-methylbenzoyl)pentanenitrile (4b)



**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.92 – 7.76 (m, 2H), 7.39 – 7.26 (m, 2H), 3.74 (ddd, *J* = 13.1, 7.1, 6.1 Hz, 1H), 2.72 (dd, *J* = 16.9, 6.7 Hz, 1H), 2.63 (dd, *J* = 16.9, 7.4 Hz, 1H), 2.43 (s, 3H), 1.97 – 1.87 (m, 1H), 1.84 – 1.71 (m, 1H), 0.90 (t, *J* = 7.5 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 199.52 , 144.91 , 133.19 , 129.75 , 128.66 , 118.86 , 43.83 , 25.41 , 21.84 , 18.13 , 10.55 .

**Physical State**: colorless oil.

**HRMS** (ESI): calcd for $C_{13}H_{16}NO$ [M+H]$^+$ 202.1232; found 202.1215.
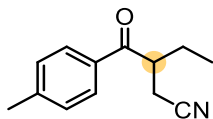
### 3-(3-methylbenzoyl)pentanenitrile (4c)



**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.78 – 7.68 (m, 2H), 7.46 – 7.34 (m, 2H), 3.80 – 3.72 (m, 1H), 2.73 (dd, *J* = 16.9, 6.7 Hz, 1H), 2.63 (dd, *J* = 17.0, 7.3 Hz, 1H), 1.99 – 1.85 (m, 1H), 1.83 – 1.71 (m, 1H), 0.90 (t, *J* = 7.5 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 200.14 , 138.96 , 135.73 , 134.66 , 129.00 , 128.89 , 125.69 , 118.80 , 43.96 , 25.27 , 21.51 , 18.05 , 10.48 .

**Physical State**: colorless oil.
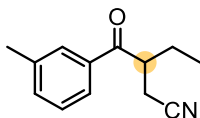
**HRMS** (ESI): calcd for $C_{13}H_{16}NO$ [M+H]⁺ 202.1232; found 202.1238.

### 3-(4-fluorobenzoyl)pentanenitrile (4d)



**¹H NMR** (400 MHz, Chloroform-*d*) δ 8.03 – 7.93 (m, 2H), 7.22 – 7.12 (m, 2H), 3.78 – 3.67 (m, 1H), 2.74 (dd, *J* = 16.9, 6.9 Hz, 1H), 2.64 (dd, *J* = 16.9, 7.1 Hz, 1H), 2.01 – 1.85 (m, 1H), 1.84 – 1.70 (m, 1H), 0.91 (t, *J* = 7.5 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 198.34 , 167.53 , 164.98 , 132.09 (d, *J* = 2.9 Hz), 131.24 (d, *J* = 9.4 Hz), 118.65 , 116.28 (d, *J* = 22.0 Hz), 44.00 , 25.38 , 18.10 , 10.54 .

**¹⁹F NMR** (377 MHz, Chloroform-*d*) δ -103.74 .

**Physical State**: colorless oil.

**HRMS** (ESI): calcd for $C_{12}H_{13}NOF$ [M+H]⁺ 206.0981; found 206.0970.

**3-(2-fluorobenzoyl)pentanenitrile (4e)**



**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.84 (td, *J* = 7.6, 1.9 Hz, 1H), 7.57 (dddd, *J* = 8.3, 7.1, 5.1, 1.9 Hz, 1H), 7.29 – 7.25 (m, 1H), 7.16 (ddd, *J* = 11.4, 8.3, 1.1 Hz, 1H), 3.68 (qd, *J* = 6.7, 5.2 Hz, 1H), 2.77 (ddd, *J* = 16.9, 6.7, 1.0 Hz, 1H), 2.63 (ddd, *J* = 16.9, 6.8, 0.9 Hz, 1H), 1.99 – 1.90 (m, 1H), 1.71 (dddd, *J* = 14.2, 7.6, 6.7, 0.9 Hz, 1H), 0.93 (t, *J* = 7.5 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 198.30 (d, *J* = 4.2 Hz), 161.52 (d, *J* = 253.8 Hz), 135.32 (d, *J* = 9.2 Hz), 131.27 (d, *J* = 2.5 Hz), 125.08 (d, *J* = 3.4 Hz), 124.59 (d, *J* = 12.5 Hz), 118.67 , 116.90 (d, *J* = 23.9 Hz), 48.51 (d, *J* = 7.6 Hz), 24.24 (d, *J* = 1.4 Hz), 17.46 , 10.55 .

**¹⁹F NMR** (377 MHz, Chloroform-*d*) δ -110.81 .

**Physical State**: colorless oil.

**HRMS** (ESI): calcd for $C_{12}H_{13}NOF$ [M+H]⁺ 206.0981; found 206.0964.

**3-(4-chlorobenzoyl)pentanenitrile (4f)**



**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.92 – 7.74 (m, 2H), 7.64 – 7.36 (m, 2H), 3.77 – 3.66 (m, 1H), 2.73 (dd, *J* = 17.0, 6.9 Hz, 1H), 2.64 (dd, *J* = 16.9, 7.1 Hz, 1H), 1.98 – 1.84 (m, 1H), 1.80 – 1.70 (m, 1H), 0.91 (t, *J* = 7.5 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 198.75 , 140.50 , 133.96 , 129.92 , 129.44 , 118.58 , 44.05 , 25.31 , 18.02 , 10.51 .

**Physical State**: colorless oil.

**HRMS** (ESI): calcd for $C_{12}H_{13}NOCl$ [M+H]⁺ 222.0686; found 222.0685.

**3-(4-bromobenzoyl)pentanenitrile (4g)**



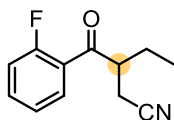**1H NMR** (400 MHz, Chloroform-*d*) δ 7.90 – 7.68 (m, 2H), 7.79 – 7.52 (m, 2H), 3.75 – 3.66 (m, 1H), 2.74 (dd, *J* = 16.9, 6.9 Hz, 1H), 2.64 (dd, *J* = 16.9, 7.1 Hz, 1H), 1.98 – 1.88 (m, 1H), 1.80 – 1.73 (m, 1H), 0.91 (t, *J* = 7.5 Hz, 3H).

**13C NMR** (101 MHz, Chloroform-*d*) δ 198.83 , 134.24 , 132.32 , 129.87 , 129.14 , 118.43 , 77.22 , 43.91 , 25.17 , 17.88 , 10.37 .

**Physical State**: colorless oil.

**HRMS** (ESI): calcd for $C_{12}H_{13}NOBr$ [M+H]$^+$ 266.0181; found 266.0200.
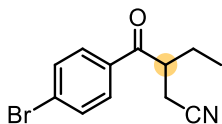
**3-(4-methoxybenzoyl)pentanenitrile (4h)**



**1H NMR** (400 MHz, Chloroform-*d*) δ 8.11 – 7.89 (m, 2H), 6.97 (dd, *J* = 9.0, 2.4 Hz, 2H), 3.88 (s, 1H), 3.78 – 3.66 (m, 1H), 2.72 (ddd, *J* = 16.9, 6.8, 2.2 Hz, 1H), 2.63 (ddd, *J* = 17.0, 7.3, 2.3 Hz, 1H), 1.98 – 1.86 (m, 1H), 1.85 – 1.71 (m, 1H), 0.91 (td, *J* = 7.5, 2.2 Hz, 3H).

**13C NMR** (101 MHz, Chloroform-*d*) δ 198.33 , 164.17 , 130.90 , 128.67 , 118.93 , 114.23 , 55.70 , 43.59 , 25.57 , 18.28 , 10.63 .

**Physical State**: colorless oil.

**HRMS** (ESI): calcd for $C_{13}H_{16}NO_2$ [M+H]$^+$ 218.1181; found 218.1163.

**3-(thiophene-2-carbonyl)pentanenitrile (4i)**



**1H NMR** (400 MHz, Chloroform-*d*) δ 7.74 (ddd, *J* = 15.2, 4.4, 1.1 Hz, 2H), 7.18 (dd, *J* = 5.0, 3.8 Hz, 1H), 3.62 – 3.50 (m, 1H), 2.64 (dd, *J* = 17.0, 7.4 Hz, 1H), 1.96 (ddd, *J* = 14.3, 7.6, 6.8 Hz, 1H), 1.84 (ddd, *J* = 12.4, 7.5, 6.0 Hz, 1H), 0.95 (t, *J* = 7.5 Hz, 3H).

$^{13}$**C NMR** (101 MHz, Chloroform-*d*) δ 192.53 , 143.05 , 135.07 , 132.60 , 128.47 , 118.34 , 99.98 , 77.22 , 45.58 , 25.89 , 18.48 , 10.79 .

**Physical State**: colorless oil.
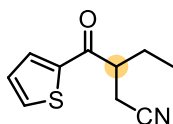
**HRMS** (ESI): calcd for $C_{10}H_{12}NOS$ [M+H]$^{+}$ 194.0640; found 194.0646.

### 3-(4-hydroxybenzoyl)pentanenitrile (4j)
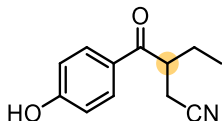


$^{1}$**H NMR** (400 MHz, Chloroform-*d*) δ 7.84 (d, *J* = 10.0 Hz, 2H), 6.93 – 6.86 (m, 2H), 3.77 – 3.63 (m, 1H), 2.74 (dd, *J* = 16.6, 7.5 Hz, 1H), 2.62 (dd, *J* = 16.6, 7.1 Hz, 1H), 1.95 – 1.81 (m, 1H), 1.80 – 1.65 (m, 1H), 0.90 (t, *J* = 7.7 Hz, 3H).

$^{13}$**C NMR** (101 MHz, Chloroform-*d*) δ 199.13 , 161.79 , 131.26 , 128.06 , 118.98 , 115.95 , 43.57 , 25.70 , 18.26 , 10.72 .

**Physical State**: colorless oil.

**HRMS** (ESI): calcd for $C_{12}H_{14}NO_2$ [M+H]$^{+}$ 204.1025; found 204.1011.

### 3-methyl-4-oxo-4-phenylbutanenitrile (4k)
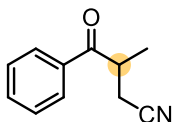


$^{1}$**H NMR** (400 MHz, Chloroform-*d*) δ 7.91 – 7.83 (m, 2H), 7.59 – 7.50 (m, 1H), 7.48 – 7.38 (m, 2H), 3.81 – 3.67 (m, 1H), 2.68 (dd, *J* = 16.9, 5.8 Hz, 1H), 2.55 (dd, *J* = 16.9, 8.1 Hz, 1H), 1.33 (d, *J* = 7.2 Hz, 3H).

$^{13}$**C NMR** (101 MHz, Chloroform-*d*) δ 199.90 , 134.69 , 133.82 , 128.97 , 128.48 , 118.61 , 77.23 , 38.11 , 20.40 , 18.12 .

**Physical State**: colorless oil.

**HRMS** (ESI): calcd for $C_{11}H_{12}NO$ [M+H]$^{+}$ 174.0919; found 174.0914.

**3-benzoylhexanenitrile (4l)**
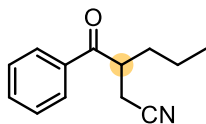


**¹H NMR** (400 MHz, Chloroform-*d*) δ 8.01 – 7.87 (m, 2H), 7.67 – 7.57 (m, 1H), 7.55 – 7.46 (m, 2H), 3.83 – 3.77 (m, 1H), 2.75 (dd, *J* = 16.9, 6.8 Hz, 1H), 2.63 (dd, *J* = 16.9, 6.9 Hz, 1H), 1.91 – 1.77 (m, 1H), 1.72 – 1.60 (m, 1H), 1.41 – 1.26 (m, 2H), 0.91 (t, *J* = 7.3 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 200.12 , 135.69 , 133.91 , 129.09 , 128.53 , 118.76 , 42.78 , 34.47 , 19.78 , 18.65 , 14.09 .

**Physical State**: colorless oil.

**HRMS** (ESI): calcd for $C_{13}H_{16}NO$ [M+H]$^+$ 202.1232; found 202.1257.

**3-methyl-4-oxo-4-(p-tolyl)butanenitrile (4m)**



**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.84 (d, *J* = 8.3 Hz, 2H), 7.36 – 7.26 (m, 2H), 3.83 – 3.74 (m, 1H), 2.73 (dd, *J* = 16.9, 5.8 Hz, 1H), 2.61 (dd, *J* = 16.9, 8.2 Hz, 1H), 2.43 (s, 3H), 1.39 (d, *J* = 7.2 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 199.63 , 144.95 , 132.31 , 129.78 , 128.74 , 118.84 , 38.08 , 21.85 , 20.55 , 18.34 .

**Physical State**: colorless oil.

**HRMS** (ESI): calcd for $C_{12}H_{14}NO$ [M+H]$^+$ 188.1075; found 188.1080.

## HPLC Calibration Curves

Calibration curves of synthesized reference compounds were created for the determination of yield and TTN. For each substrate, four different concentrations of product (1, 2, 4, and 6 mM) with 2 mM internal standard in 375-μL acetonitrile solutions were mixed each with 375 μL water. The mixtures were vortexed and then analyzed by HPLC based on UV absorbance at 210 nm. All data points represent the average of duplicate runs. The calibration curves depict the ratio of product area to internal standard area (x-axis) against

product concentration in mM (y-axis). Notes: Pdt = product area, IS = internal standard area, [Pdt] = product concentration in reaction, [PC] = protein concentration in reaction, Avg. TTN = average total turnover number, SD TTN = standard deviation of TTN, Avg. Yield = average yield, SD Yield = standard deviation of yield.

**Note:** Data points represent average of duplicate runs; Pdt, product; IS, internal standard; PC, protein concentration; SD, standard deviation; Avg., average.

### (*E*)-1-phenylbut-1-en-1-yl acetate ((*E*)-1a)
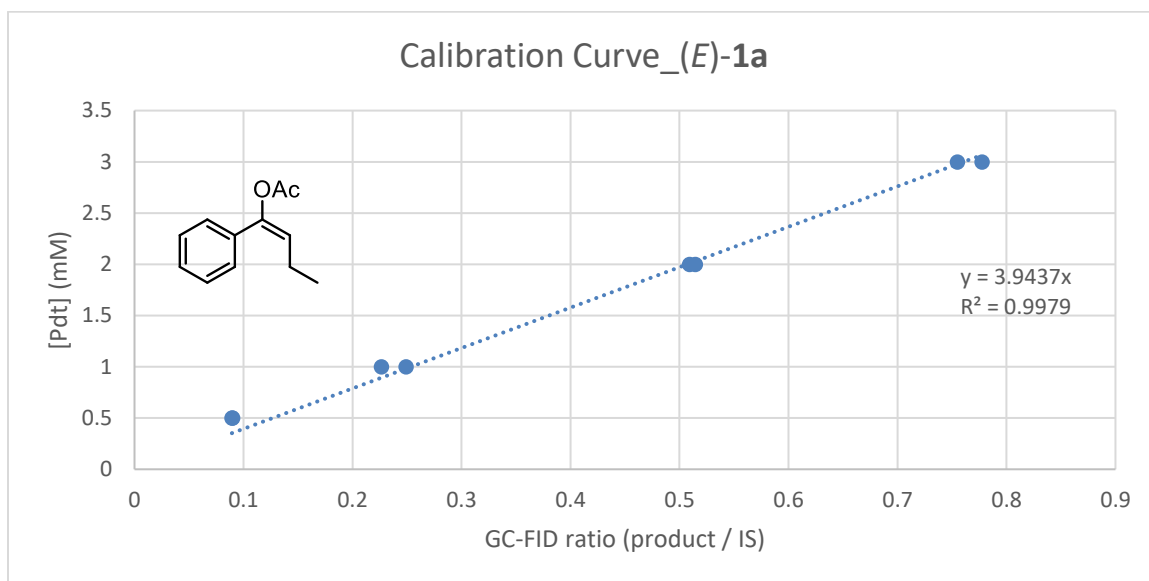


### (*1S,2S,3S*)-2-cyano-3-ethyl-1-phenylcyclopropyl acetate (3a)
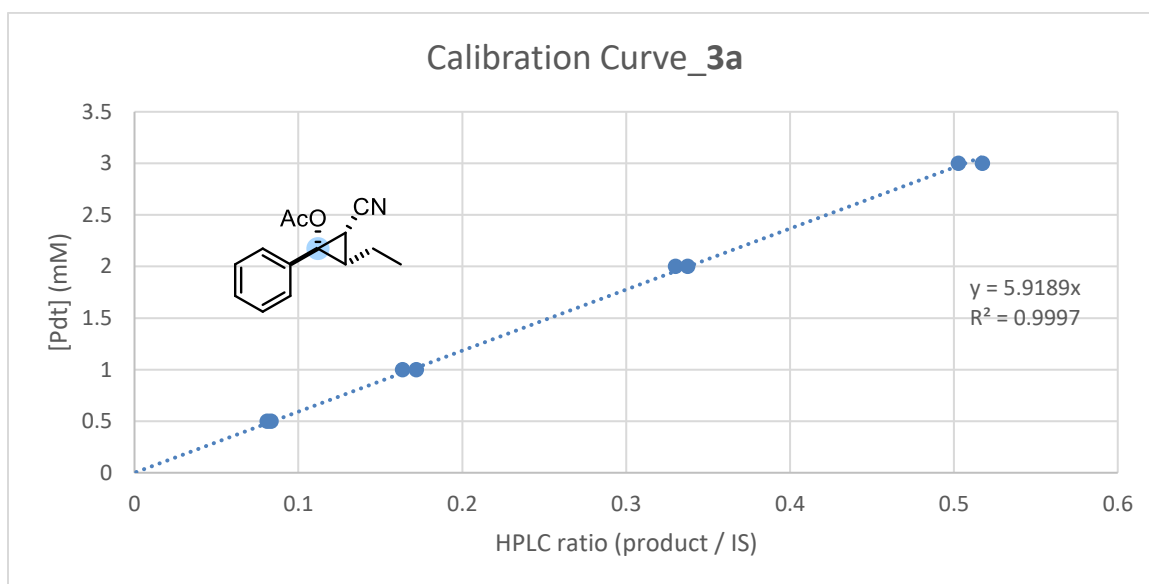
**3-benzoylpentanenitrile (4a)**
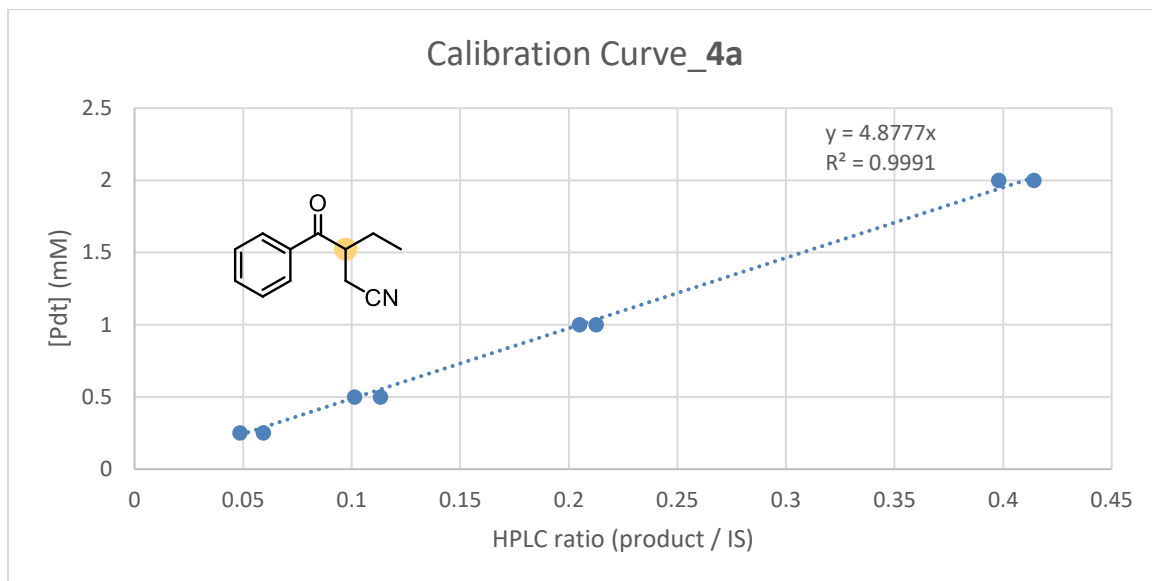


Calibration Curve_**4a**

$y = 4.8777x$
$R^2 = 0.9991$

[Pdt] (mM) vs HPLC ratio (product / IS)

**(*1S,2S,3S*)-2-cyano-3-ethyl-1-(p-tolyl)cyclopropyl acetate (3b)**



Calibration Curve_**3b**

$y = 5.7815x$
$R^2 = 0.9998$

[Pdt] (mM) vs HPLC ratio (product / IS)
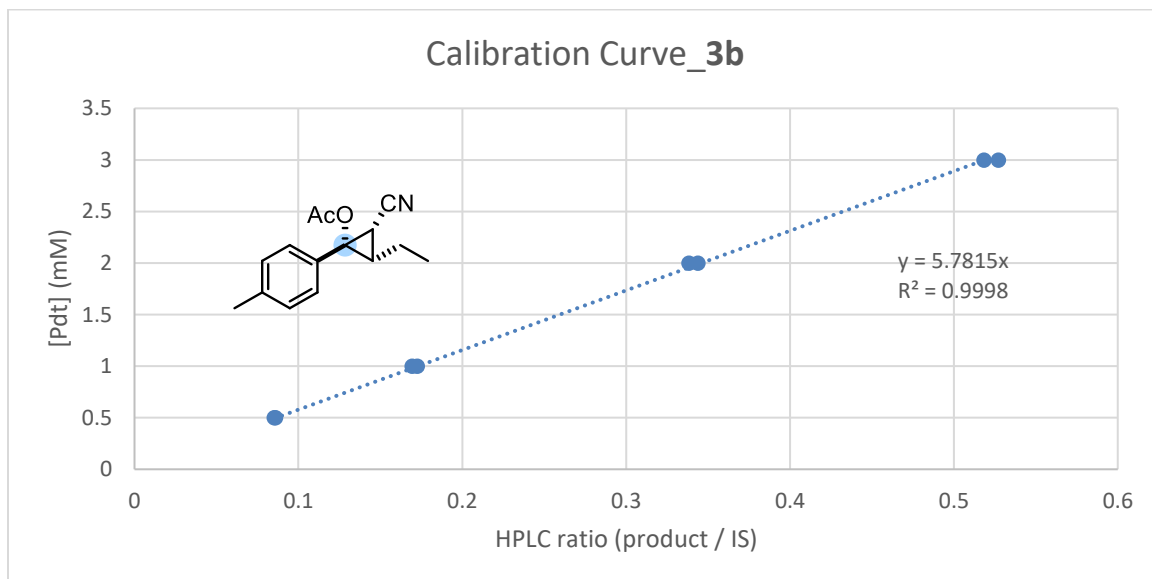
**3-(4-methylbenzoyl)pentanenitrile (4b)**
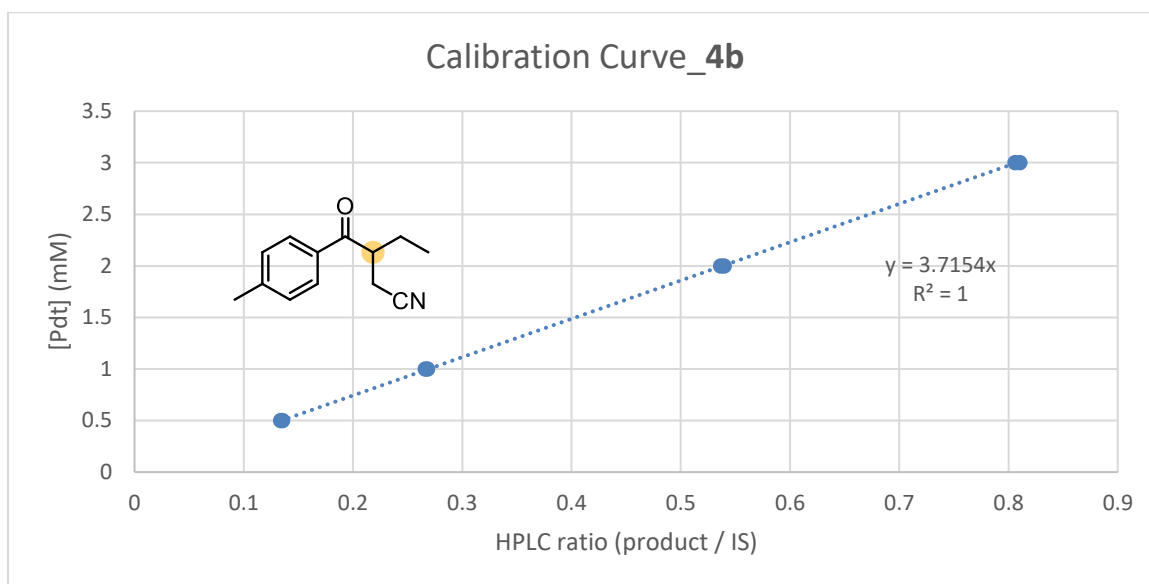


**(*1S,2S,3S*)-2-cyano-3-ethyl-1-(m-tolyl)cyclopropyl acetate (3c)**
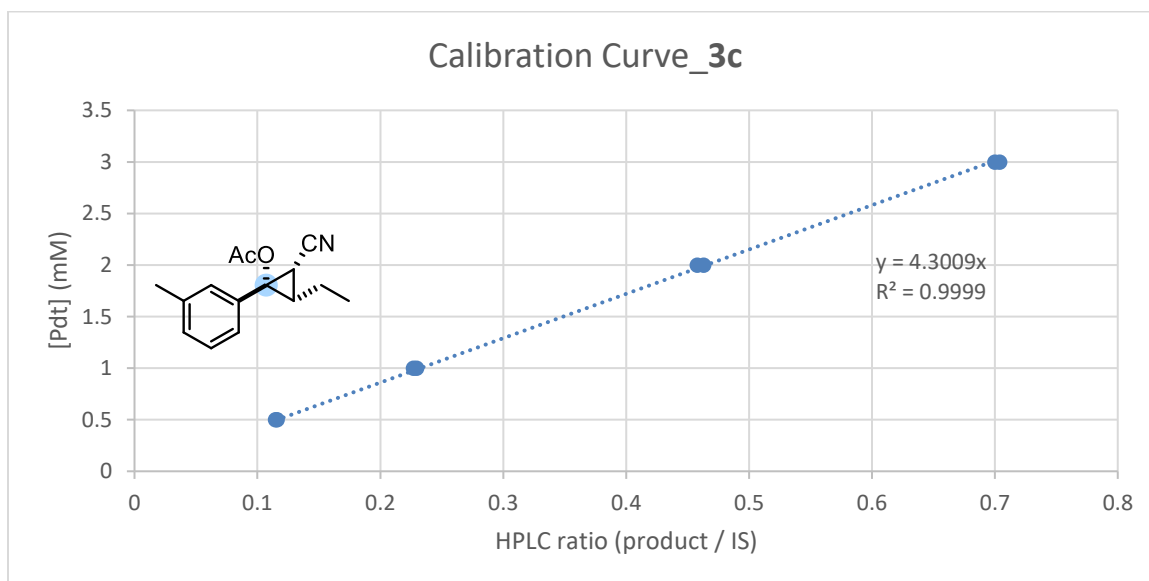
**3-(3-methylbenzoyl)pentanenitrile (4c)**
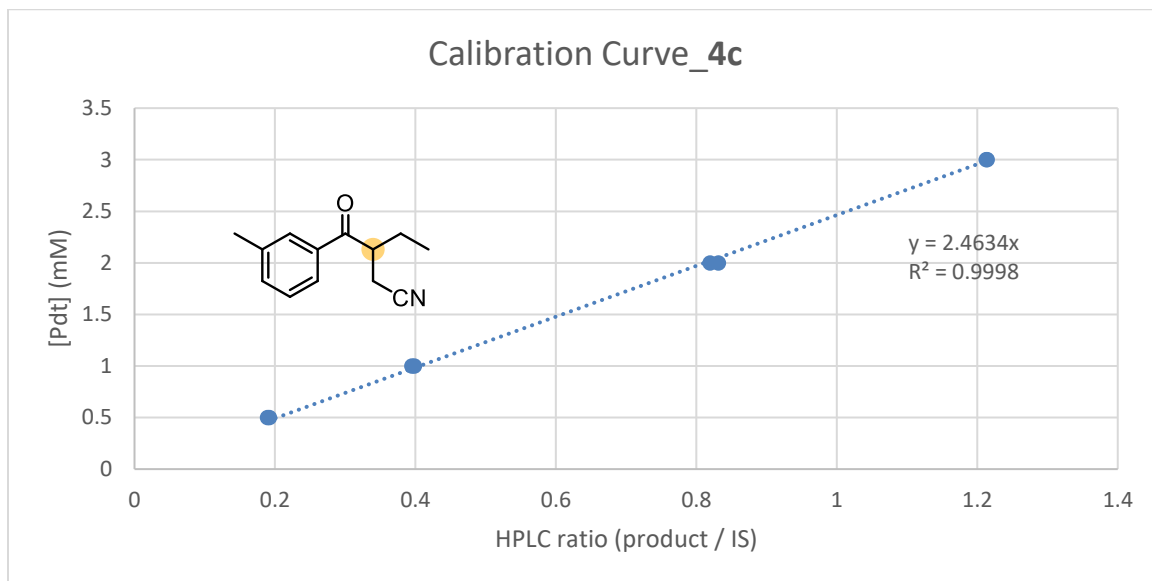


**(*1S,2S,3S*)-2-cyano-3-ethyl-1-(4-fluorophenyl)cyclopropyl acetate (3d)**
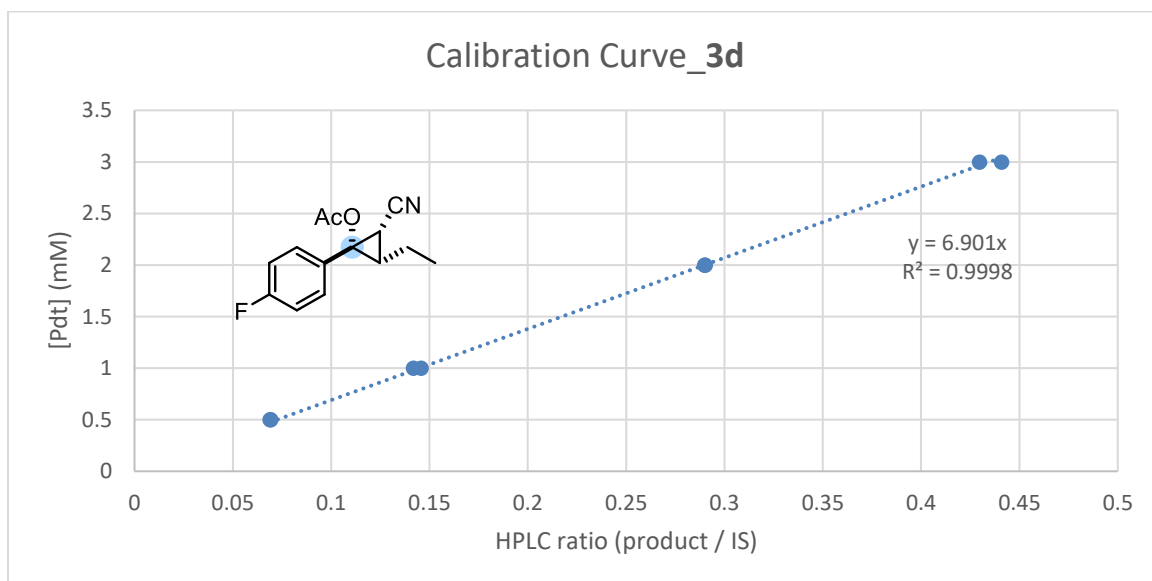
**3-(4-fluorobenzoyl)pentanenitrile (4d)**
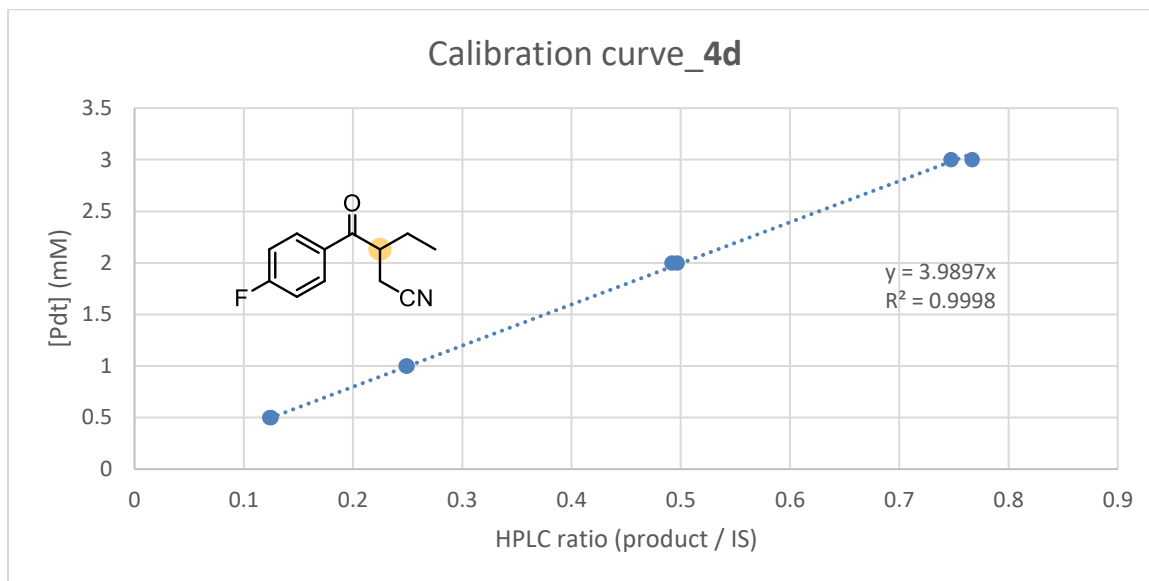


**(1S,2S,3S)-2-cyano-3-ethyl-1-(2-fluorophenyl)cyclopropyl acetate (3e)**
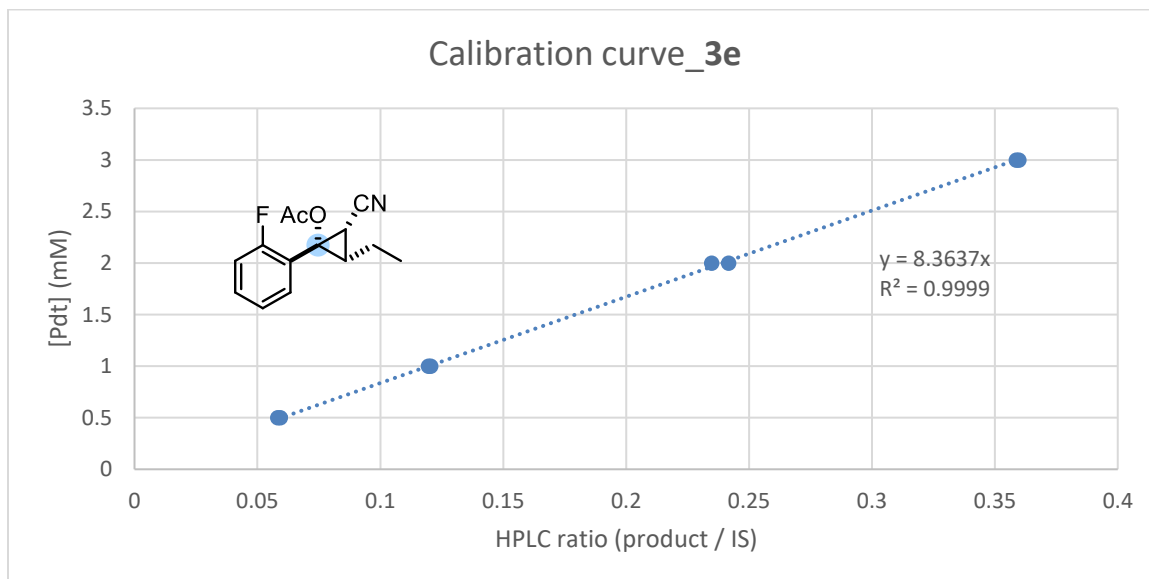
**(*1S,2S,3S*)-2-cyano-3-ethyl-1-(4-chlorophenyl)cyclopropyl acetate (3f)**
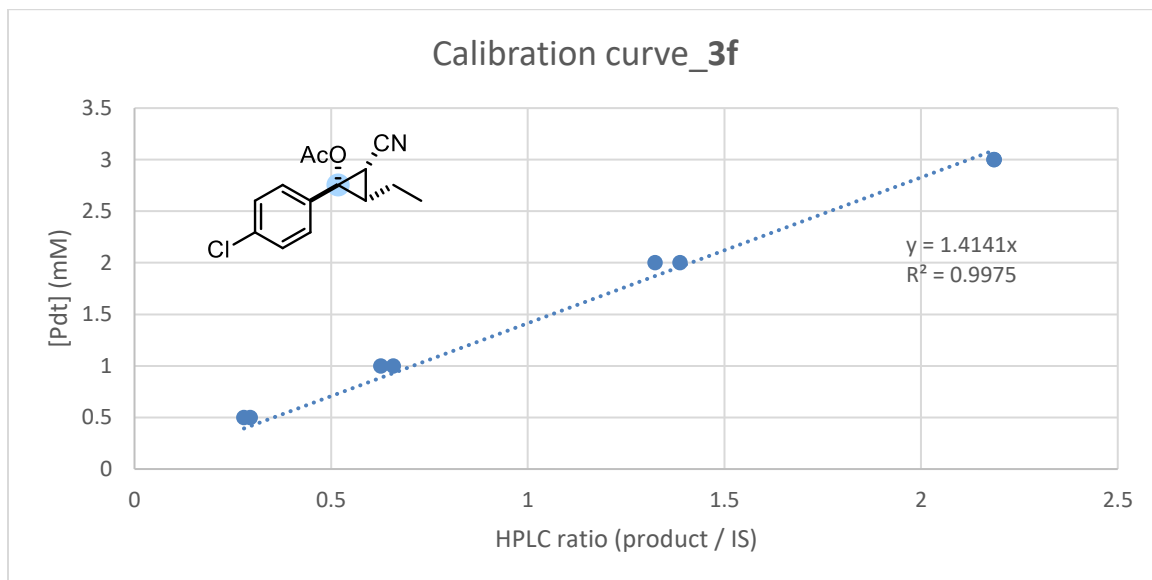


**3-(4-chlorobenzoyl)pentanenitrile (4f)**
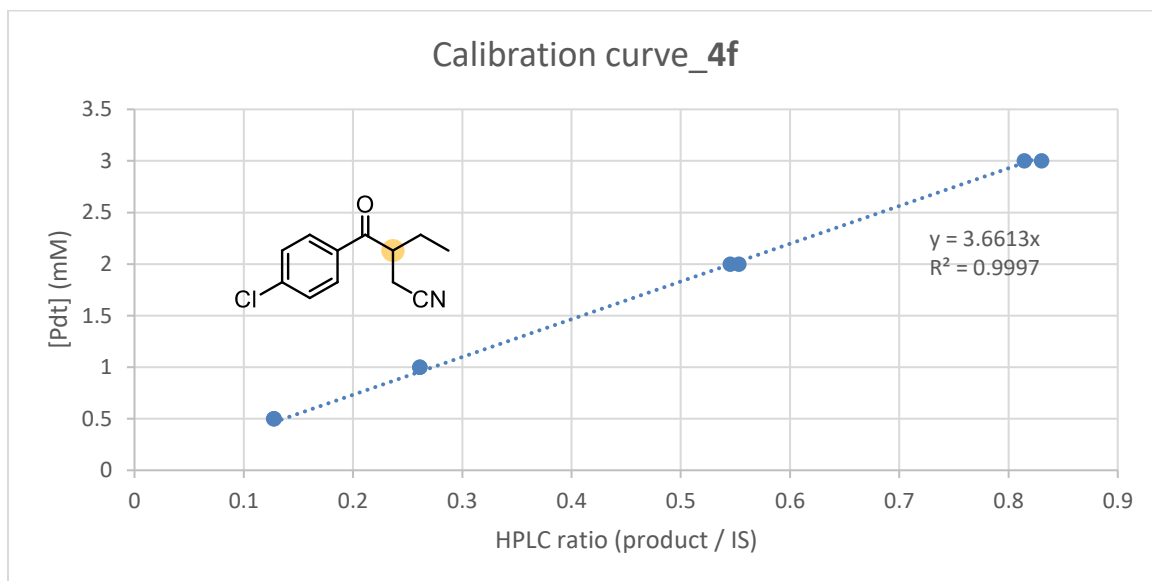
**(*1S,2S,3S*)-2-cyano-3-ethyl-1-(4-bromophenyl)cyclopropyl acetate (3g)**
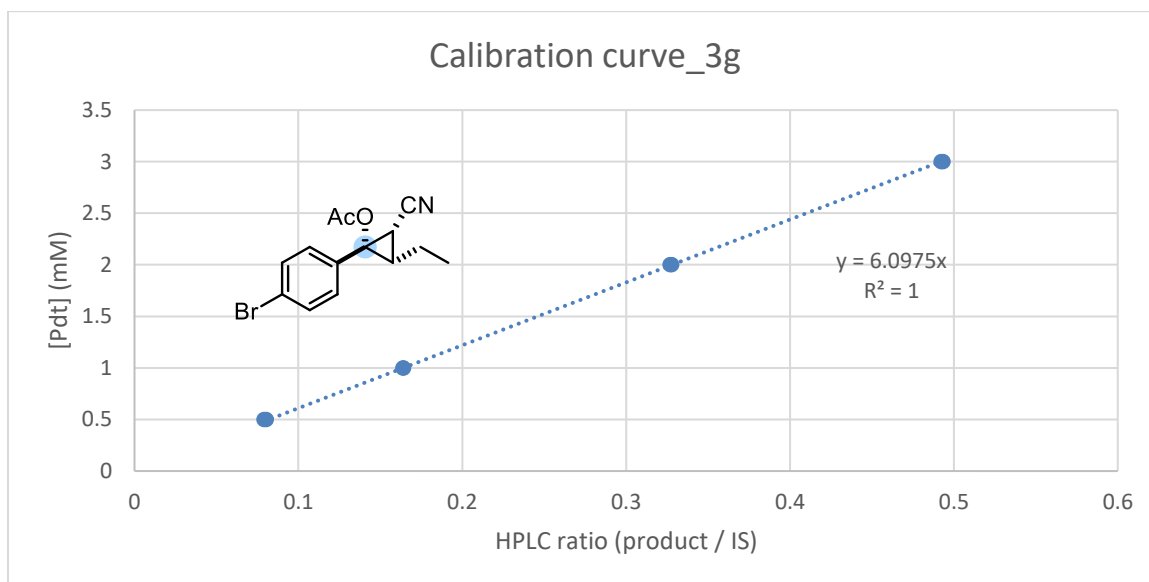


**3-(4-bromobenzoyl)pentanenitrile (4g)**
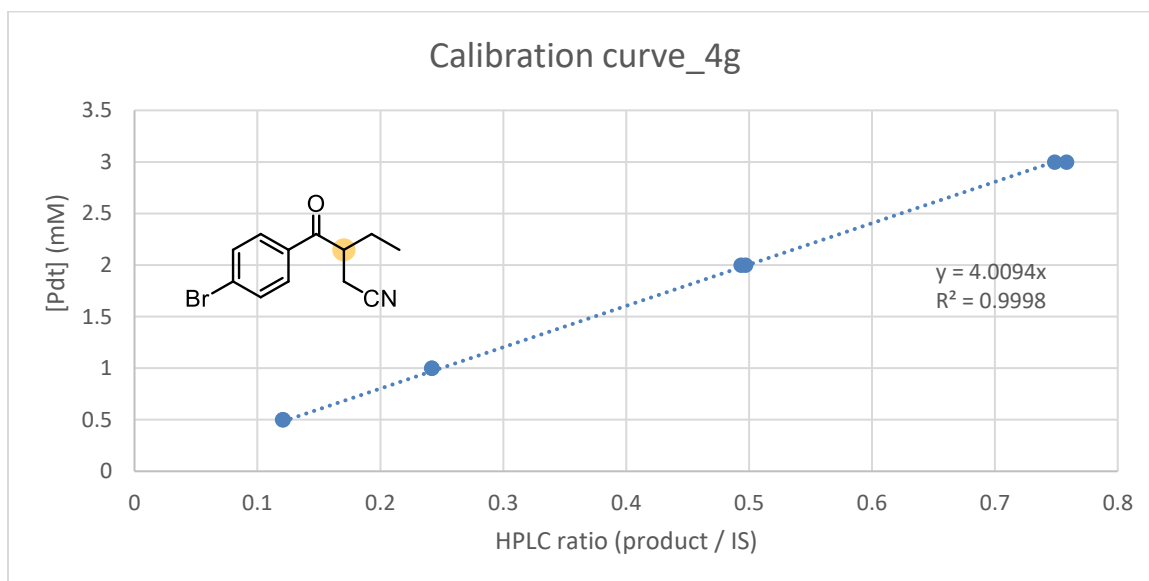
**(*1S,2S,3S*)-2-cyano-3-ethyl-1-(4-methoxyphenyl)cyclopropyl acetate (3h)**
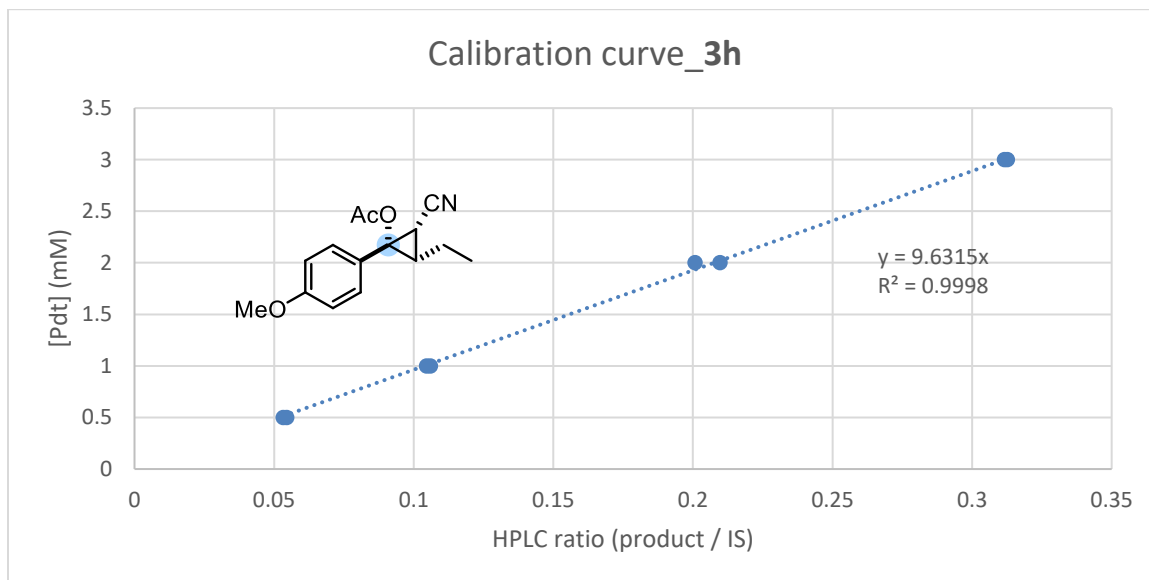


**3-(4-methoxybenzoyl)pentanenitrile (4h)**
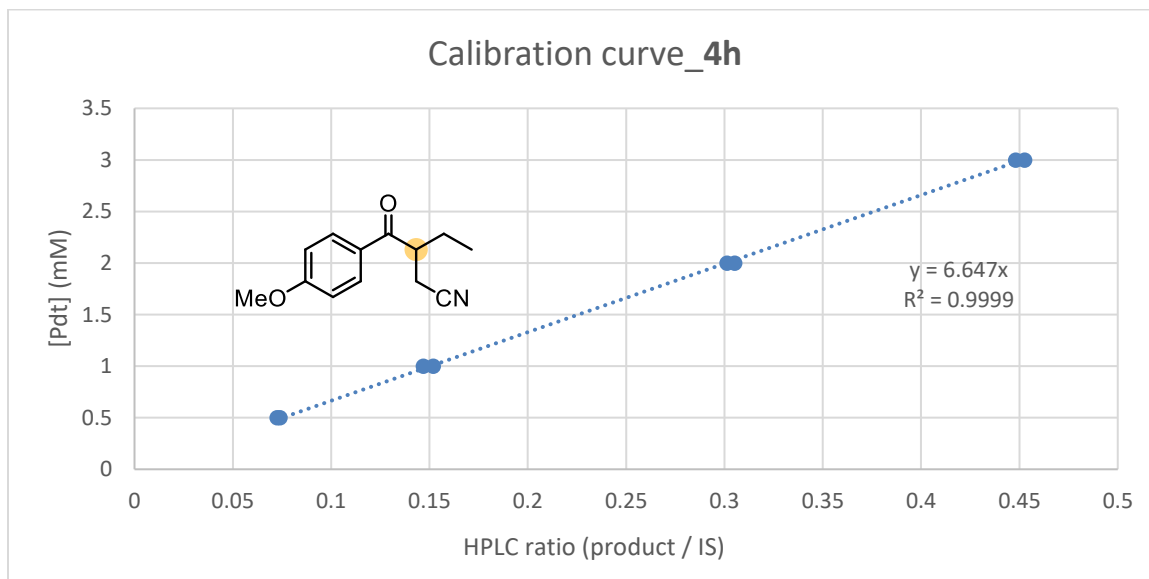
**(*1S,2S,3S*)-2-cyano-3-ethyl-1-(thiophen-2-yl)cyclopropyl acetate (3i)**
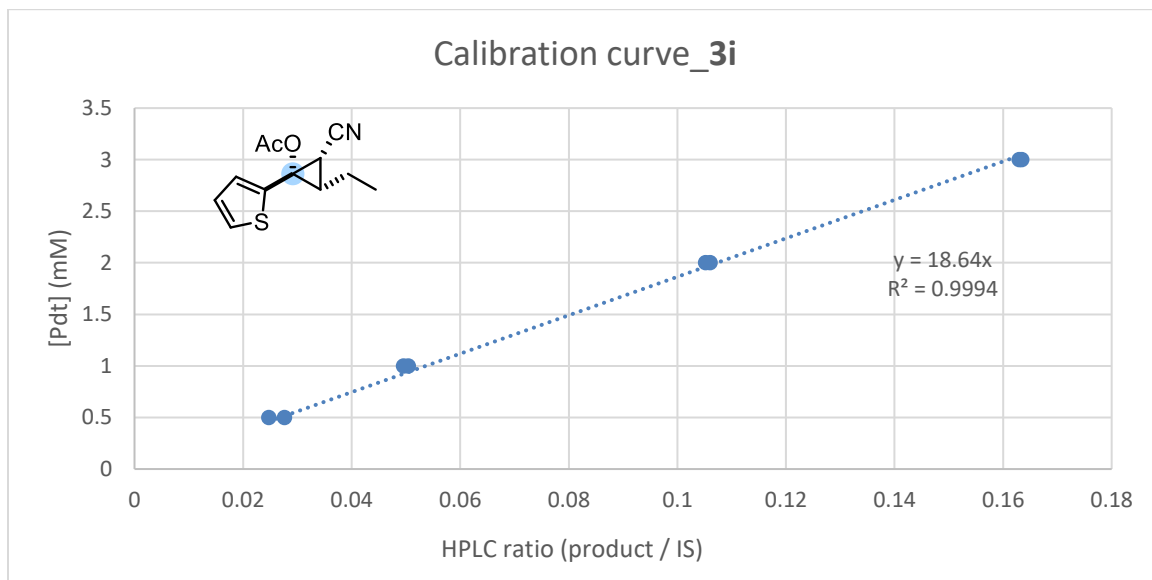


**3-(thiophene-2-carbonyl)pentanenitrile (4i)**
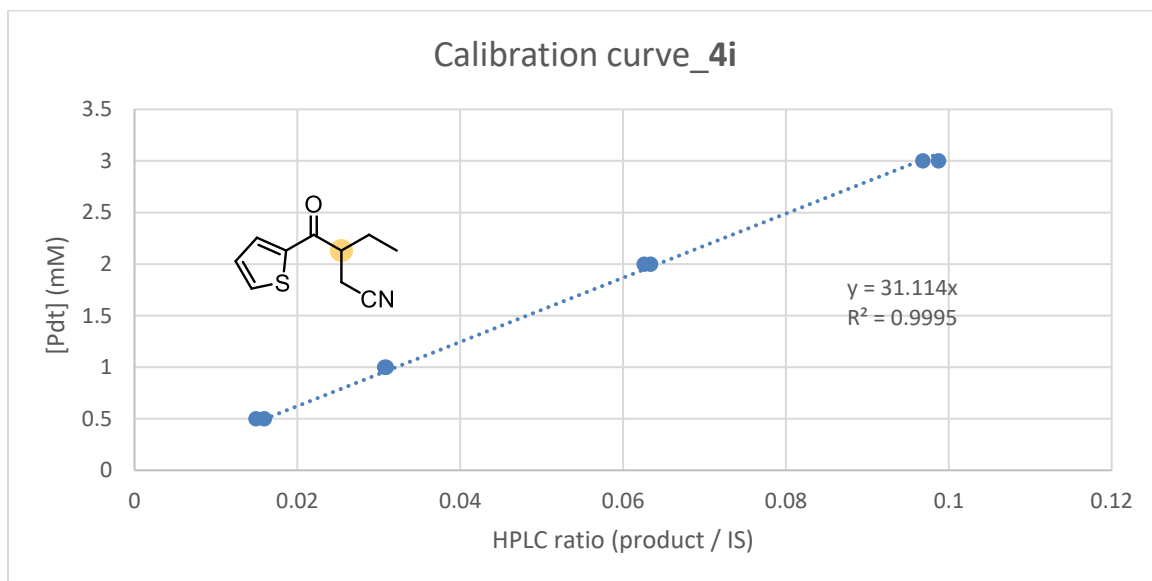
**(*1S,2S,3S*)-2-cyano-3-ethyl-1-(4-hydroxyphenyl)cyclopropyl acetate (3j)**
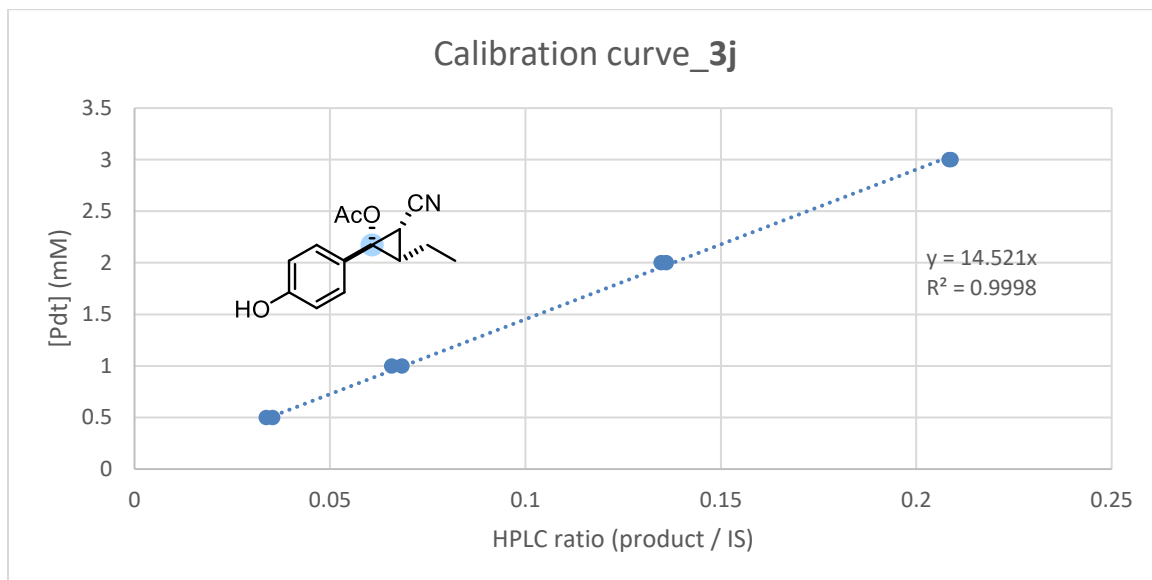


**3-(4-hydroxybenzoyl)pentanenitrile (4j)**
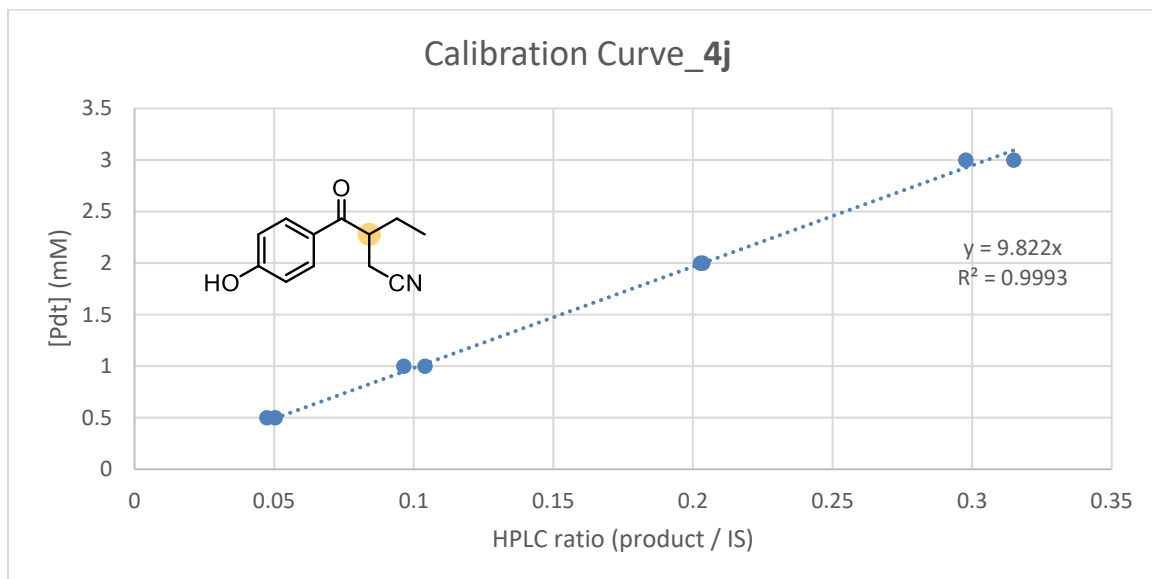
**(*1S,2S,3S*)-2-cyano-3-methyl-1-phenylcyclopropyl acetate (3k)**
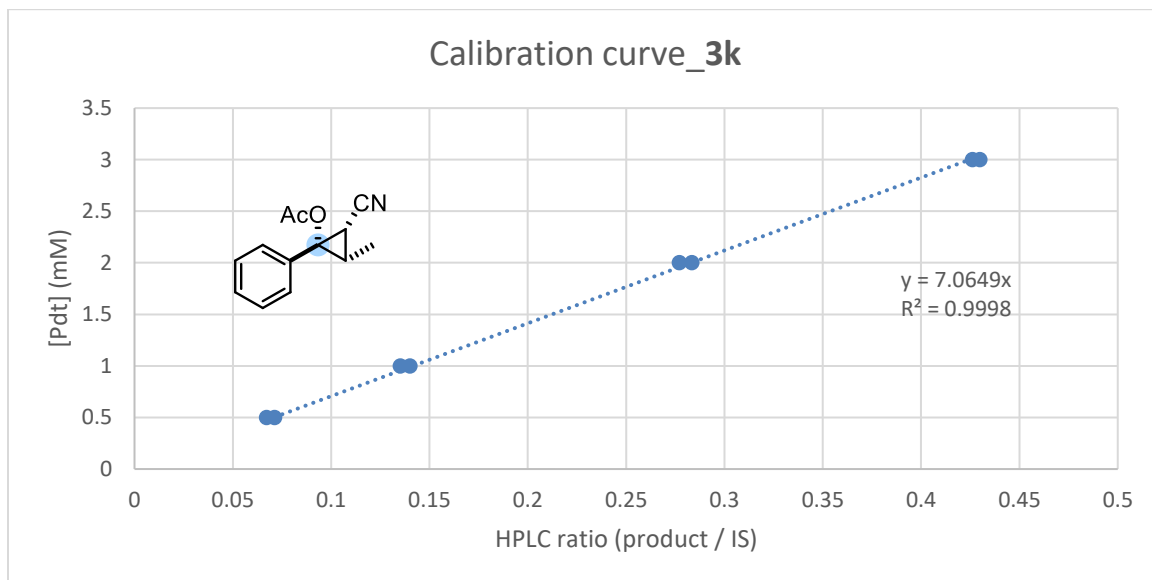


**3-methyl-4-oxo-4-phenylbutanenitrile (4k)**

**(*1S,2S,3S*)-2-cyano-1-phenyl-3-propylcyclopropyl acetate (3l)**



**3-benzoylhexanenitrile (4l)**
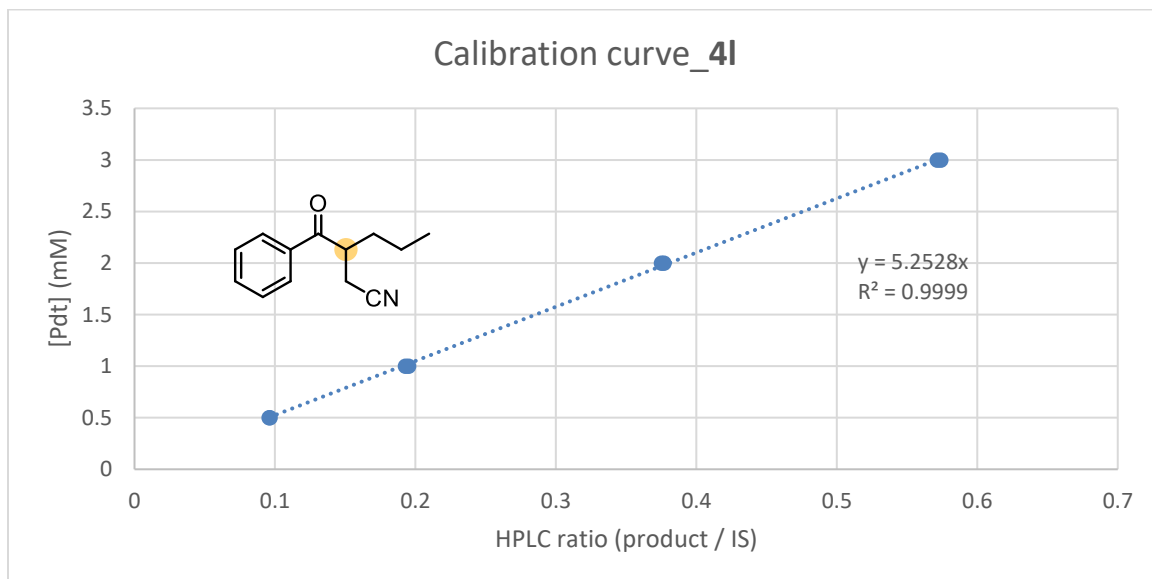
**(*1S,2S,3S*)-2-cyano-3-methyl-1-(p-tolyl)cyclopropyl acetate (3m)**
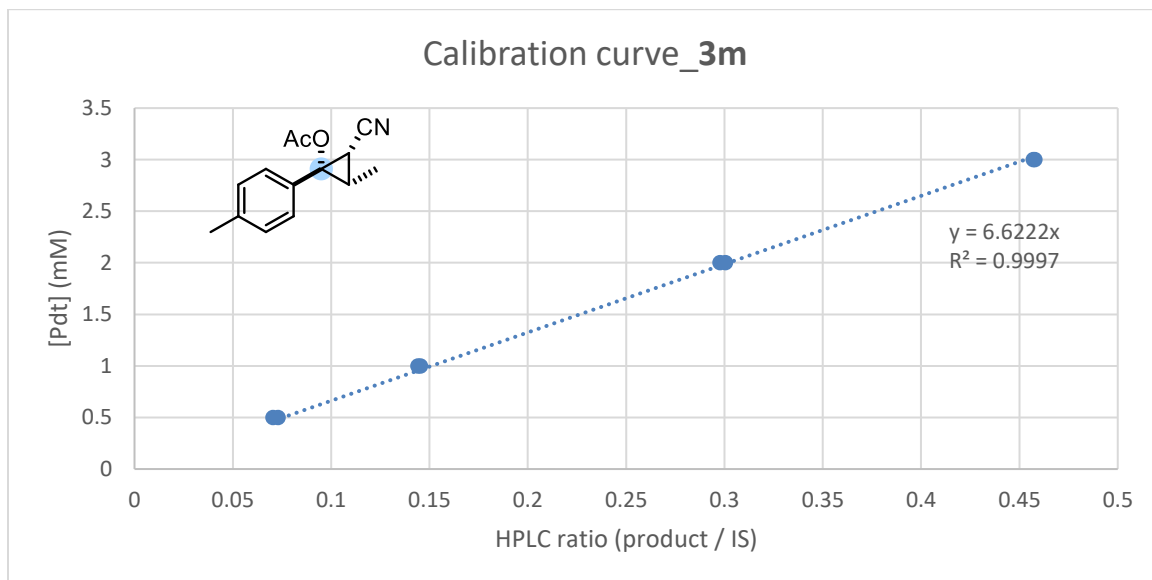


**3-methyl-4-oxo-4-(p-tolyl)butanenitrile (4m)**
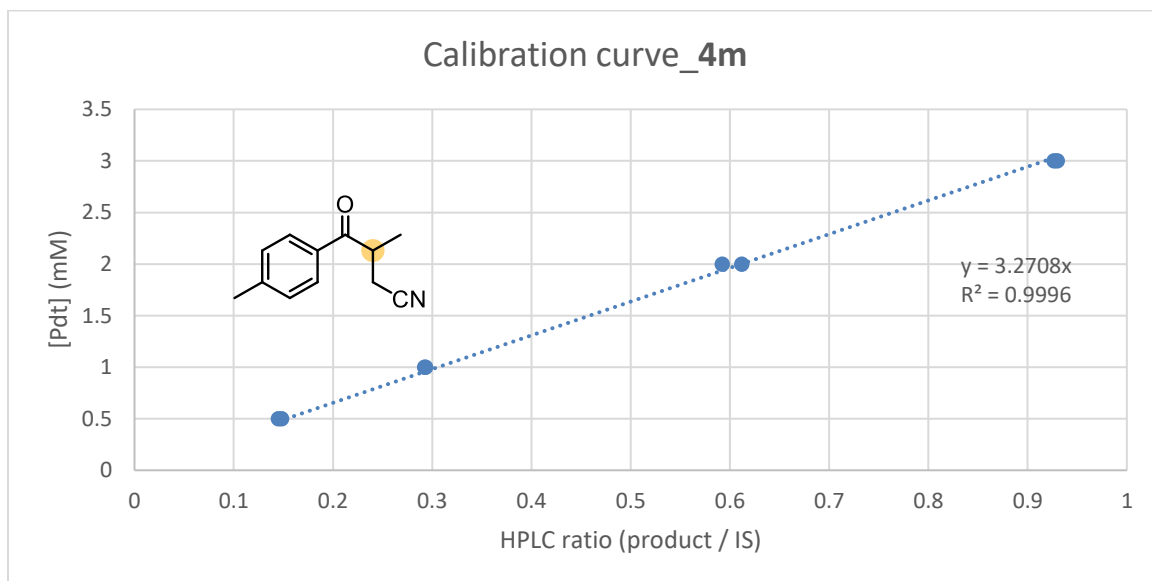
**(*1S,2S,3S*)-2-cyano-3-ethyl-1-phenylcyclopropyl pivalate (3n)**
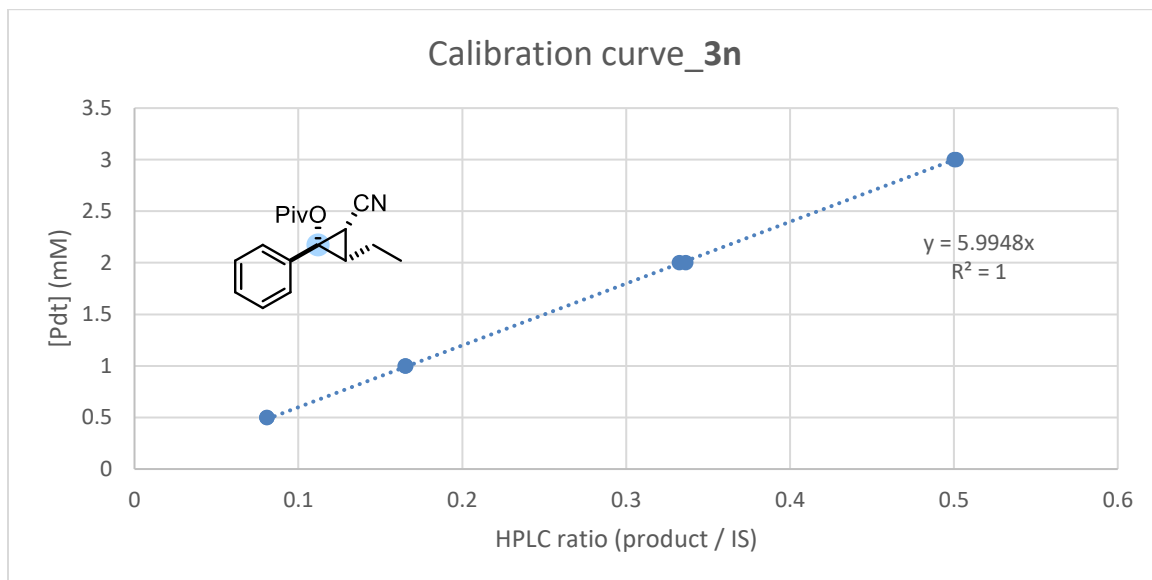


## Enantioselectivity Determinations

**Stereopurity determination of 1a before and after** P411-INC-5185**-catalyzed cyclopropanation of a *Z/E* mixture of 1a.**

The stereochemistry of mixed **1a** was determined by comparing the elution order of each two diastereomers with pure (*Z*)-**1a** and (*E*)-**1a** that obtained from Chiral chromatographic separations (see 0). GC-FID was used to determine the *Z/E* ratios as below.

**(*Z*)-1-phenylbut-1-en-1-yl acetate ((*Z*)-1a)**
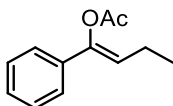


**GC-FID Column:** HP-5ms Ultra Inert Column (Part Number:19091S-413UIJ&W; HP-5ms Ultra Inert GC Column, 30 m, 0.32 mm, 0.25 µm, 7 inch cage);

**GC-FID method:** Hold at 100 ºC for 2 min, ramp to 300 ºC at 50 ºC / min, and hold at 300 ºC for 2 min.

**(*E*)-1-phenylbut-1-en-1-yl acetate ((*E*)-1a)**
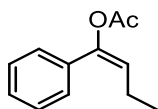


**GC-FID Column:** HP-5ms Ultra Inert Column (Part Number:19091S-413UIJ&W; HP-5ms Ultra Inert GC Column, 30 m, 0.32 mm, 0.25 µm, 7 inch cage);

**GC-FID method:** Hold at 100 °C for 2 min, ramp to 300 °C at 50 °C / min, and hold at 300 °C for 2 min.
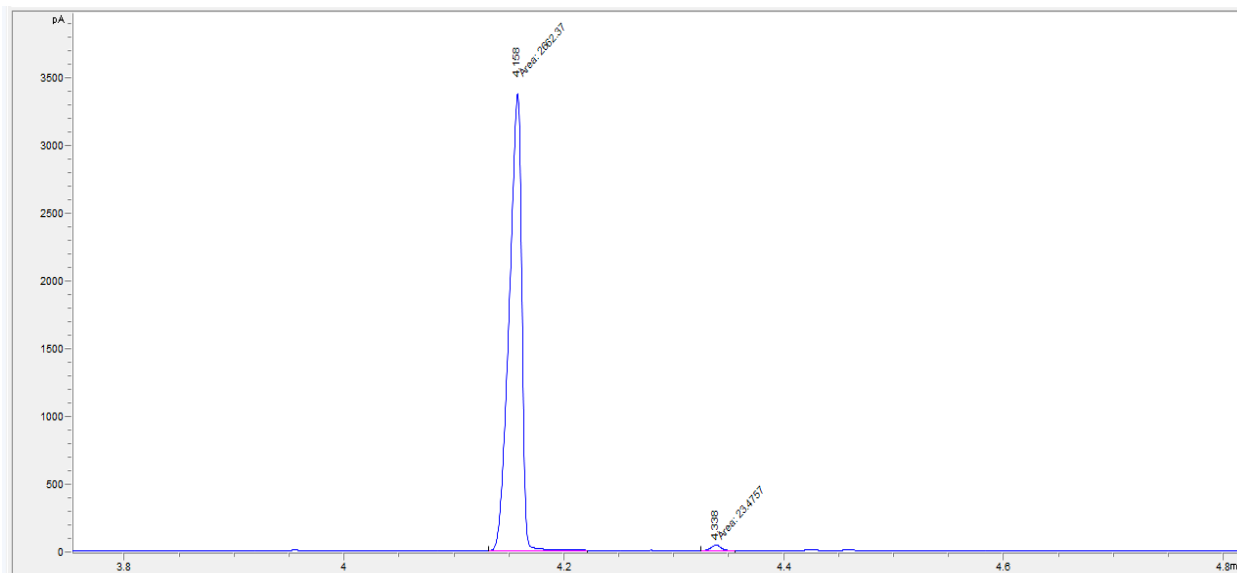
**(Z)- and (E)-1-phenylbut-1-en-1-yl acetate ((Z)- and (E)-1a)**



(Z)-**1a**          and          (E)-**1a**

**GC-FID Column:** HP-5ms Ultra Inert Column (Part Number:19091S-413UIJ&W; HP-5ms Ultra Inert GC Column, 30 m, 0.32 mm, 0.25 µm, 7 inch cage);

**GC-FID method:** Hold at 100 ºC for 2 min, ramp to 300 ºC at 50 ºC / min, and hold at 300 ºC for 2 min.

*Z/E* ratio of **1a** before P411-INC-5185-catalyzed cyclopropanation: *Z* (right peak) */E* (left peak) = 1:1
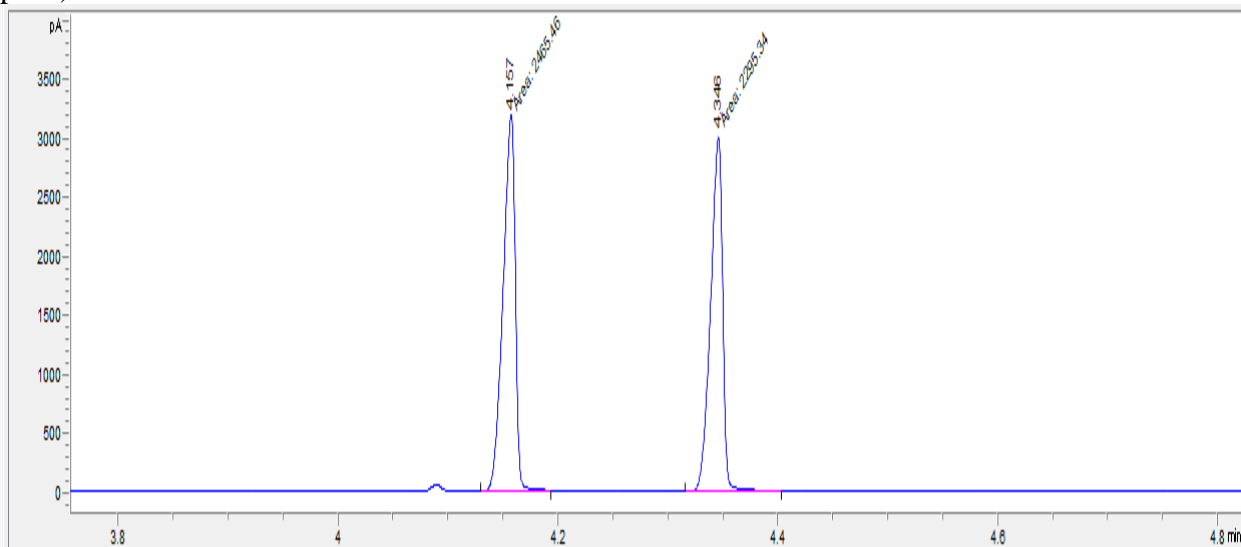


*Z/E* ratio of **1a** after P411-INC-5185-catalyzed cyclopropanation: *Z* (right peak) */E* (left peak) = 2:98
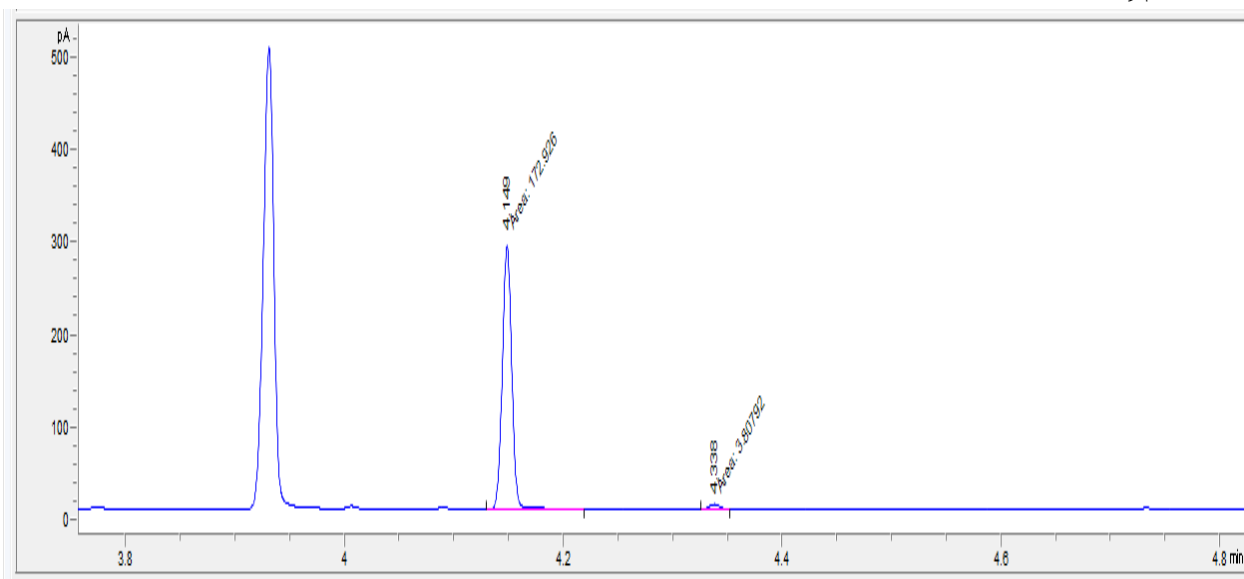
## General procedure F: Enantioselectivity determinations of cyclopropanated products 3
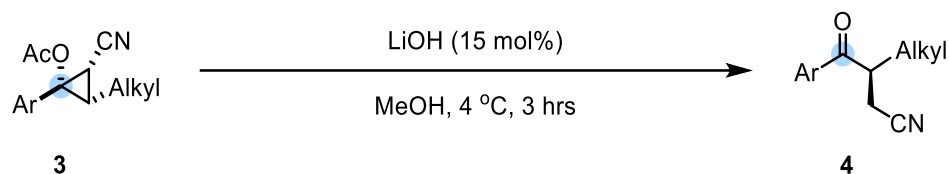


A solution of **3** (100 µL, 13.32 mM in methanol) was prepared and transferred to a 2-mL screw-cap vial. After cooling down to 4 ℃, a solution of lithium hydroxide (0.4 µL, LiOH, 0.5 M in $H_2O$, 15 mol%) was added, and the resulting solution was shaken at 600 rpm and 4 ℃ for 3 hours unless otherwise noted. After that, the mixture was quenched with a mixture of hexane and $H_2O$ (1:1, 1000 µL in total) in 1.7-mL Eppendorf tube before vortexing (20 s × 2) and centrifugation (14,000 × g, 5 min, 4 °C). The organic phase was then carefully transferred to a 2-mL screw-cap vial, concentrated and subjected to normal-phase HPLC to determine the e.e. of **4**.

To determine the enantiomeric excess (e.e.), cyclopropanated products **3** were converted into α-branched ketone products **4**. The absolute stereochemistry for enzymatic products **3j** and **3k** was assigned as *S, S, S* through x-ray crystallography (see **Section 12**). The other polysubstituted cyclopropyl acetates **3** were assigned by analogy.

**Note:**

1. In most cases, the racemization of **4** occurs when extending the reaction time to more than 4 h under standard reaction conditions.

2. For some samples, we found the retention time drifts in normal-phase HPLC.

**(*S*)-3-benzoylpentanenitrile (4a; obtained from 3a)**
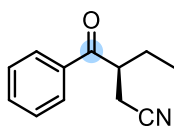


**Chiral-phase HPLC conditions:** Chiralpak IB, 15% *i*-PrOH in hexane, 1.0 mL/min, 25 ℃, 254 nm

**Racemic 4a:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 7.759 | 971.2 | 108.9 | 0.1486 | 49.921 | 0.851 |
| 2 | 9.565 | 974.3 | 84.9 | 0.1912 | 50.079 | 0.821 |

**Enzymatic synthesis with** P411-INC-5185**:** 95% e.e.



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 7.756 | 21.6 | 2.5 | 0.1455 | 2.414 | 0.748 |
| 2 | 9.593 | 874.9 | 77.5 | 0.188 | 97.586 | 0.861 |

**Enzymatic synthesis with** P411-INC-5186**:** 98% e.e.



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 7.666 | 5.3 | 9.6E-1 | 0.0924 | 0.986 | 0.443 |
| 2 | 9.331 | 531.6 | 50.7 | 0.1746 | 99.014 | 0.852 |
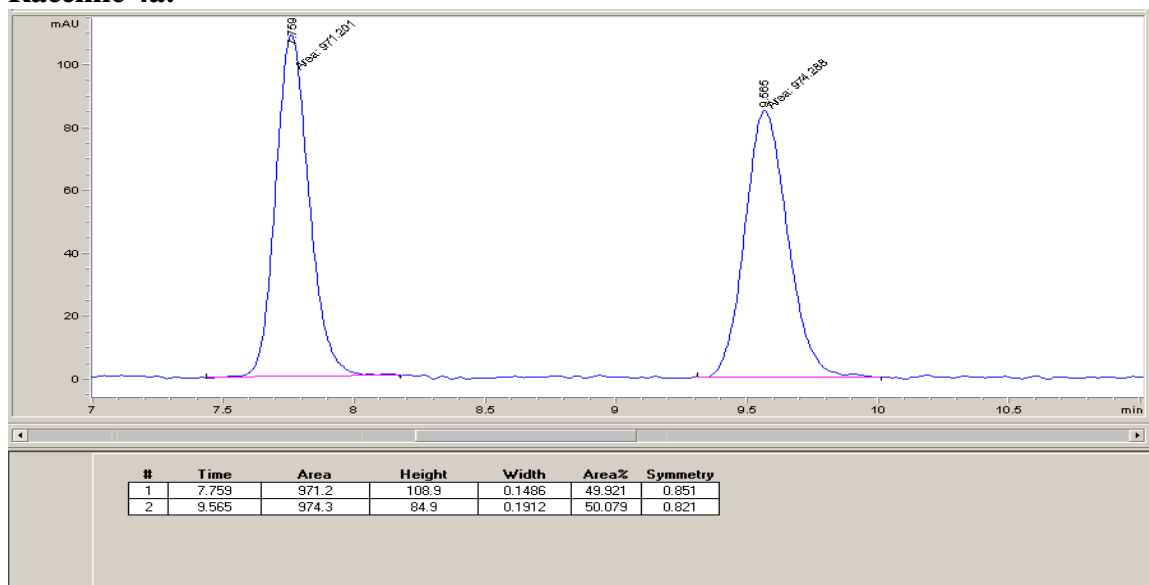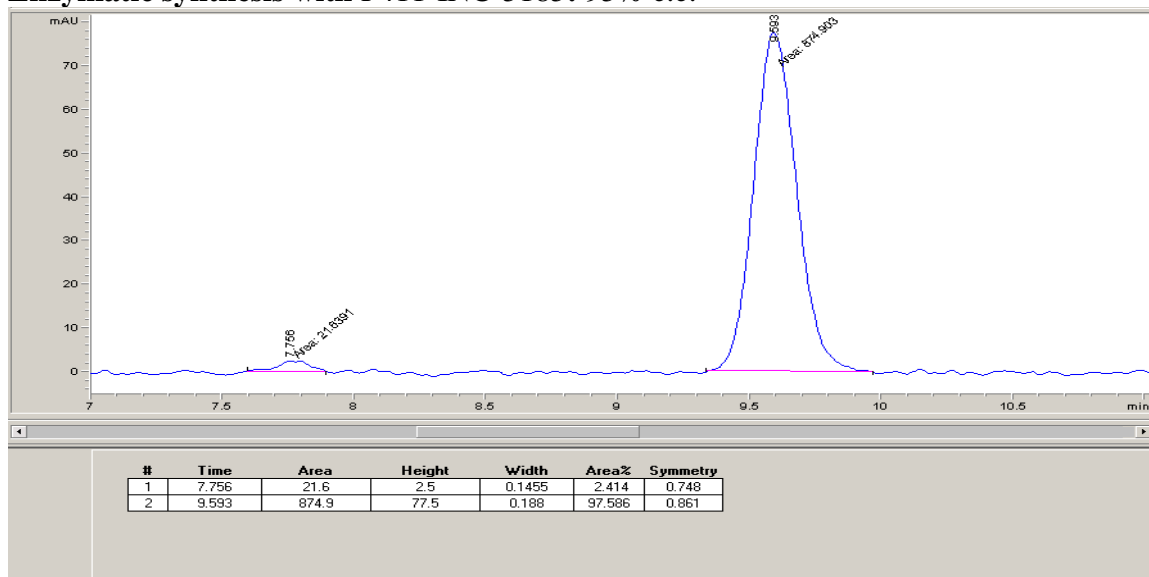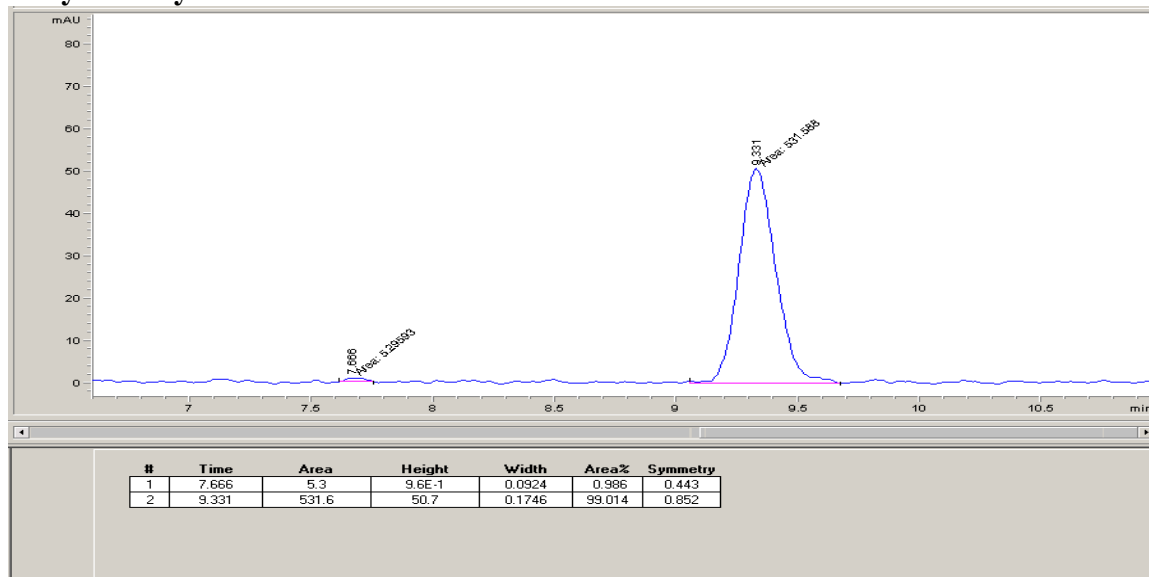
**(*S*)-3-(4-methylbenzoyl)pentanenitrile (4b; obtained from 3b)**



**Chiral-phase HPLC conditions:** Chiralpak IA, 5% *i*-PrOH in hexane, 1.0 mL/min, 25 °C, 245 nm

**Racemic 4b:**
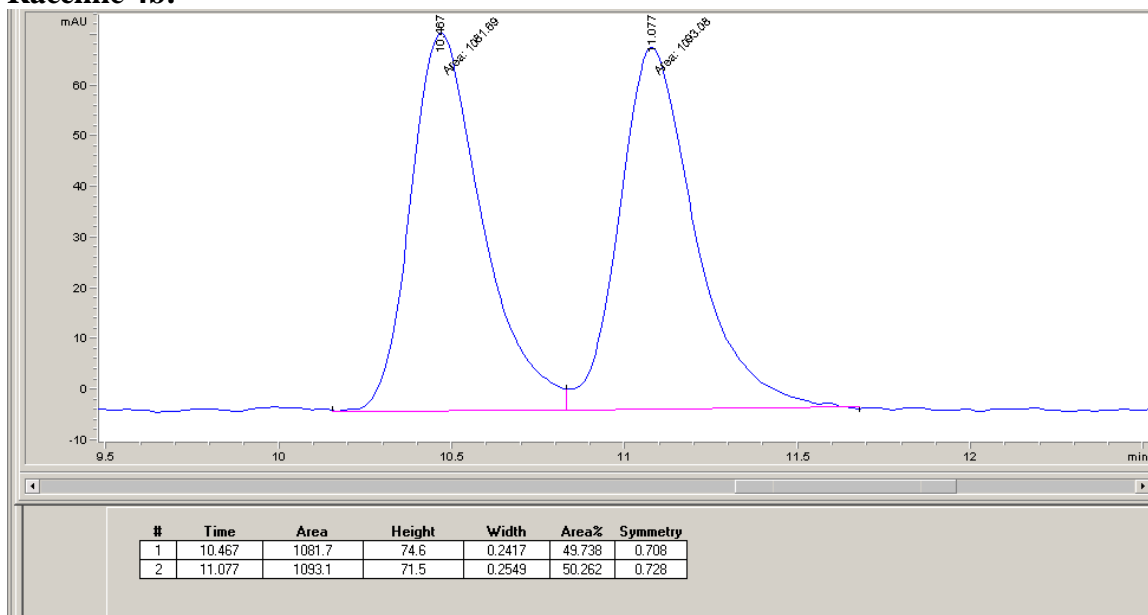


| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 10.467 | 1081.7 | 74.6 | 0.2417 | 49.738 | 0.708 |
| 2 | 11.077 | 1093.1 | 71.5 | 0.2549 | 50.262 | 0.728 |

**Enzymatic synthesis with** P411-INC-5185**:** 94% e.e.



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 10.394 | 6.5 | 8.1E-1 | 0.1334 | 3.196 | 1.452 |
| 2 | 11.015 | 196.3 | 13.1 | 0.2497 | 96.804 | 0.73 |

**Enzymatic synthesis with** P411-INC-5186**:** 95% e.e.



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 10.261 | 9.4 | 9.7E-1 | 0.1615 | 2.316 | 0.895 |
| 2 | 10.857 | 397.7 | 26 | 0.2547 | 97.684 | 0.685 |
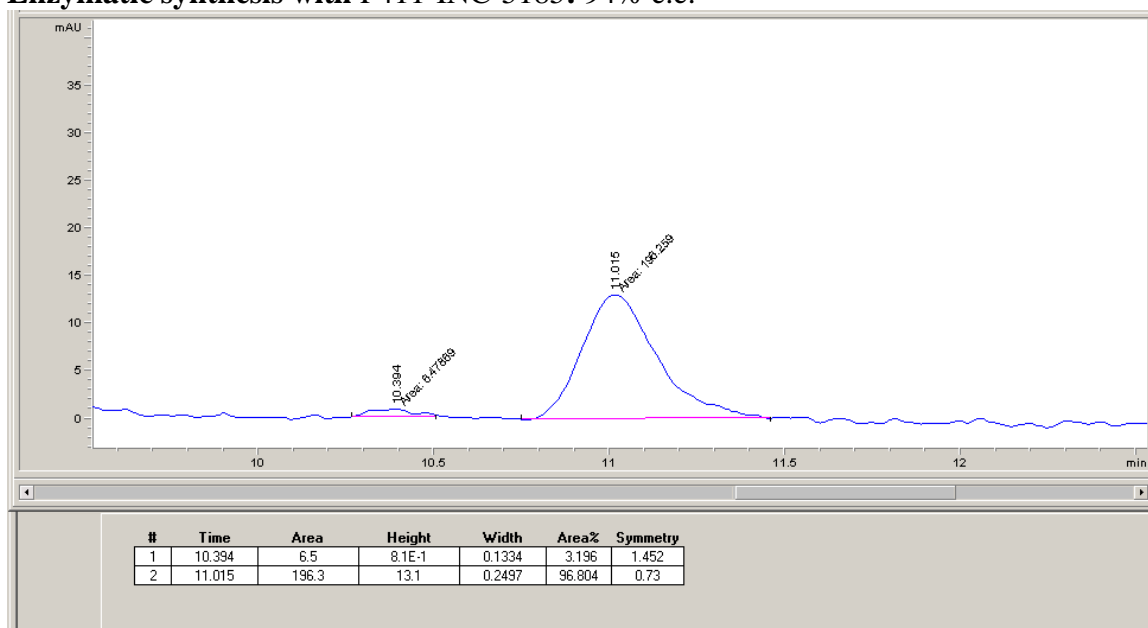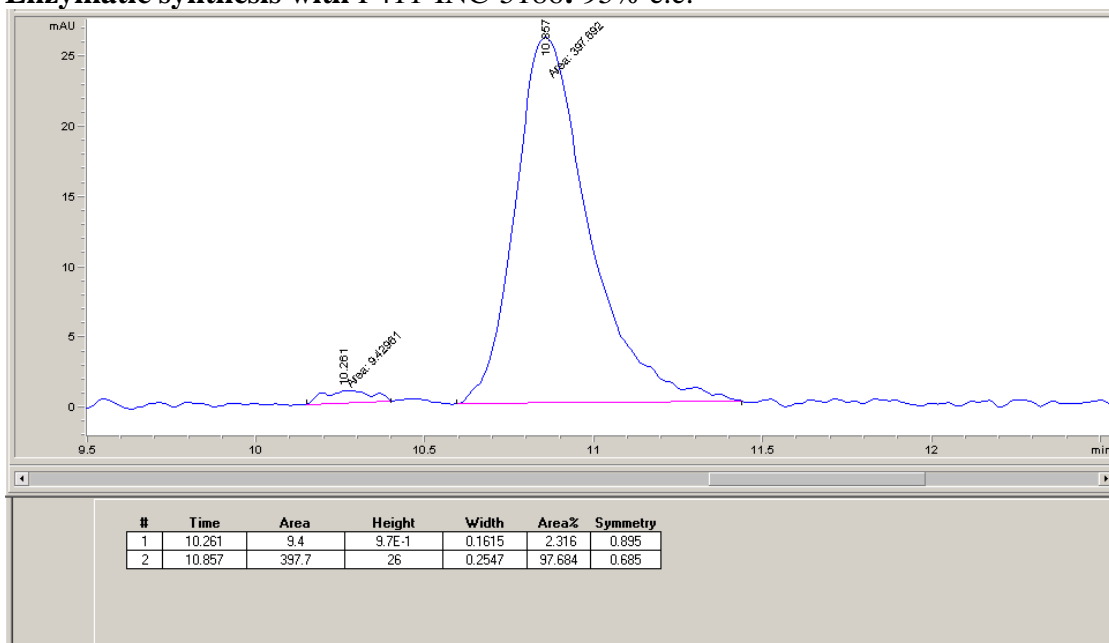
**(*S*)-3-(3-methylbenzoyl)pentanenitrile (4c; obtained from 3c)**



**Chiral-phase HPLC conditions:** Chiralpak IB, 15% *i*-PrOH in hexane, 1.0 mL/min, 25 ℃, 245 nm

**Racemic 4c:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 6.757 | 5229 | 675 | 0.1291 | 49.807 | 0.804 |
| 2 | 7.962 | 5269.6 | 588.2 | 0.1493 | 50.193 | 0.81 |

**Enzymatic synthesis with** P411-INC-5185**:** >99% e.e.



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 8.128 | 353.2 | 37.8 | 0.1557 | 100.000 | 0.879 |

**Enzymatic synthesis with** P411-INC-5186**:** 95% e.e.



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 6.738 | 24.6 | 3.4 | 0.1212 | 2.384 | 0.86 |
| 2 | 7.908 | 1007.6 | 110.6 | 0.1519 | 97.616 | 0.863 |

**(*S*)-3-(4-fluorobenzoyl)pentanenitrile (4d; obtained from 3d)**



**Chiral-phase HPLC conditions:** Chiralpak IA, 5% *i*-PrOH in hexane, 1.0 mL/min, 25 ℃, 245 nm

**Racemic 4d:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 11.238 | 1336.4 | 86.9 | 0.2563 | 49.999 | 0.708 |
| 2 | 11.944 | 1336.4 | 82.2 | 0.2709 | 50.001 | 0.726 |

**Enzymatic synthesis with** P411-INC-5185**:** 93% e.e.



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 11.301 | 6.6 | 6.9E-1 | 0.1588 | 3.438 | 0.793 |
| 2 | 11.969 | 184.4 | 10.7 | 0.2885 | 96.562 | 0.896 |

**Enzymatic synthesis with** P411-INC-5186**:** 96% e.e.



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 11.178 | 115.8 | 8.4 | 0.2302 | 1.994 | 1.187 |
| 2 | 11.834 | 5692.5 | 259.5 | 0.3656 | 98.006 | 0.697 |

**(*S*)-3-(2-fluorobenzoyl)pentanenitrile (4e; obtained from 3e)**
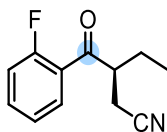


**Chiral-phase HPLC conditions:** Chiralpak IA, 5% *i*-PrOH in hexane, 1.0 mL/min, 25 ºC, 245 nm

**Racemic 4e:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 8.359 | 282.5 | 25 | 0.188 | 50.421 | 0.715 |
| 2 | 9.153 | 277.8 | 22.8 | 0.2033 | 49.579 | 0.75 |

**Enzymatic synthesis with** P411-INC-5185**: >99% e.e.**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 8.414 | 15.4 | 1.7 | 0.1513 | 0.494 | 0.525 |
| 2 | 9.134 | 3110.2 | 224 | 0.2314 | 99.506 | 0.663 |

**(*S*)-3-(4-chlorobenzoyl)pentanenitrile (4f; obtained from 3f)**



**Chiral-phase HPLC conditions:** Chiralpak IA, 5% *i*-PrOH in hexane, 1.0 mL/min, 25 ℃, 254 nm

**Racemic 4f:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 11.938 | 4993.4 | 304.4 | 0.2734 | 49.163 | 0.672 |
| 2 | 12.626 | 5163.3 | 291.4 | 0.2953 | 50.837 | 0.688 |

**Enzymatic synthesis with** P411-INC-5186**:** >99% e.e.



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 12.68 | 723.7 | 43.2 | 0.2554 | 100.000 | 0.713 |
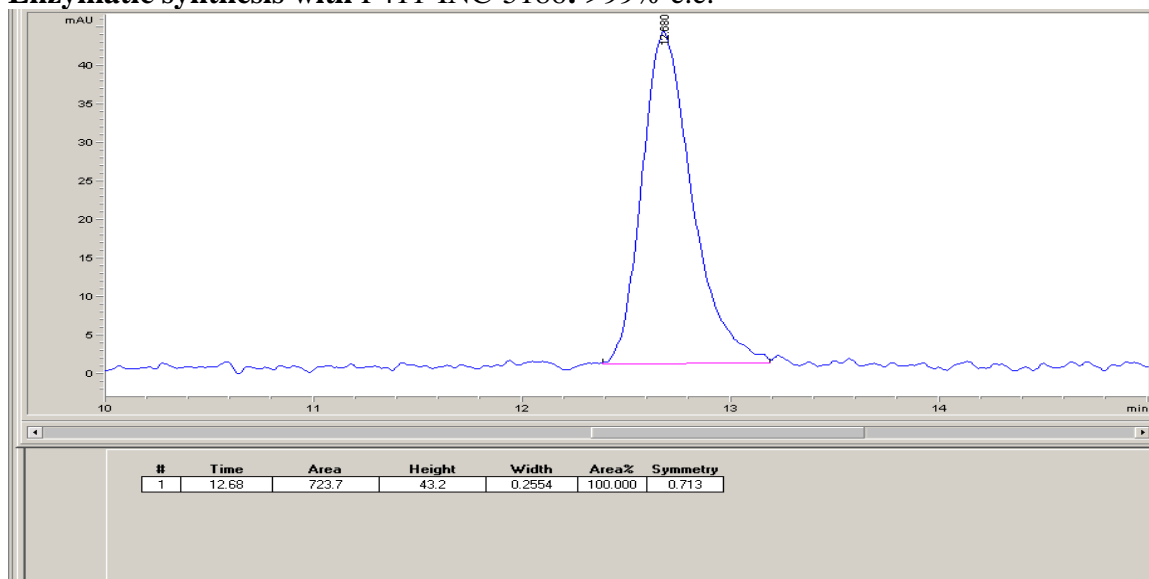
**(*S*)-3-(4-bromobenzoyl)pentanenitrile (4g; from 3g)**



**Chiral-phase HPLC conditions:** Chiralpak IA, 5% *i*-PrOH in hexane, 1.0 mL/min, 25 ºC, 245 nm

**Racemic 4g:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 13.376 | 855.1 | 44.5 | 0.3202 | 49.554 | 0.683 |
| 2 | 14.827 | 870.5 | 41.5 | 0.3499 | 50.446 | 0.705 |

**Enzymatic synthesis with** P411-INC-5185**:** 98% e.e.



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 13.381 | 15.5 | 1.2 | 0.2159 | 1.253 | 1.91 |
| 2 | 14.702 | 1220.7 | 58.6 | 0.3474 | 98.747 | 0.66 |

**Enzymatic synthesis with** P411-INC-5186**:** 98% e.e.



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 13.316 | 140.1 | 8.5 | 0.2752 | 1.048 | 1.03 |
| 2 | 14.557 | 13231.3 | 540.1 | 0.4083 | 98.952 | 0.486 |
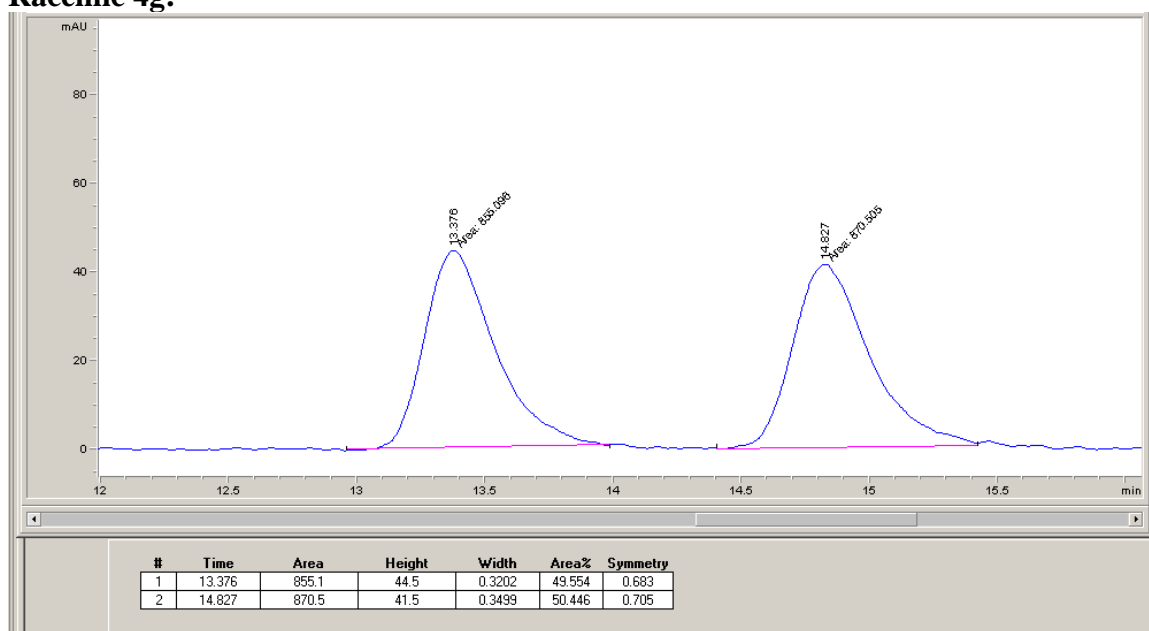
**(*S*)-3-(4-methoxybenzoyl)pentanenitrile (4h; from 3h)**



**Chiral-phase HPLC conditions:** Chiralpak IA, 5% *i*-PrOH in hexane, 1.0 mL/min, 25 ℃, 245 nm

**Racemic 4h:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 17.61 | 1589.1 | 58.7 | 0.4516 | 49.587 | 0.524 |
| 2 | 19.968 | 1615.6 | 52.6 | 0.5117 | 50.413 | 0.622 |
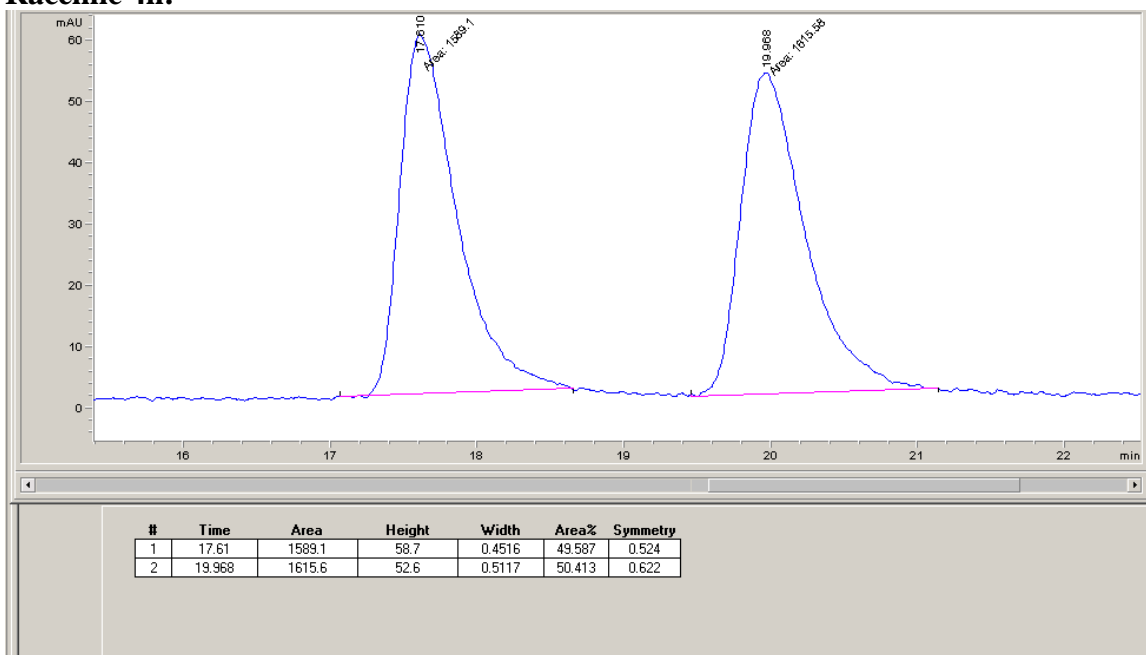
**Enzymatic synthesis with** P411-INC-5185**:** >99% e.e.

**Enzymatic synthesis with** P411-INC-5186**:** 97% e.e.



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 17.291 | 39.7 | 2.5 | 0.2691 | 1.486 | 0.594 |
| 2 | 19.379 | 2635.3 | 92.3 | 0.476 | 98.514 | 0.591 |
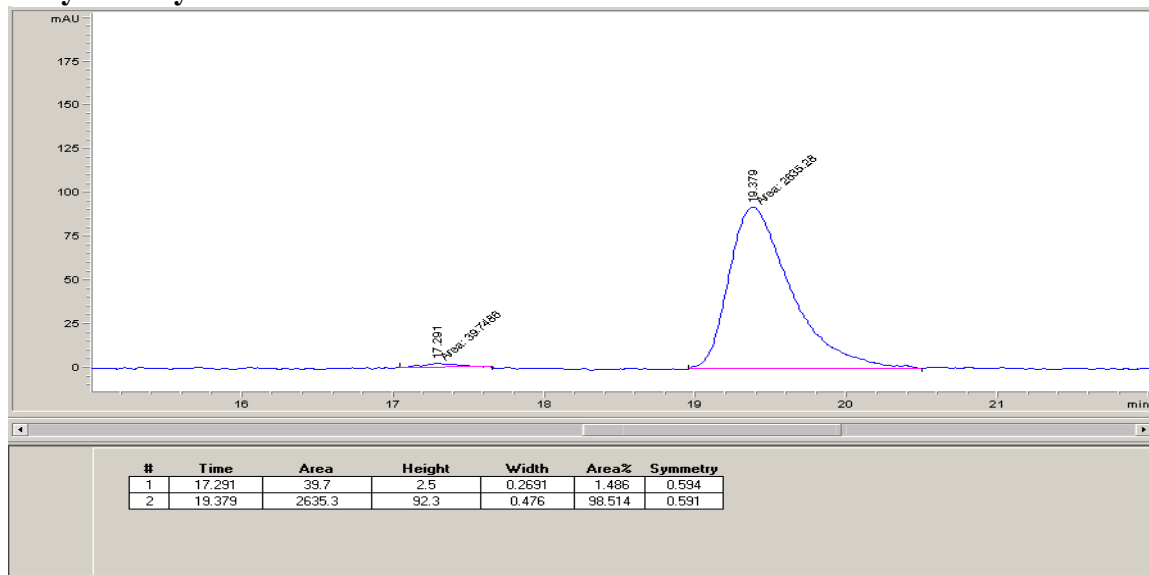
**(*S*)-3-(thiophene-2-carbonyl)pentanenitrile (4i; from 3i)**
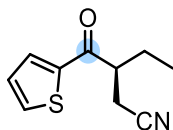


**Chiral-phase HPLC conditions:** Chiralpak IB, 15% *i*-PrOH in hexane, 1.0 mL/min, 25 ℃, 245 nm

**Racemic 4i:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 10.566 | 1296.4 | 100.9 | 0.2141 | 50.382 | 0.755 |
| 2 | 11.533 | 1276.7 | 93.6 | 0.2274 | 49.618 | 0.754 |

**Enzymatic synthesis with** P411-INC-5185**:** >99% e.e.



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 11.513 | 906.6 | 64.9 | 0.2329 | 100.000 | 0.761 |

**Chiral-phase HPLC conditions:** Chiralpak IB, 20% *i*-PrOH in hexane, 1.0 mL/min, 25 ºC, 254 nm

**Racemic 4i:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 8.732 | 1067.7 | 99.3 | 0.1792 | 49.940 | 0.772 |
| 2 | 9.44 | 1070.3 | 94.5 | 0.1888 | 50.060 | 0.824 |

**Enzymatic synthesis with** P411-INC-5186**: >99% e.e.**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 9.461 | 508.7 | 45.6 | 0.186 | 100.000 | 0.864 |
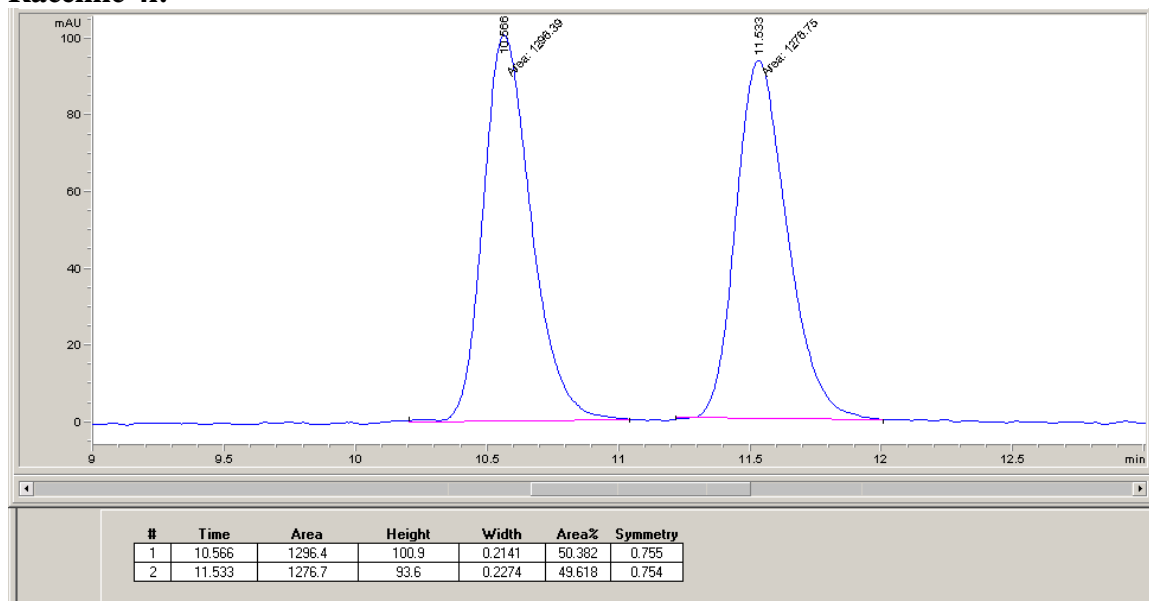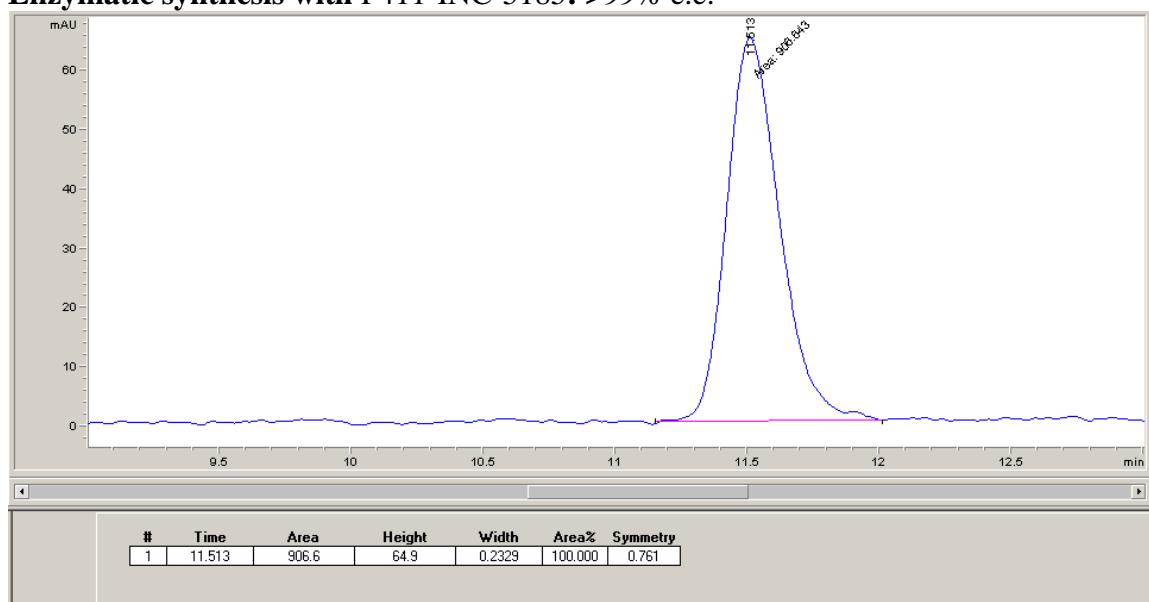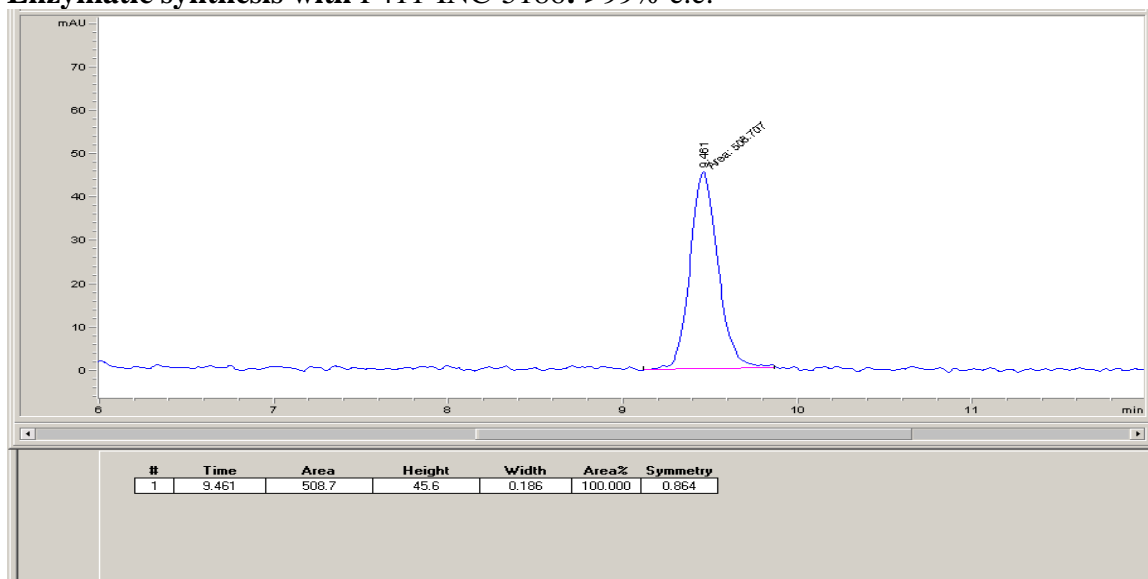
**(*S*)-3-methyl-4-oxo-4-phenylbutanenitrile (4k; from 3k)**
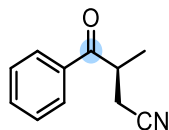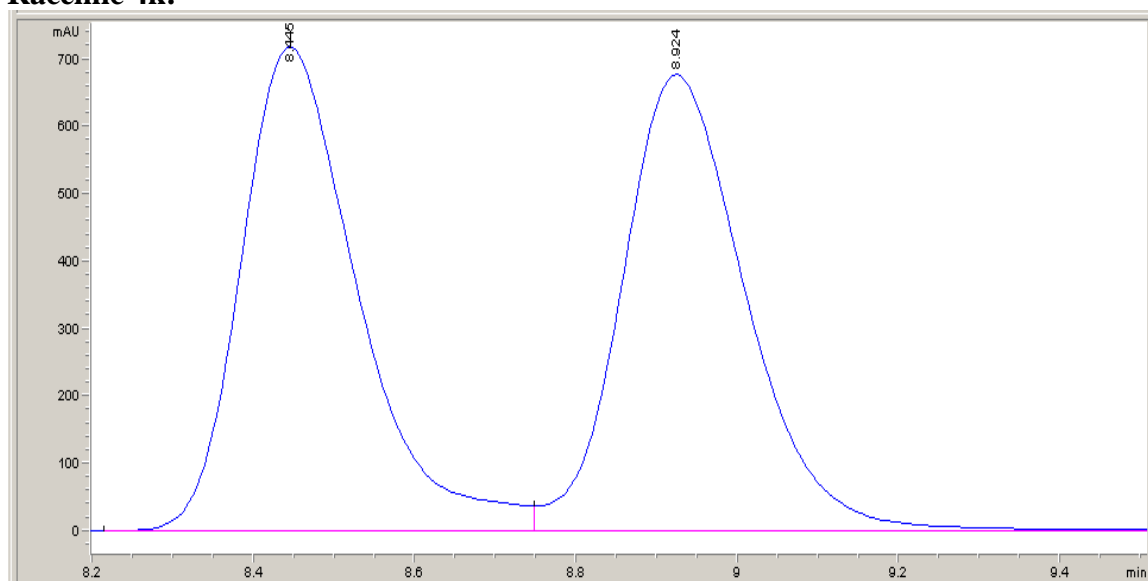


**Chiral-phase HPLC conditions:** Chiralpak IB, 15% *i*-PrOH in hexane, 1.0 mL/min, 25 ºC, 245 nm

**Racemic 4k:**

**Enzymatic synthesis with** P411-INC-5185**:** .>99% e.e.



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 8.887 | 168 | 15.2 | 0.1838 | 100.000 | 0.917 |

**Chiral-phase HPLC conditions:** Chiralpak IB, 15% *i*-PrOH in hexane, 1.0 mL/min, 25 ºC, 254 nm

**Racemic 4k:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 8.426 | 7419.7 | 771.4 | 0.1603 | 49.860 | 0.71 |
| 2 | 10.135 | 7461.3 | 654.5 | 0.19 | 50.140 | 0.764 |

**Enzymatic synthesis with** P411-INC-5186**:** 93% e.e.



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|---|---|---|---|---|---|
| 1 | 8.631 | 43.1 | 6 | 0.1194 | 3.480 | 1.093 |
| 2 | 10.321 | 1196.7 | 107.5 | 0.1856 | 96.520 | 0.858 |
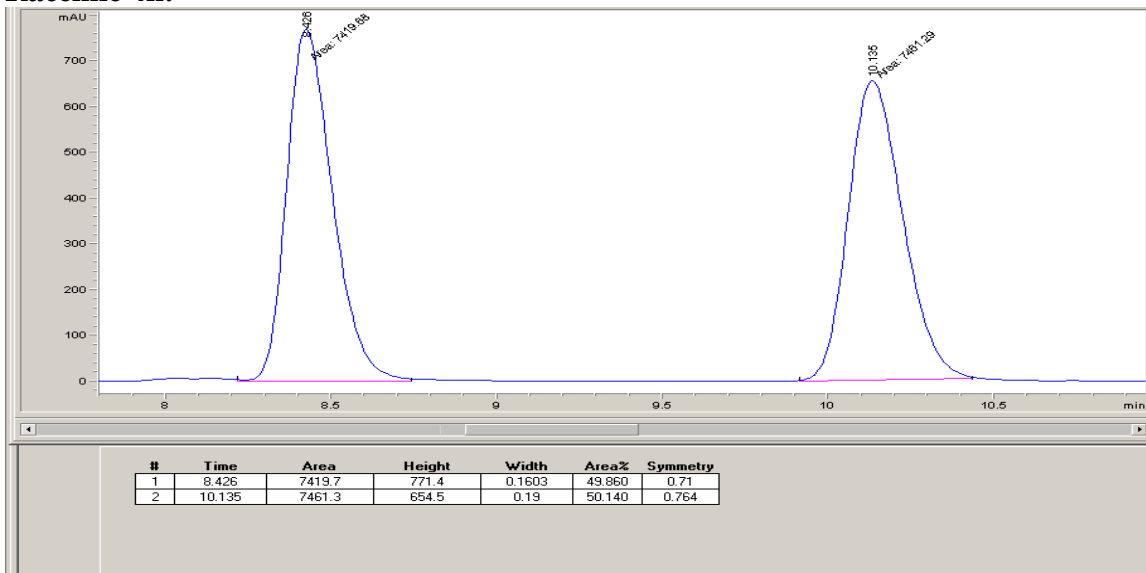
**(*S*)-3-benzoylhexanenitrile (4l; from 3l)**
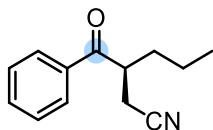


**Chiral-phase HPLC conditions:** Chiralpak IB, 15% *i*-PrOH in hexane, 1.0 mL/min, 25 ℃, 245 nm
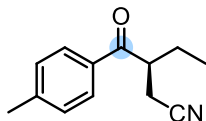
**Racemic 4l:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|---|---|---|---|---|---|
| 1 | 7.193 | 6087.4 | 744 | 0.1364 | 50.220 | 0.783 |
| 2 | 9.564 | 6034.1 | 518.6 | 0.1939 | 49.780 | 0.858 |

**Enzymatic synthesis with** P411-INC-5185**:** 94% e.e.



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 7.225 | 72 | 10.4 | 0.1159 | 2.958 | 1.084 |
| 2 | 9.597 | 2361.4 | 206.8 | 0.1903 | 97.042 | 0.846 |

**Chiral-phase HPLC conditions:** Chiralpak IB, 15% *i*-PrOH in hexane, 1.0 mL/min, 25 ºC, 254 nm

**Racemic 4l:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 6.92 | 8897.6 | 1172.7 | 0.1265 | 49.577 | 0.787 |
| 2 | 8.89 | 9049.2 | 882.6 | 0.1709 | 50.423 | 0.834 |

**Enzymatic synthesis with** P411-INC-5186**:** 91% e.e.



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 6.939 | 66.7 | 9.6 | 0.1152 | 4.357 | 0.795 |
| 2 | 8.916 | 1463 | 141 | 0.173 | 95.643 | 0.855 |
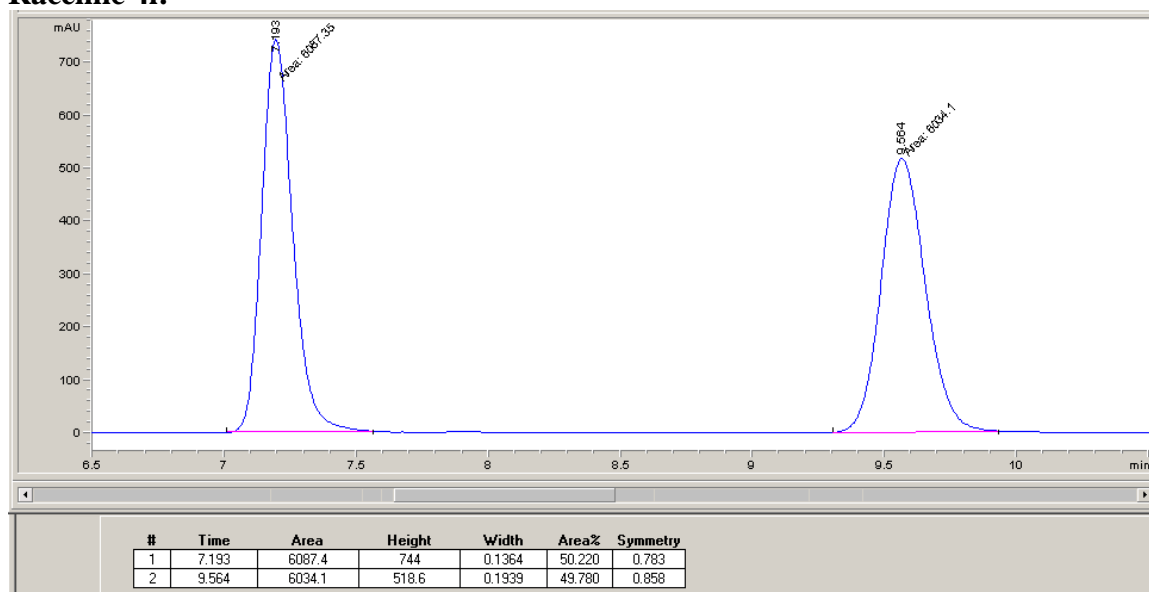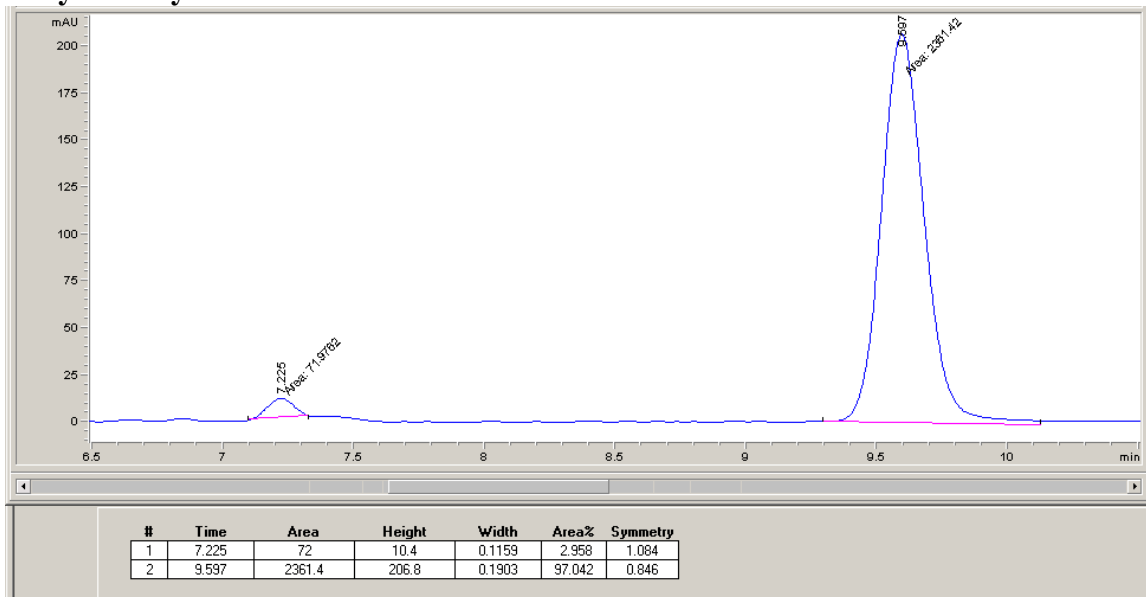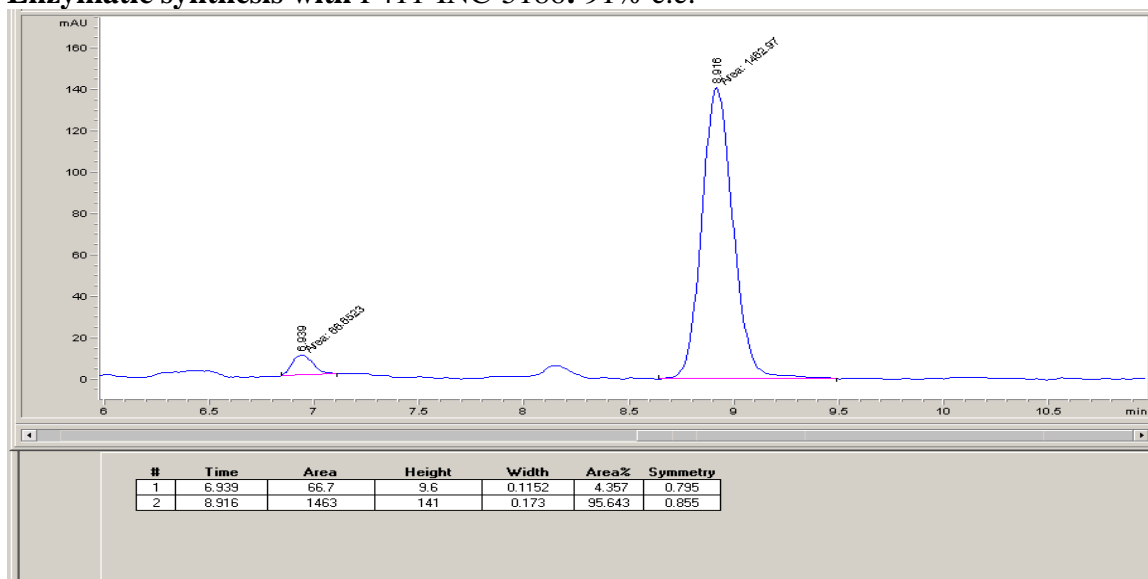
**(*S*)-3-methyl-4-oxo-4-(p-tolyl)butanenitrile (4m; from 3m)**
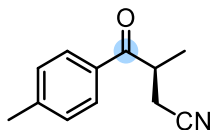


**Chiral-phase HPLC conditions:** Chiralpak IA, 3% *i*-PrOH in hexane, 1.0 mL/min, 25 ℃, 254 nm

**Racemic 4m:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 16.198 | 1104.3 | 48.5 | 0.3796 | 49.881 | 0.767 |
| 2 | 17.647 | 1109.6 | 45.8 | 0.4036 | 50.119 | 0.698 |

**Enzymatic synthesis with** P411-INC-5185**:** 93% e.e.



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 15.739 | 46.4 | 2.6 | 0.3019 | 3.483 | 0.813 |
| 2 | 17.214 | 1286.4 | 52.8 | 0.4059 | 96.517 | 0.609 |
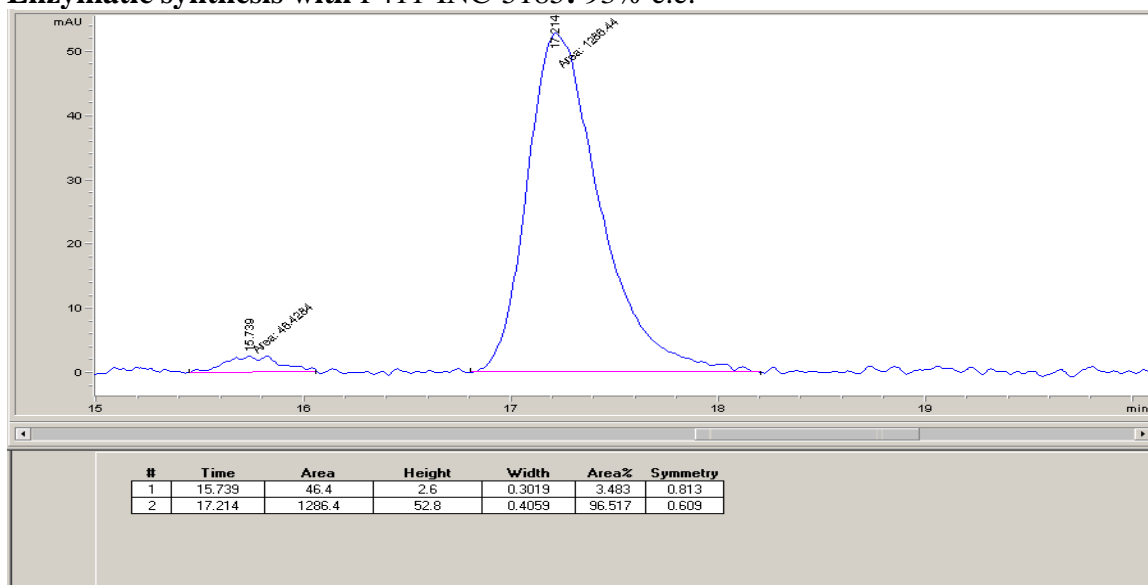
**Enzymatic synthesis with** P411-INC-5186**:** 93% e.e.



| | File Information |
|---|---|
| LC-File | 012-0501.D |
| File Path | C:\CHEM32\1\DATA\RM\221202_IA_3C-PME 1980-02-17 00-20-1 |
| Date | 17-Feb-80, 01:56:31 |
| Sample | 3C-PMe_RO-en |
| Sample Info | |
| Barcode | |
| Operator | RM |
| Method | LZ_3IPA_1ML_25MIN.M |
| Analysis Time | 25 min |
| Sampling Rate | 0.0067 min (0.402 sec), 3751 datapoints |

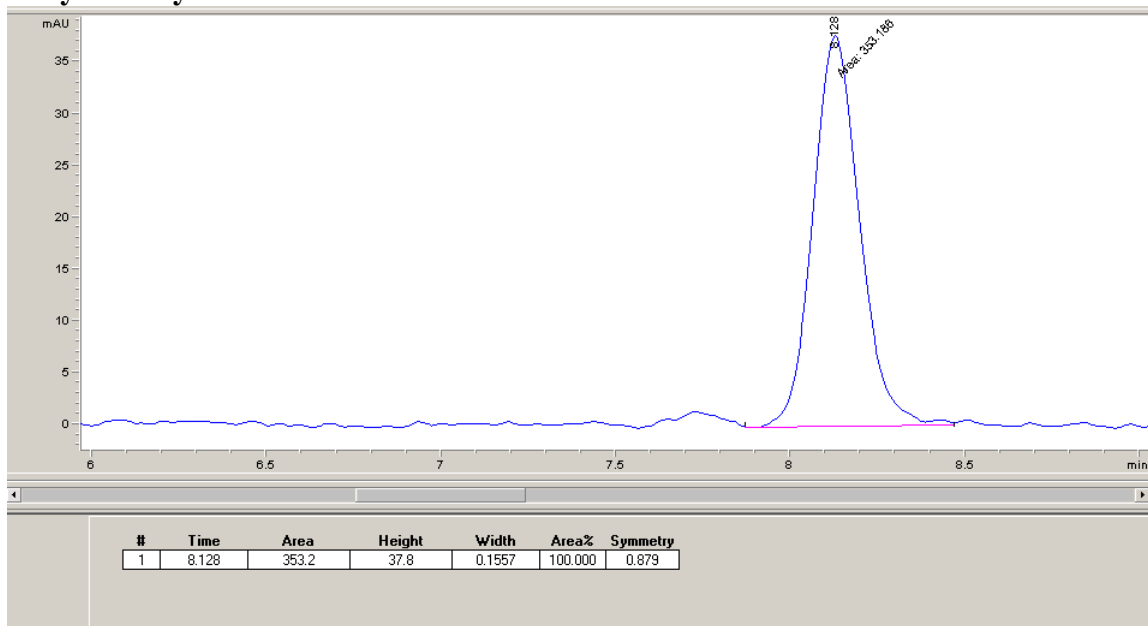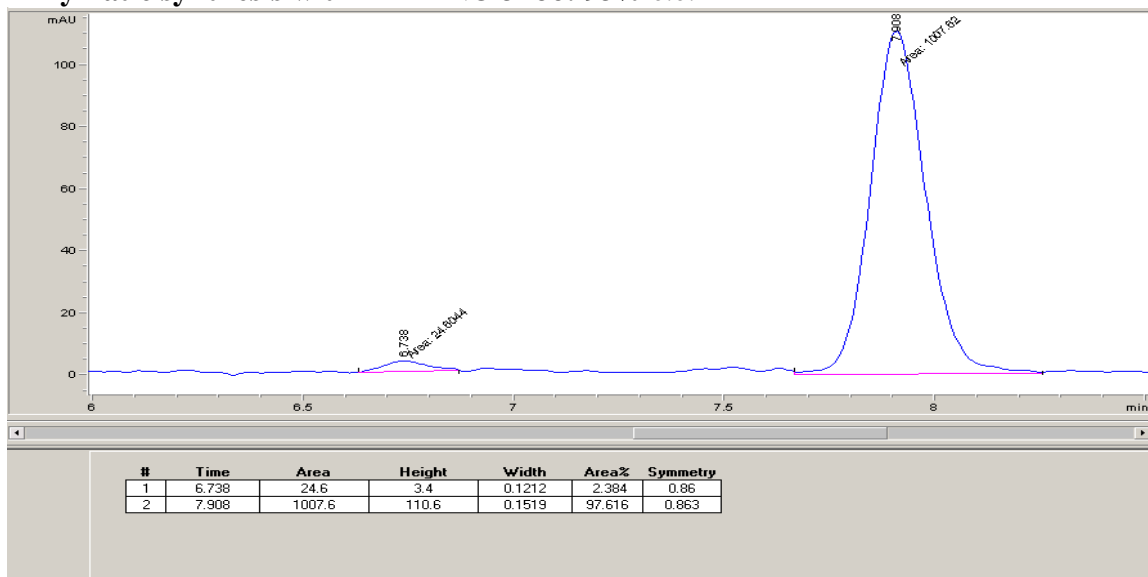| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 15.851 | 54.9 | 3.1 | 0.297 | 3.334 | 1.054 |
| 2 | 17.329 | 1591.4 | 65.5 | 0.4051 | 96.666 | 0.681 |
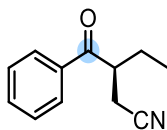
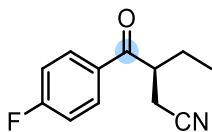**(*S*)-3-benzoylpentanenitrile (4a; obtained from 3n)**



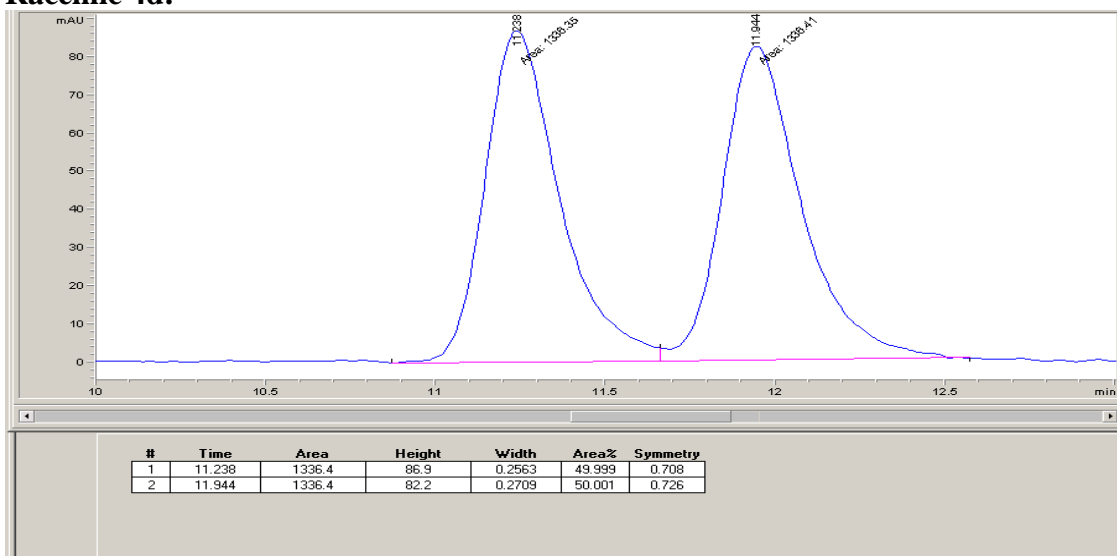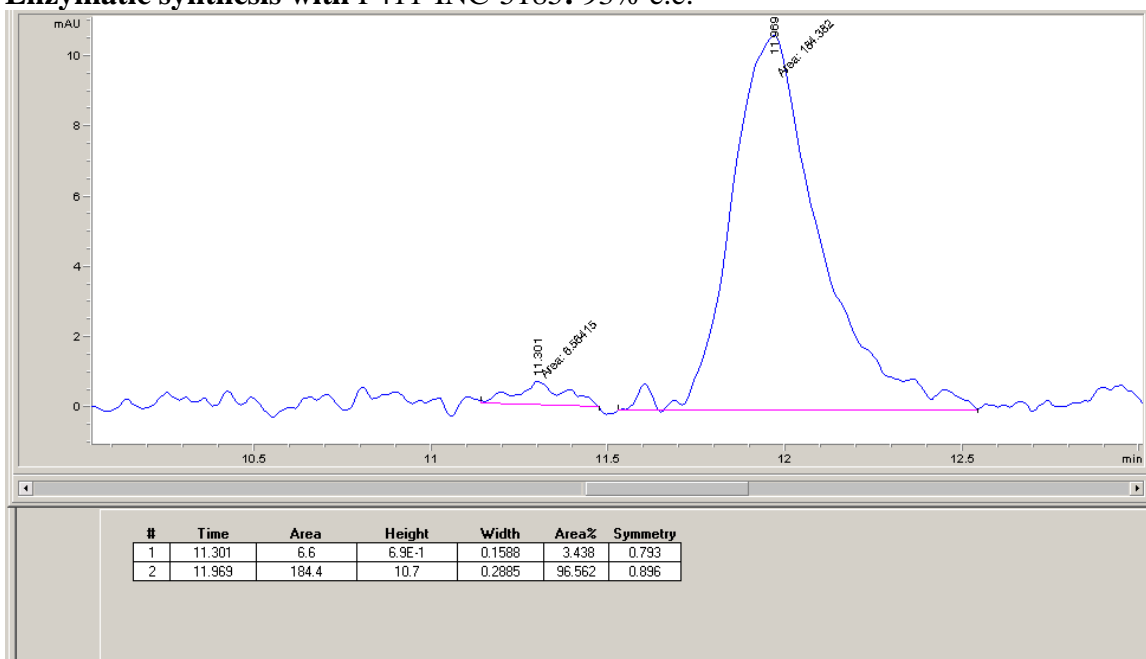**Chiral-phase HPLC conditions:** Chiralpak IB, 15% *i*-PrOH in hexane, 1.0 mL/min, 25 ºC, 254 nm
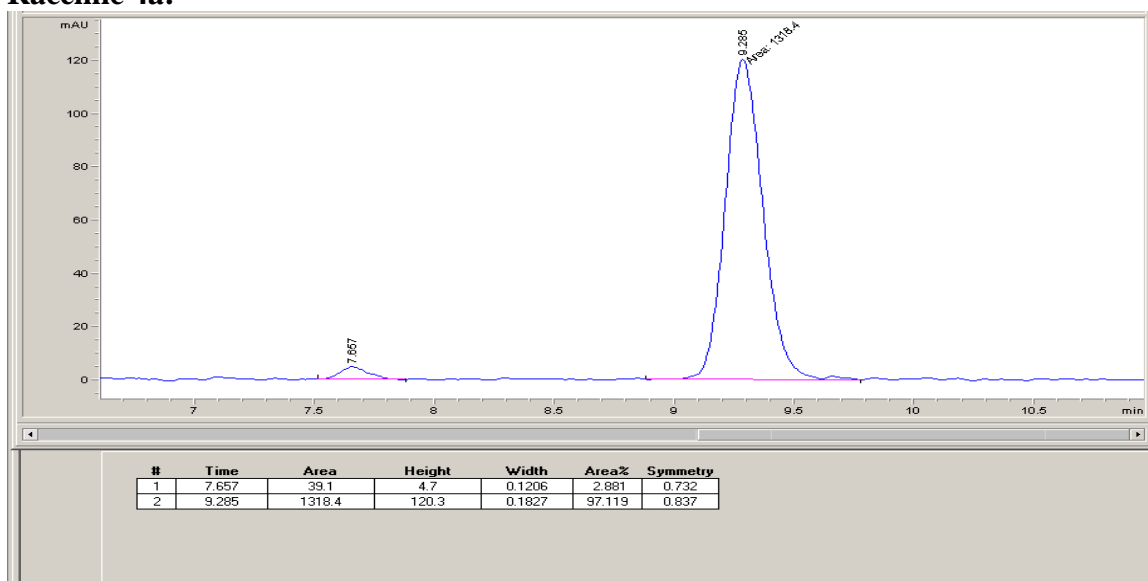
**Racemic 4a:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 7.657 | 39.1 | 4.7 | 0.1206 | 2.881 | 0.732 |
| 2 | 9.285 | 1318.4 | 120.3 | 0.1827 | 97.119 | 0.837 |

**Enzymatic synthesis with** P411-INC-5186**:** 94% e.e.



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 7.657 | 39.1 | 4.7 | 0.1206 | 2.881 | 0.732 |
| 2 | 9.285 | 1318.4 | 120.3 | 0.1827 | 97.119 | 0.837 |
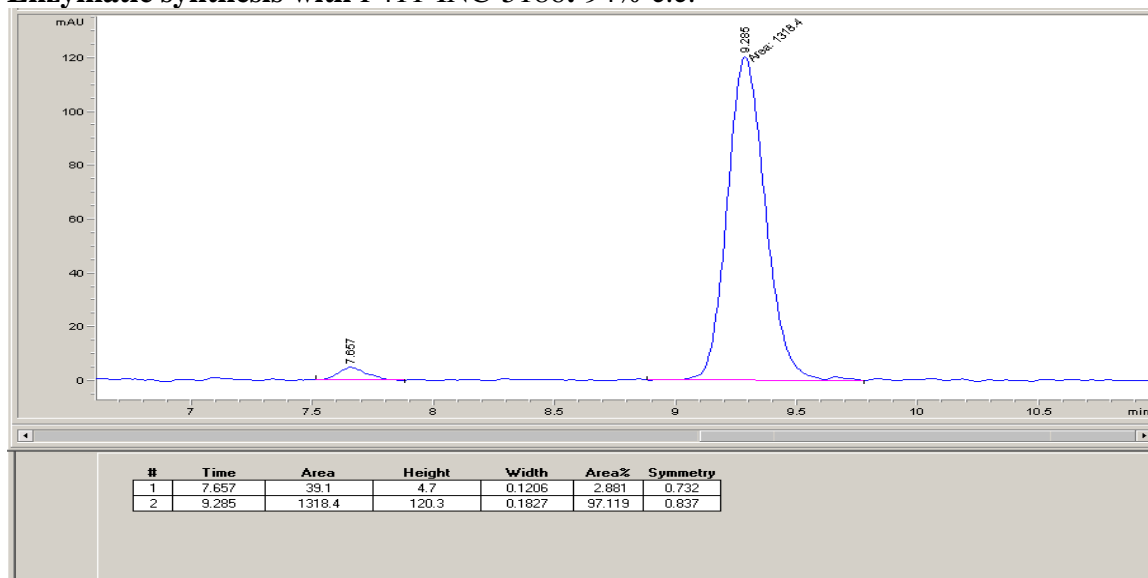
**Enantioselectivity determinations of *α*-branched ketone products 4**

The absolute stereochemistry for enzymatic products **4j** and **4k** were assigned to be *S* by comparing the elution order of each of the two enantiomers with **4j** and **4k** that was obtained from **3j** and **3k** according to the General procedure F:. The other *α*-branched ketone products **4** were assigned by analogy.

**Note:** For some samples, we found the retention time drifts in normal-phase HPLC.

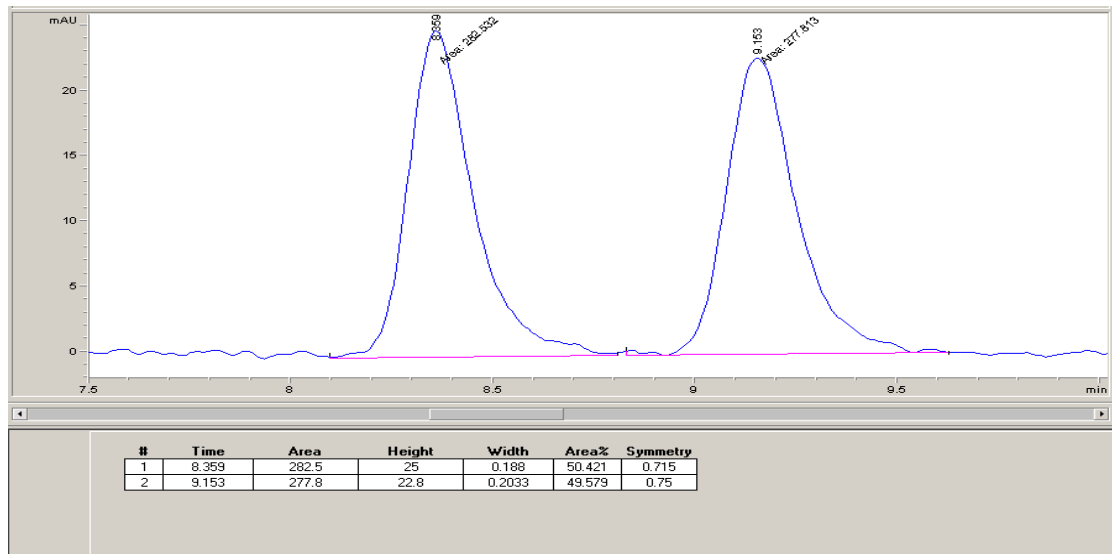**(*S*)-3-benzoylpentanenitrile (4a)**



**Chiral-phase HPLC conditions:** Chiralpak IB, 15% *i*-PrOH in hexane, 1.0 mL/min, 25 ºC, 254 nm

**Racemic 4a:**

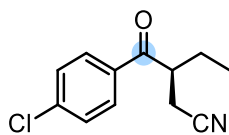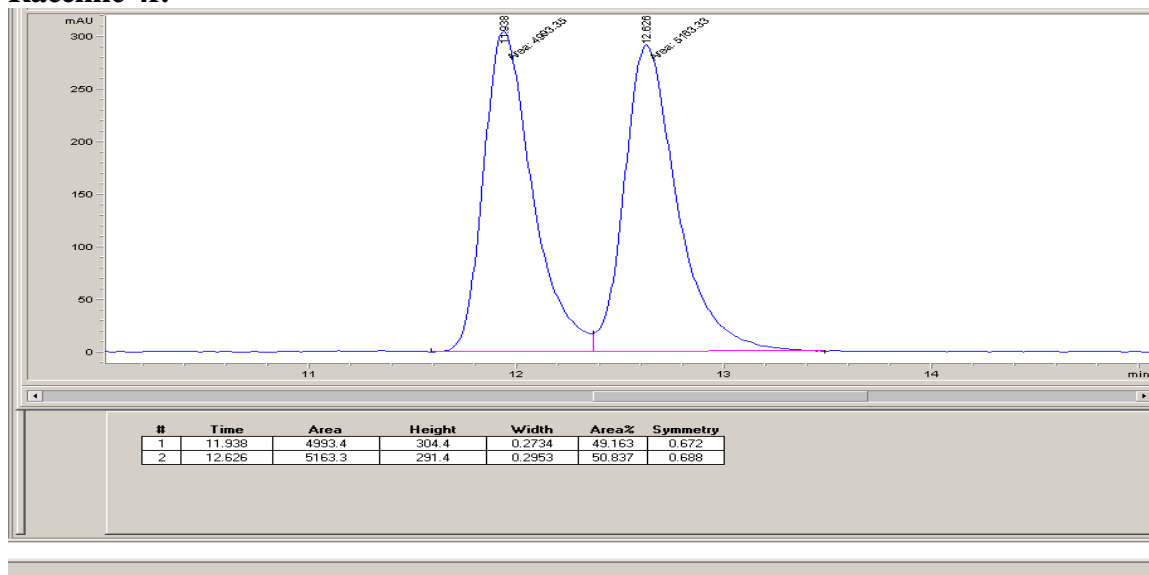

| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 7.64 | 8746.6 | 985.1 | 0.1376 | 49.667 | 0.709 |
| 2 | 9.293 | 8863.7 | 826.9 | 0.1671 | 50.333 | 0.739 |

**Enzymatic preparation of 4a with** P411-INC-5186**:** 97% e.e.



| File Information | |
| --- | --- |
| LC-File | 001-0301.D |
| File Path | C:\CHEM32\1\DATA\RM\20220905_E-4C-PDT_EE_2 1980-05-23 |
| Date | 23-May-80, 09:28:29 |

| # | Time | Area | Height | Width | Area% | Symmetry |
| --- | --- | --- | --- | --- | --- | --- |
| 1 | 7.411 | 15.4 | 2.1 | 0.1197 | 1.500 | 1.145 |
| 2 | 9.08 | 1009.2 | 100.5 | 0.1673 | 98.500 | 0.905 |

**(*S*)-3-(4-methylbenzoyl)pentanenitrile (4b)**



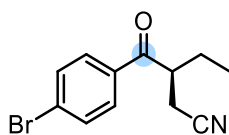**Chiral-phase HPLC conditions:** Chiralpak IA, 5% *i*-PrOH in hexane, 1.0 mL/min, 25 ℃, 245 nm

**Racemic 4b:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 10.467 | 1081.7 | 74.6 | 0.2417 | 49.738 | 0.708 |
| 2 | 11.077 | 1093.1 | 71.5 | 0.2549 | 50.262 | 0.728 |
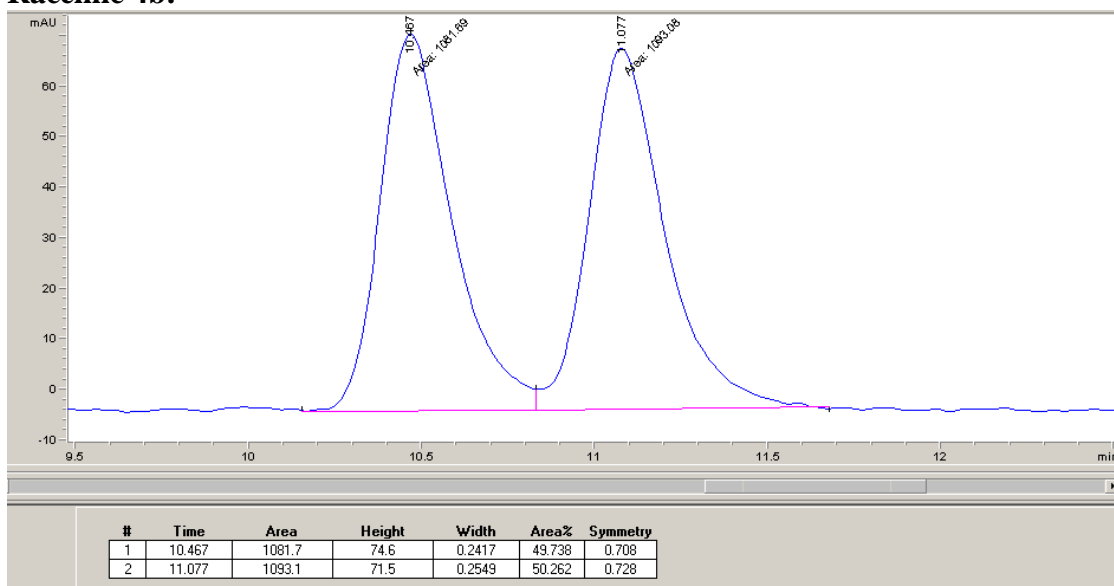
**Enzymatic preparation of 4b with** P411-INC-5186**:** >99% e.e.
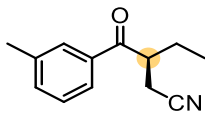


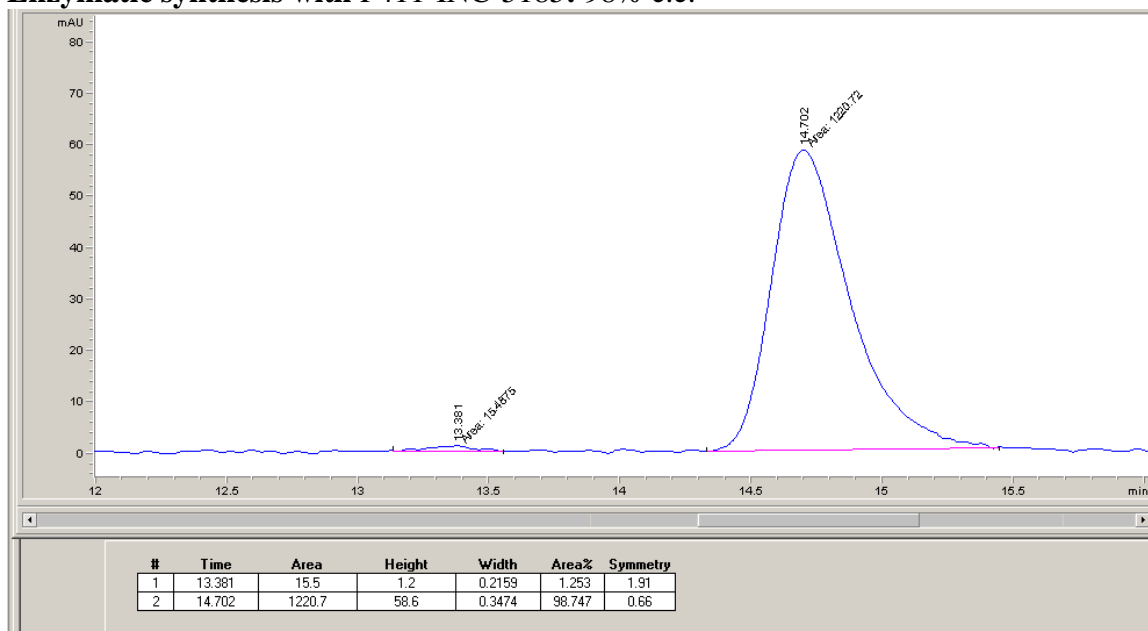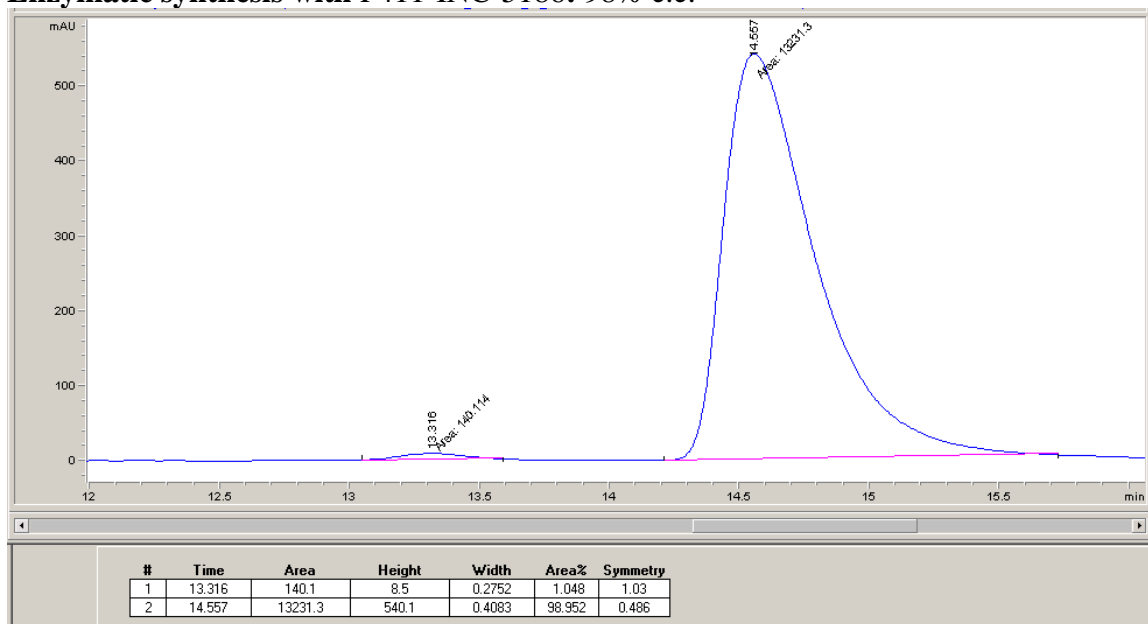| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 11.036 | 351.3 | 23.5 | 0.2493 | 100.000 | 0.785 |

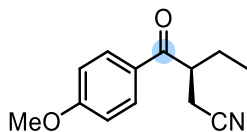**(*S*)-3-(3-methylbenzoyl)pentanenitrile (4c)**



**Chiral-phase HPLC conditions:** Chiralpak IB, 15% *i*-PrOH in hexane, 1.0 mL/min, 25 ℃, 254 nm

**Racemic 4c:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 5.819 | 17427.8 | 2027.3 | 0.1433 | 49.416 | 0.945 |
| 2 | 6.633 | 17840 | 1960.8 | 0.1516 | 50.584 | 0.669 |

**Enzymatic preparation of 4c with** P411-INC-5186**:** >99% e.e.



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 6.636 | 5648.6 | 714.3 | 0.1318 | 100.000 | 0.8 |

## (*S*)-3-(4-fluorobenzoyl)pentanenitrile (4d)



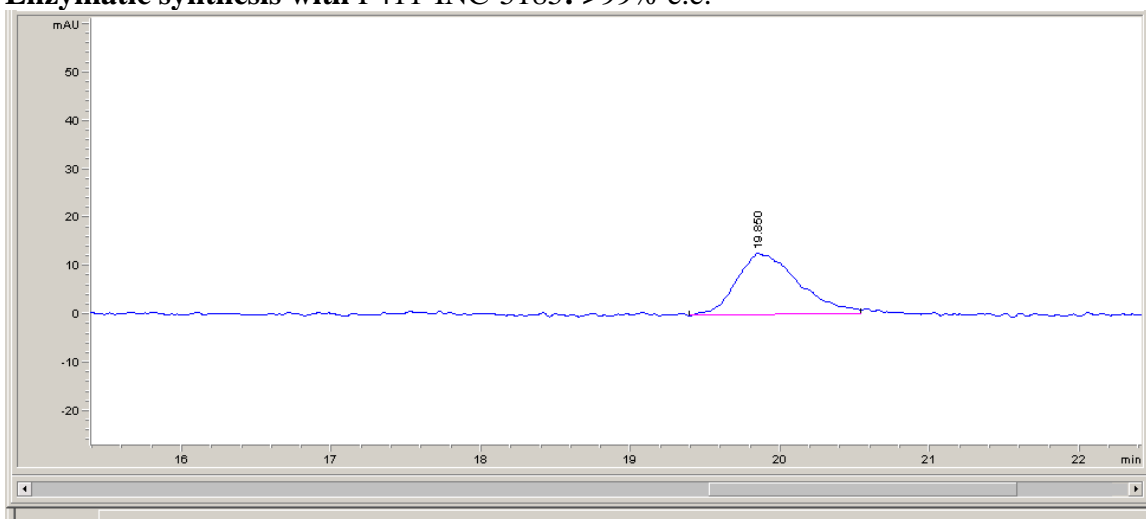**Chiral-phase HPLC conditions:** Chiralpak IA, 5% *i*-PrOH in hexane, 1.0 mL/min, 25 ℃, 254 nm

### Racemic 4d:



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 10.71 | 339.9 | 23.8 | 0.2385 | 49.928 | 0.803 |
| 2 | 11.334 | 340.9 | 22 | 0.2585 | 50.072 | 0.71 |

### Enzymatic preparation of 4d with P411-INC-5186: >99% e.e.



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 11.31 | 726.7 | 48.7 | 0.2485 | 100.000 | 0.731 |

**(*S*)-3-(4-chlorobenzoyl)pentanenitrile (4f)**



**Chiral-phase HPLC conditions:** Chiralpak IA, 5% *i*-PrOH in hexane, 1.0 mL/min, 25 ℃,

245 nm

**Racemic 4f:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 12.391 | 3257.9 | 184.1 | 0.2949 | 49.161 | 0 |
| 2 | 13.078 | 3369.2 | 177.9 | 0.3156 | 50.839 | 0.642 |
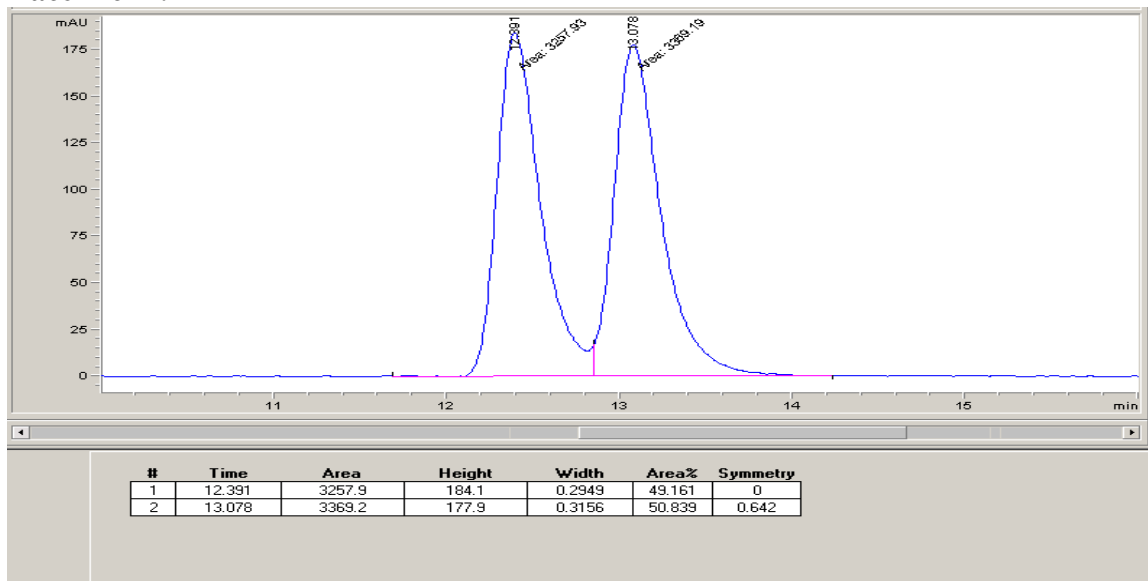
**Enzymatic preparation of 4f with** P411-INC-5186**: >99% e.e.**
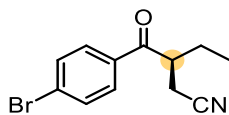


| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 13.107 | 781.8 | 41.2 | 0.3161 | 100.000 | 0.661 |

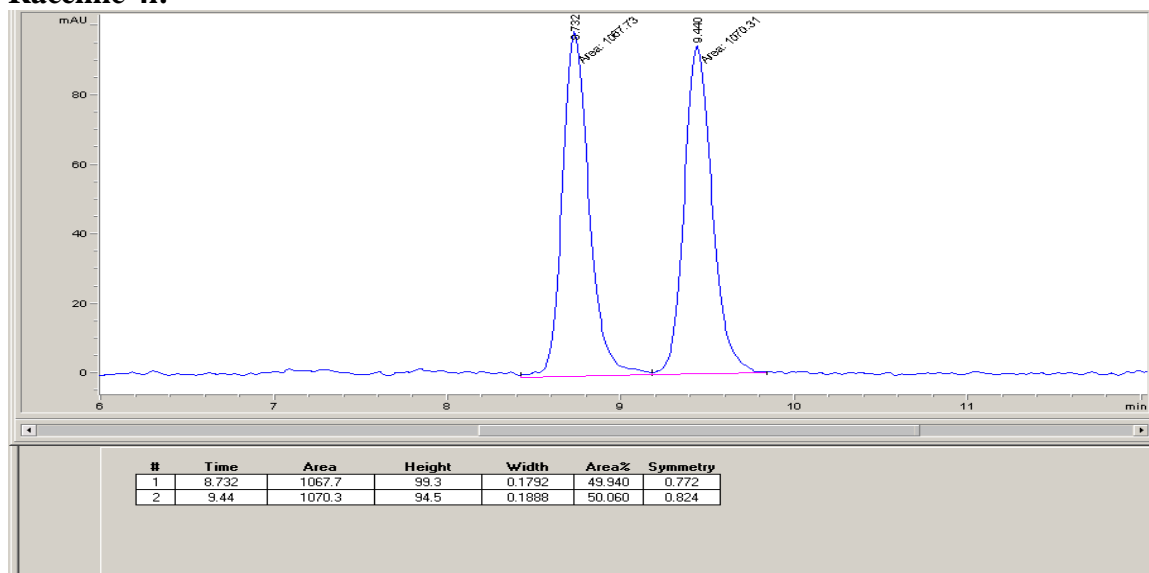**(*S*)-3-(4-bromobenzoyl)pentanenitrile (4g)**



**Chiral-phase HPLC conditions:** Chiralpak IA, 5% *i*-PrOH in hexane, 1.0 mL/min, 25 ℃, 245 nm

**Racemic 4g:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 13.376 | 855.1 | 44.5 | 0.3202 | 49.554 | 0.683 |
| 2 | 14.827 | 870.5 | 41.5 | 0.3499 | 50.446 | 0.705 |
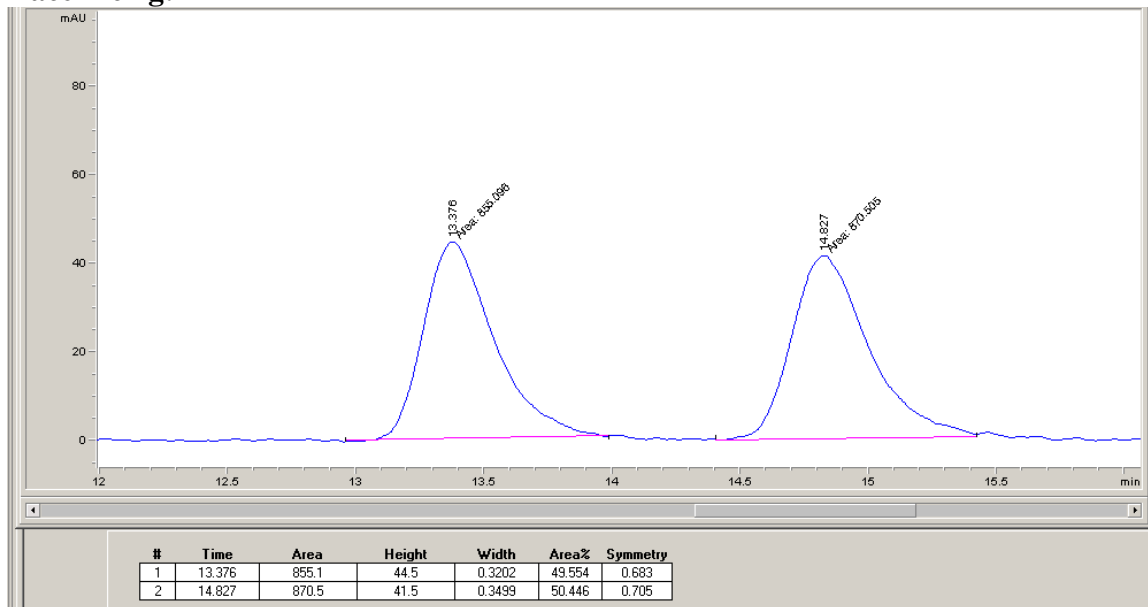
**Enzymatic preparation of 4g with** P411-INC-5186**:** >99% e.e.
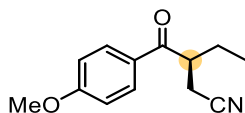


| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 14.712 | 1282.7 | 61.7 | 0.3466 | 100.000 | 0.657 |

**(*S*)-3-(4-methoxybenzoyl)pentanenitrile (4h)**



**Chiral-phase HPLC conditions:** Chiralpak IA, 5% *i*-PrOH in hexane, 1.0 mL/min, 25 ℃, 254 nm

**Racemic 4h:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 17.016 | 9235.9 | 355.4 | 0.4331 | 49.902 | 0.514 |
| 2 | 19.242 | 9272.1 | 315 | 0.4906 | 50.098 | 0.525 |
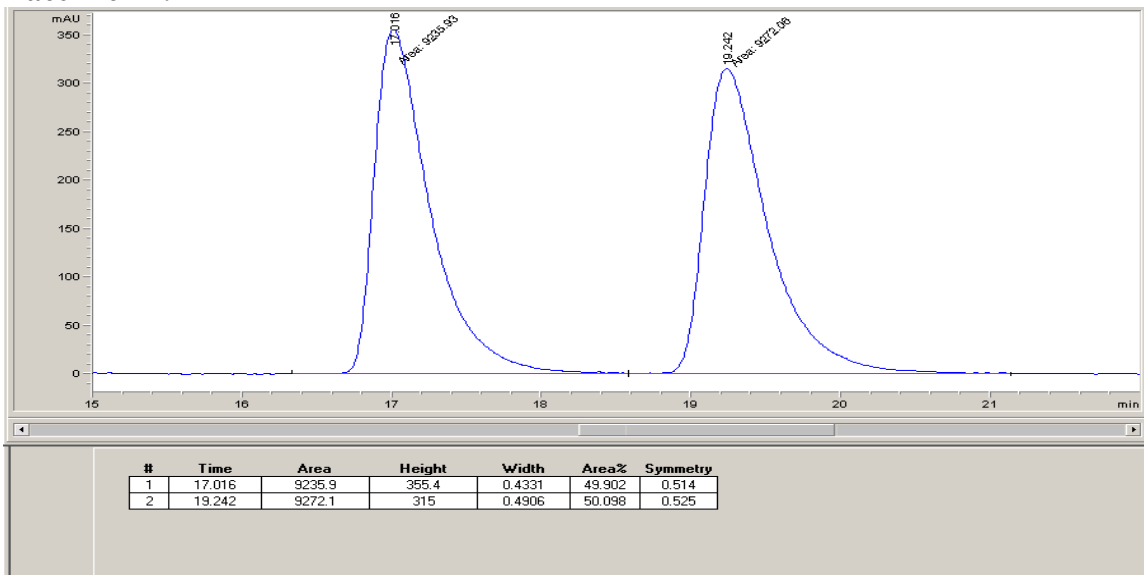
**Enzymatic preparation of 4h with** P411-INC-5186**:** >99% e.e.
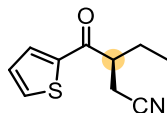


| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 19.389 | 1615.7 | 56.1 | 0.4801 | 100.000 | 0.639 |

**(*S*)-3-(thiophene-2-carbonyl)pentanenitrile (4i)**



**Chiral-phase HPLC conditions:** Chiralpak IB, 20% *i*-PrOH in hexane, 1.0 mL/min, 25 ºC, 254 nm

**Racemic 4i:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 8.732 | 1067.7 | 99.3 | 0.1792 | 49.940 | 0.772 |
| 2 | 9.44 | 1070.3 | 94.5 | 0.1888 | 50.060 | 0.824 |
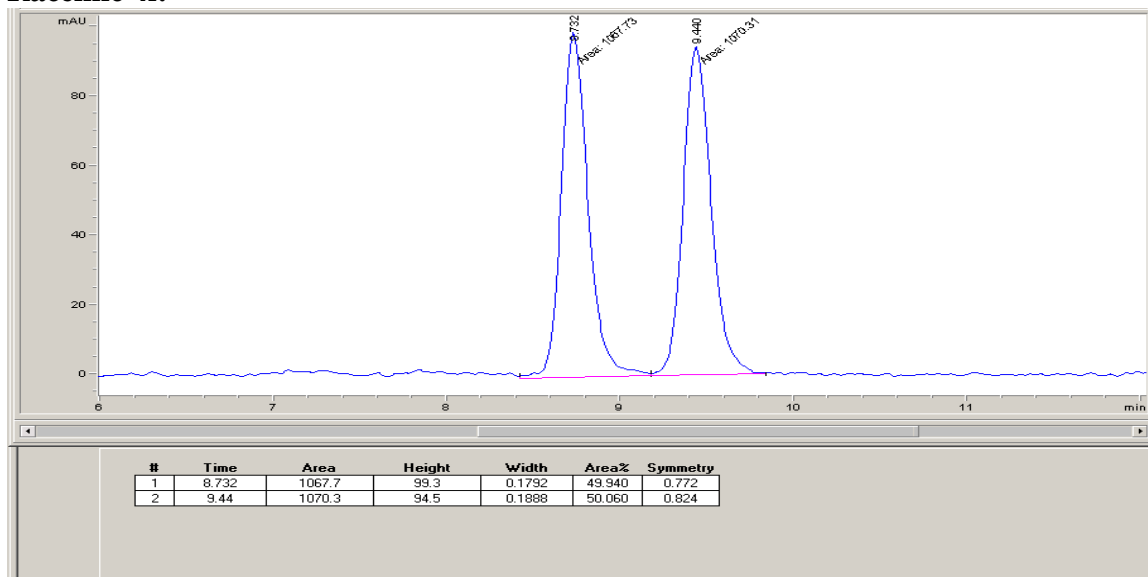
**Enzymatic preparation of 4i with** P411-INC-5186**: >99% e.e.**
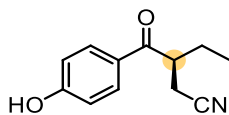


| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 8.955 | 1628.4 | 161.6 | 0.168 | 100.000 | 0.836 |

**(*S*)-3-(4-hydroxybenzoyl)pentanenitrile (4j)**



**Chiral-phase HPLC conditions:** Chiralpak IA, 3% *i*-PrOH in hexane, 1.0 mL/min, 25 ℃, 254 nm
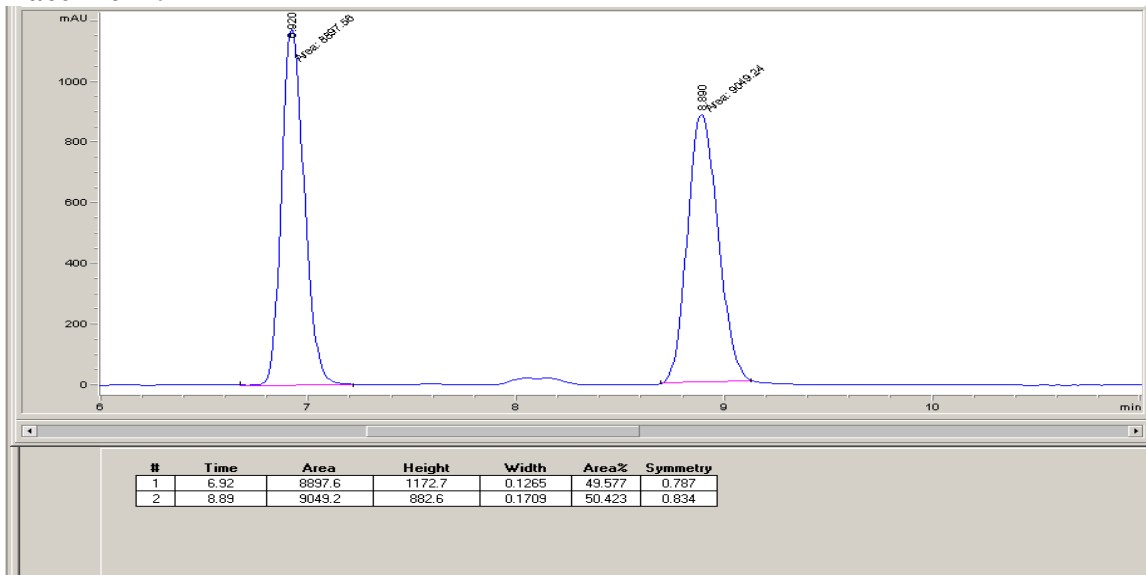
**Racemic 4j:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 17.718 | 2444.4 | 86.8 | 0.4693 | 49.793 | 0.608 |
| 2 | 20.286 | 2464.7 | 80.2 | 0.5124 | 50.207 | 0.614 |
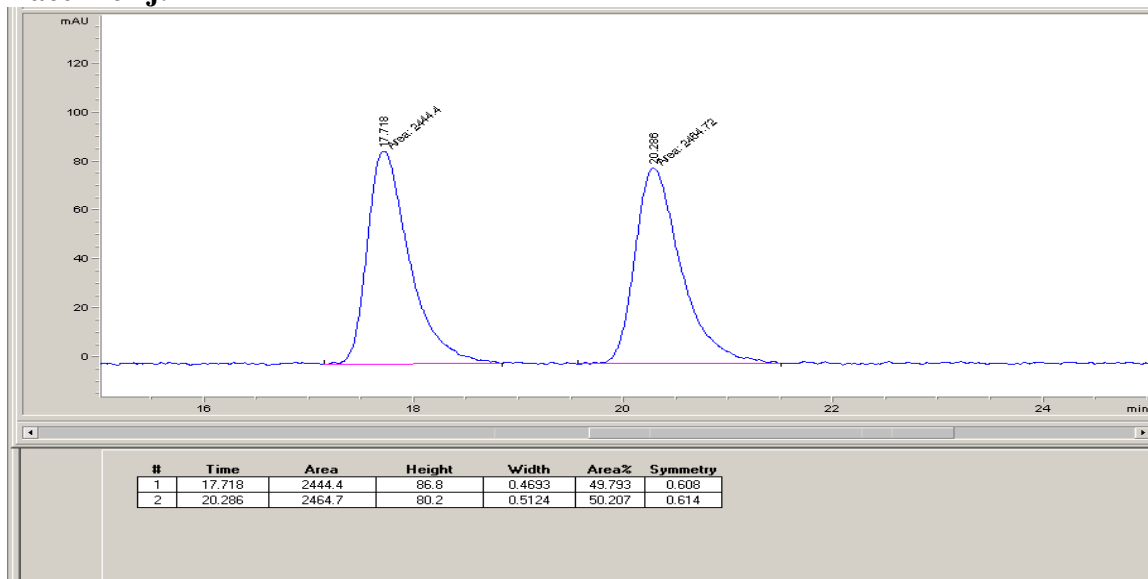
**Enzymatic preparation of 4j with** P411-INC-5186**: >99% e.e.**
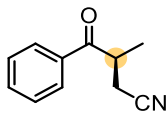


| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 21.189 | 2454.2 | 82.4 | 0.4964 | 100.000 | 0.645 |

**(*S*)-3-methyl-4-oxo-4-phenylbutanenitrile (4k)**



**Chiral-phase HPLC conditions:** Chiralpak IB, 15% *i*-PrOH in hexane, 1.0 mL/min, 25 ℃, 254 nm

**Racemic 4k:**



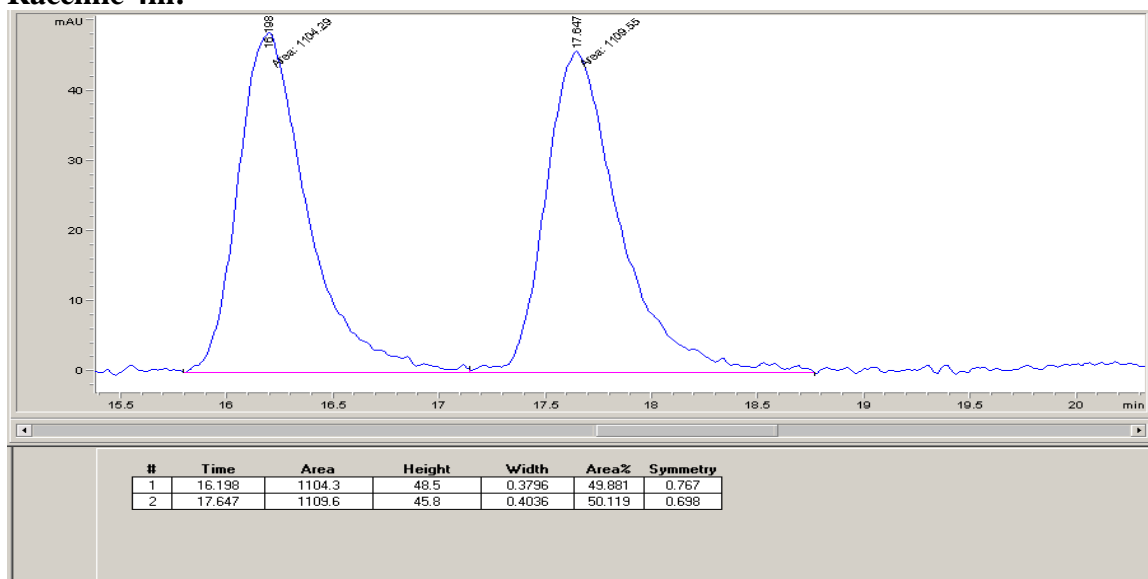| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 8.426 | 7419.7 | 771.4 | 0.1603 | 49.860 | 0.71 |
| 2 | 10.135 | 7461.3 | 654.5 | 0.19 | 50.140 | 0.764 |
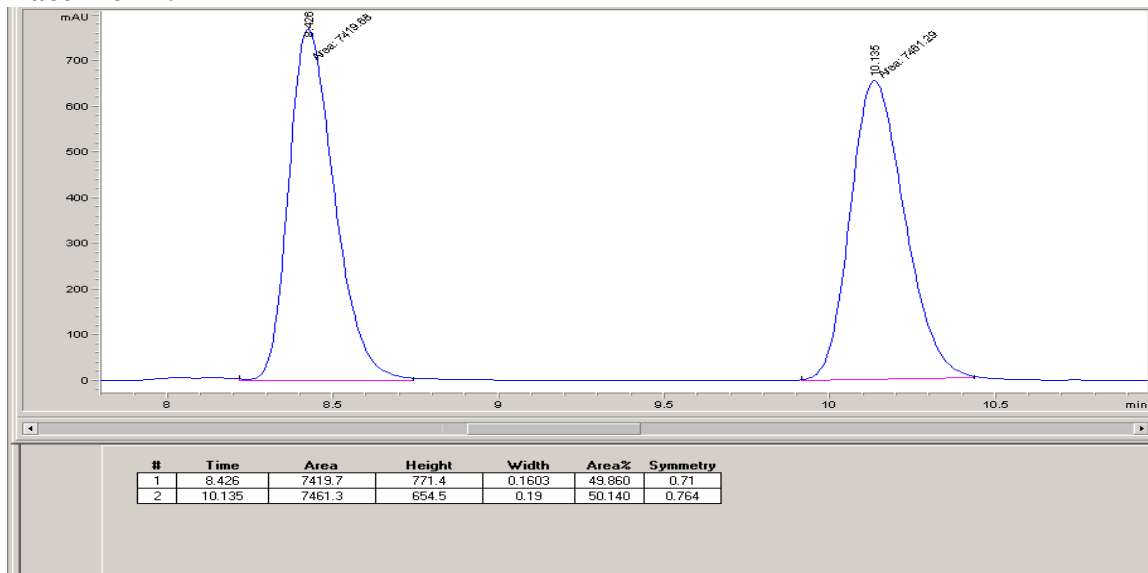
**Enzymatic preparation of 4k with** P411-INC-5186**: >99% e.e.**
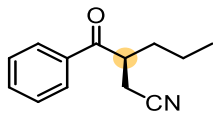


| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 10.156 | 279 | 24.3 | 0.1737 | 100.000 | 0.856 |

**(*S*)-3-benzoylhexanenitrile (4l)**



**Chiral-phase HPLC conditions:** Chiralpak IB, 15% *i*-PrOH in hexane, 1.0 mL/min, 25 ℃, 254 nm

**Racemic 4l:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 7.259 | 3960.4 | 453.8 | 0.1338 | 49.884 | 0.762 |
| 2 | 9.577 | 3978.8 | 342.4 | 0.1816 | 50.116 | 0.822 |

**Enzymatic preparation of 4l with** P411-INC-5186**:** 92% e.e.



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 7.286 | 104.6 | 14.2 | 0.1226 | 3.970 | 0.897 |
| 2 | 9.617 | 2529.6 | 241.4 | 0.1746 | 96.030 | 0.865 |
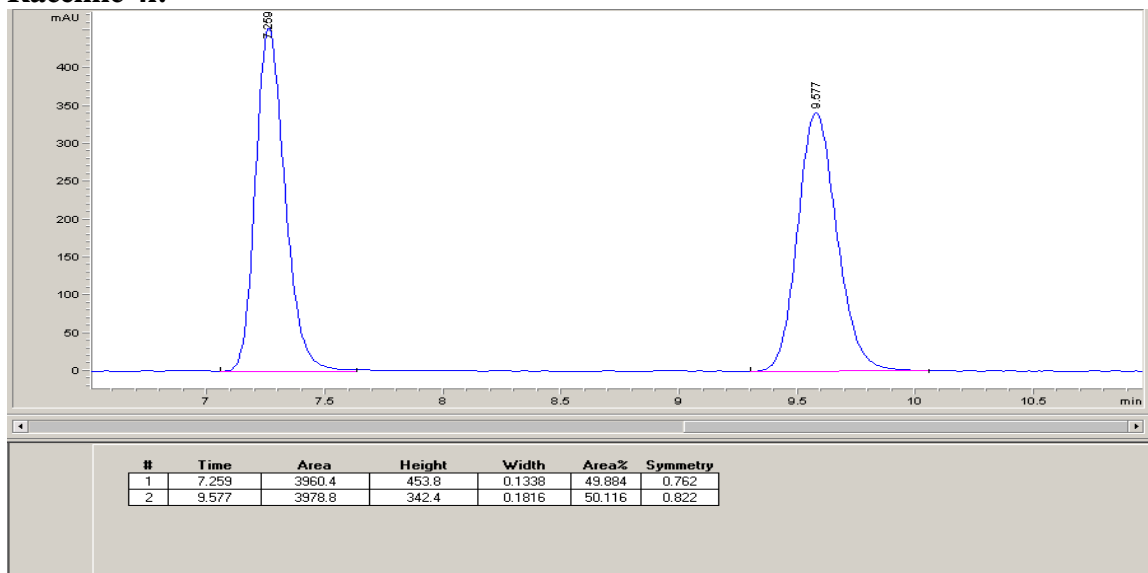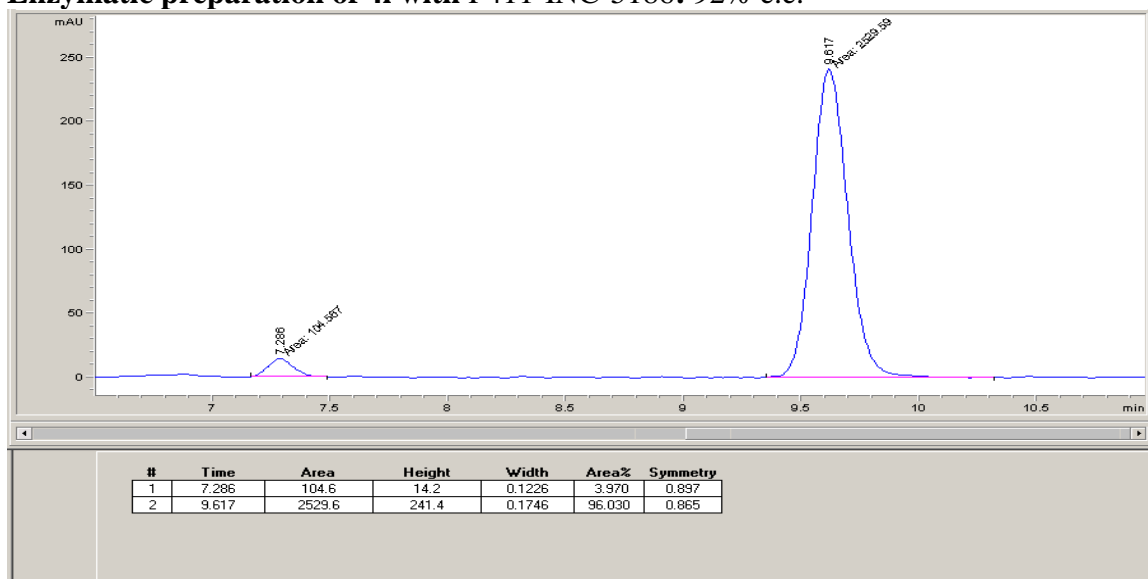
**(*S*)-3-methyl-4-oxo-4-(p-tolyl)butanenitrile (4m)**



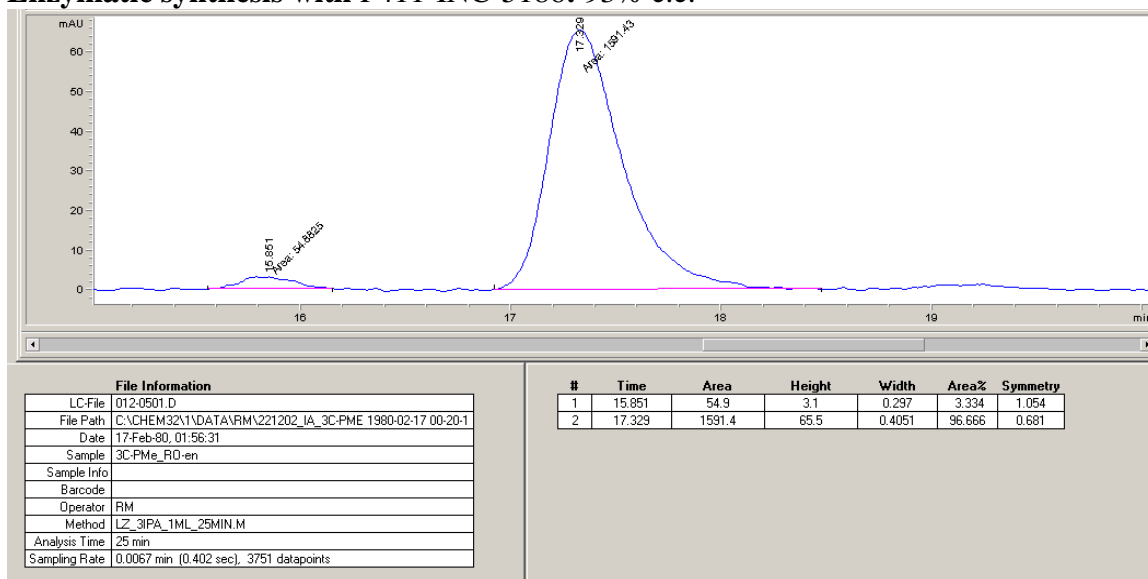**Chiral-phase HPLC conditions:** Chiralpak IA, 3% *i*-PrOH in hexane, 1.0 mL/min, 25 ℃, 245 nm
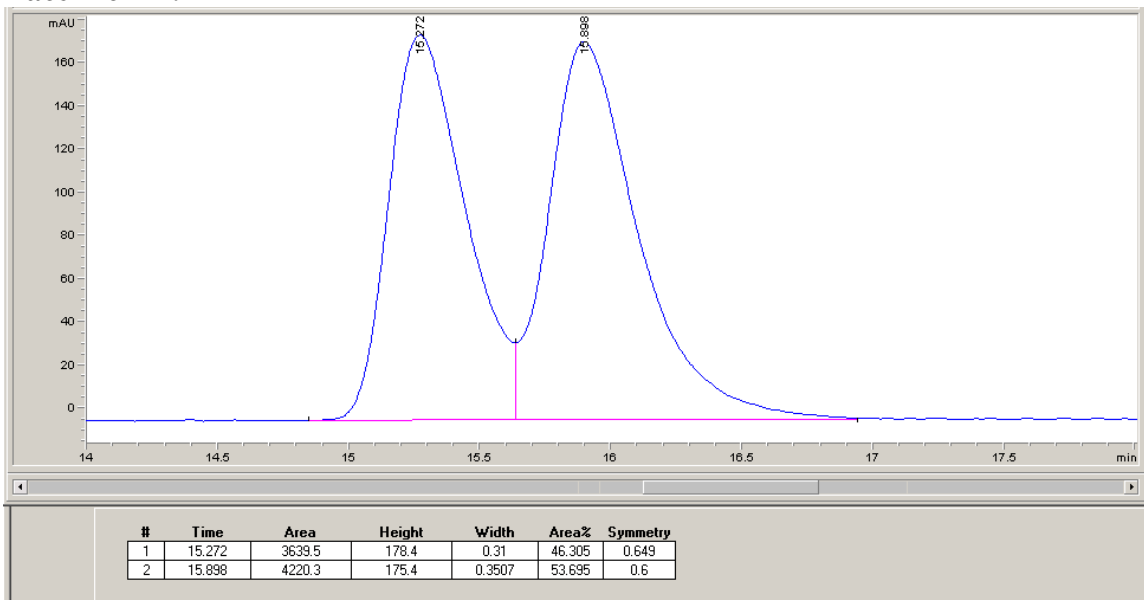
**Racemic 4m:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 15.272 | 3639.5 | 178.4 | 0.31 | 46.305 | 0.649 |
| 2 | 15.898 | 4220.3 | 175.4 | 0.3507 | 53.695 | 0.6 |

**Enzymatic preparation of 4l with** P411-INC-5186**:** 91% e.e.



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 15.519 | 11.7 | 8.3E-1 | 0.2353 | 4.946 | 0.464 |
| 2 | 16.24 | 225.8 | 9.8 | 0.3822 | 95.054 | 0.666 |
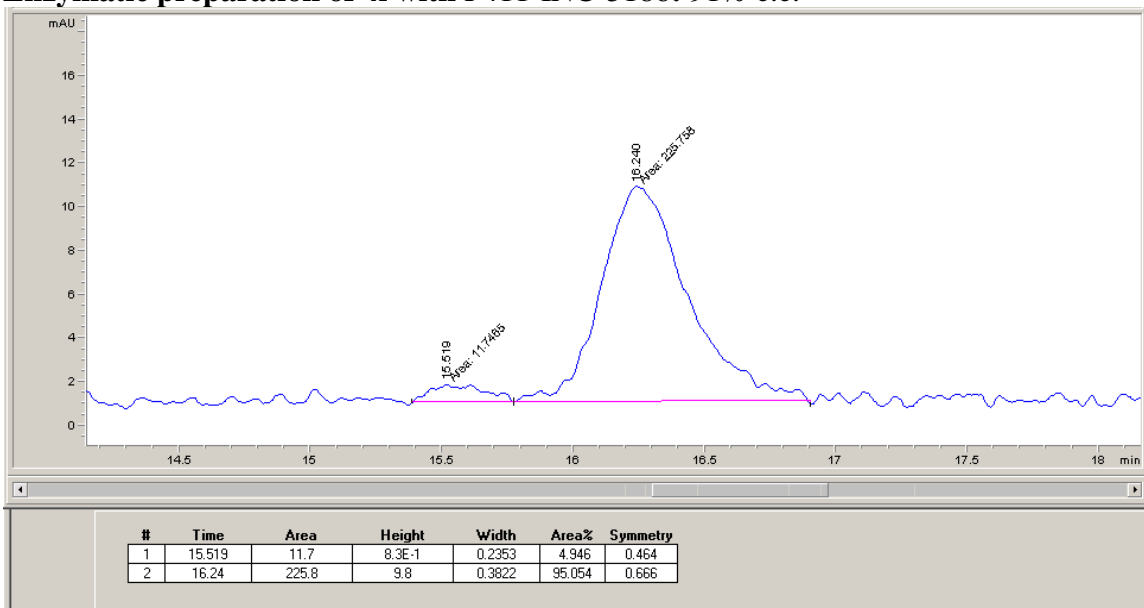
## X-ray Crystallography

Low-temperature diffraction data ($\phi$-and $\omega$-scans) were collected on a Bruker AXS D8 VENTURE KAPPA diffractometer coupled to a PHOTON II CPAD detector with Cu $K_\alpha$ radiation ($\lambda = 1.54178$ Å) from an I$\mu$S micro-source for the structure of compounds. The structure was solved by direct methods using SHELXS[41] and refined against $F^2$ on all data by full-matrix least squares with SHELXL-2019[42] using established refinement techniques.[43] All non-hydrogen atoms were refined anisotropically. Unless otherwise noted, all hydrogen atoms were included into the model at geometrically calculated positions and refined using a riding model. The isotropic displacement parameters of all hydrogen atoms were fixed to 1.2 times the $U$ value of the atoms they are linked to (1.5 times for methyl groups).
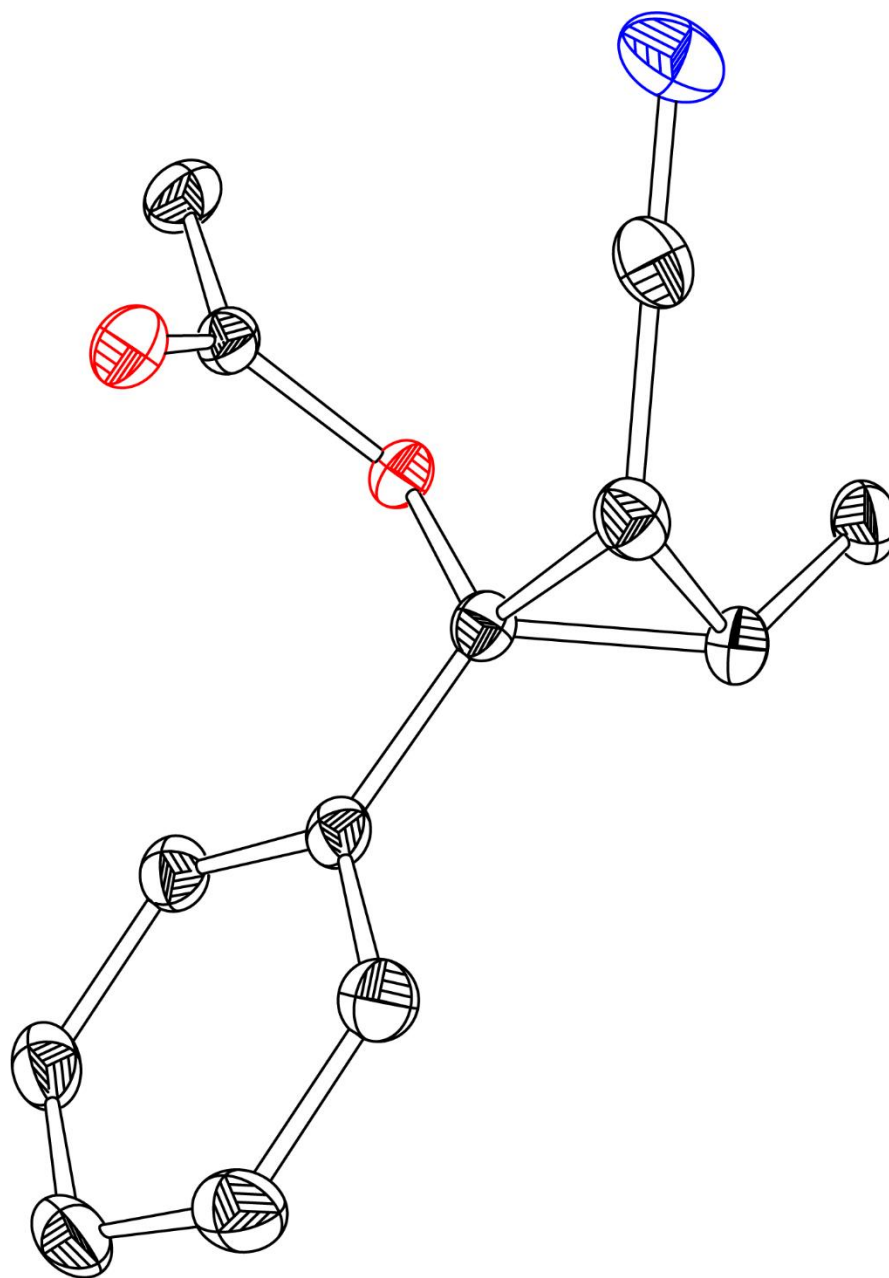
**Figure 2-S2**. Displacement ellipsoid plot for **3j** plotted at 50% probability. Single crystals were obtained from slow evaporation of **3j** dissolved in CHCl$_3$ at 4 °C. Compound **3j** crystallizes in the orthorhombic space group $P2_12_12_1$ with one molecule in the asymmetric unit. The coordinates for the hydrogen atoms bound to O3 were located in the difference Fourier synthesis and refined semi-freely with the help of a restraint on the O–H distance (0.84(4) Å).

**Table 2-S11**. X-ray experimental details of **3j** (CCDC 2218386).

| Crystal data | |
|---|---|
| Chemical formula | $C_{14}H_{15}NO_3$ |
| $M_r$ | 245.27 |
| Crystal system, space group | Orthorhombic, $P2_12_12_1$ |
| Temperature (K) | 100 |
| $a$, $b$, $c$ (Å) | 5.487 (1), 13.0261 (18), 17.561 (3) |
| $V$ (Å$^3$) | 1255.1 (3) |
| $Z$ | 4 |
| Radiation type | Cu $K\alpha$ |
| $\mu$ (mm$^{-1}$) | 0.75 |
| Crystal size (mm) | $0.25 \times 0.20 \times 0.10$ |
| **Data collection** | |
| Diffractometer | Bruker D8 VENTURE Kappa Duo PHOTON II CPAD |
| Absorption correction | Multi-scan *SADABS2016*/2 (Sheldrick, 2014) |
| $T_{min}$, $T_{max}$ | 0.618, 0.754 |
| No. of measured, independent and observed [$I > 2\sigma(I)$] reflections | 20076, 2551, 2523 |
| $R_{int}$ | 0.033 |
| $(\sin \theta/\lambda)_{max}$ (Å$^{-1}$) | 0.625 |
| **Refinement** | |
| $R[F^2 > 2\sigma(F^2)]$, $wR(F^2)$, $S$ | 0.024, 0.063, 1.08 |
| No. of reflections | 2551 |
| No. of parameters | 168 |
| No. of restraints | 1 |
| H-atom treatment | H atoms treated by a mixture of independent and constrained refinement |
| $\Gamma_{max}$, $\Gamma_{min}$ (e Å$^{-3}$) | 0.14, -0.16 |
| Absolute structure | Flack x determined using 1032 quotients [(I+)-(I-)]/[(I+)+(I-)]   (Parsons, Flack and Wagner, Acta Cryst. B69 (2013) 249-259). |
| Absolute structure parameter | 0.00 (4) |
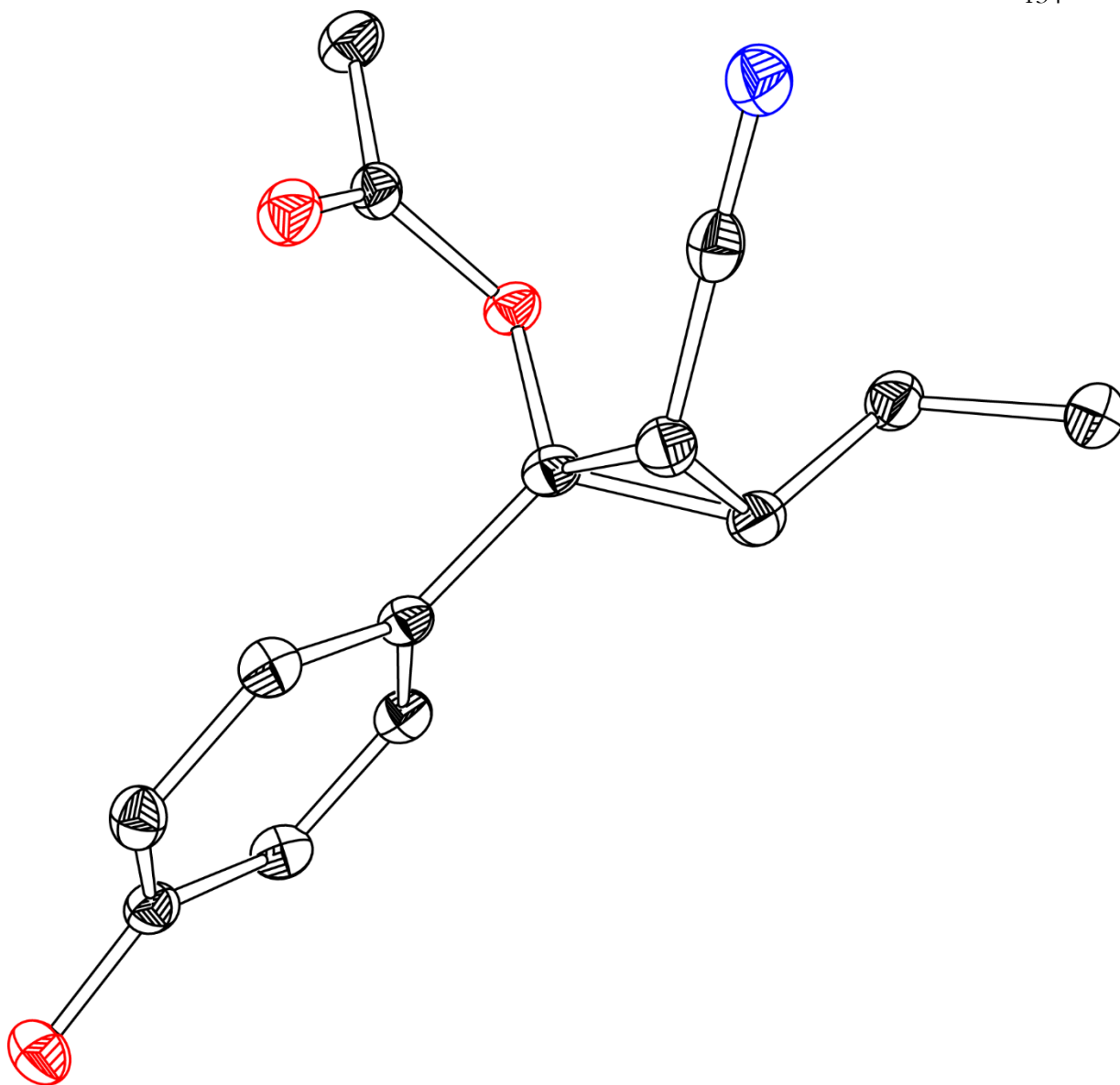
**Figure 2-S3**. Displacement ellipsoid plot for **3k** plotted at 50% probability. Single crystals were obtained from slow evaporation of **3k** dissolved in CHCl$_3$ at 4 °C. Compound **3k** crystallizes in the orthorhombic space group $P2_12_12_1$ with one molecule in the asymmetric unit.

**Table 2-S12**. X-ray experimental details of **3k** (CCDC 2218385).

| Crystal data | |
|---|---|
| Chemical formula | $C_{13}H_{13}NO_2$ |
| $M_r$ | 215.24 |
| Crystal system, space group | Orthorhombic, $P2_12_12_1$ |
| Temperature (K) | 100 |
| $a$, $b$, $c$ (Å) | 5.4031 (5), 9.823 (1), 21.363 (2) |
| $V$ (Å$^3$) | 1133.8 (2) |
| $Z$ | 4 |
| Radiation type | Cu $K\alpha$ |
| $\mu$ (mm$^{-1}$) | 0.69 |
| Crystal size (mm) | $0.20 \times 0.15 \times 0.10$ |
| **Data collection** | |
| Diffractometer | Bruker D8 VENTURE Kappa Duo PHOTON II CPAD |
| Absorption correction | Multi-scan *SADABS2016*/2 (Sheldrick, 2014) |
| $T_{min}$, $T_{max}$ | 0.662, 0.754 |
| No. of measured, independent and observed [$I > 2\sigma(I)$] reflections | 18080, 2313, 2270 |
| $R_{int}$ | 0.036 |
| $(\sin \theta/\lambda)_{max}$ (Å$^{-1}$) | 0.625 |
| **Refinement** | |
| $R[F^2 > 2\sigma(F^2)]$, $wR(F^2)$, $S$ | 0.023, 0.059, 1.08 |
| No. of reflections | 2313 |
| No. of parameters | 147 |
| H-atom treatment | H-atom parameters constrained |
| $\Gamma_{max}$, $\Gamma_{min}$ (e Å$^{-3}$) | 0.13, -0.15 |
| Absolute structure | Flack x determined using 917 quotients [(I+)-(I-)]/[(I+)+(I-)]   (Parsons, Flack and Wagner, Acta Cryst. B69 (2013) 249-259). |
| Absolute structure parameter | -0.03 (6) |

^{1}H, ^{13}C, and ^{19}F NMR spectra of the compounds and computational details can be found in the Supporting Information of the published paper.

## 2.8 References for Chapter 2

References for Sections 2.1-2.6:

1. Lebel, H.; Marcoux, J.-F.; Molinaro, C.; Charette, A. B. Stereoselective Cyclopropanation Reactions. *Chem. Rev.* **2003,** *103*, 977–1050.

2. Wessjohann, L. A.; Brandt, W.; Thiemann, T. Biosynthesis and Metabolism of Cyclopropane Rings in Natural Compounds. *Chem. Rev.* **2003,** *103*, 1625–1647.

3. Talele, T. T. The "Cyclopropyl Fragment" is a Versatile Player that Frequently Appears in Preclinical/Clinical Drug Molecules. *J. Med. Chem.* **2016,** *59*, 8712–8756.

4. Ebner, C.; Carreira, E. M. Cyclopropanation Strategies in Recent Total Syntheses. *Chem. Rev.* **2017,** *117*, 11651–11679.

5. Cohen, Y.; Cohen, A.; Marek, I. Creating Stereocenters within Acyclic Systems by C–C Bond Cleavage of Cyclopropanes. *Chem. Rev.* **2021,** *121*, 140–161.

6. McDonald, T. R.; Mills, L. R.; West, M. S.; Rousseaux, S. A. L. Selective Carbon–Carbon Bond Cleavage of Cyclopropanols. *Chem. Rev.* **2021,** *121*, 3–79.

7. Pellissier, H. Recent Developments in Asymmetric Cyclopropanation. *Tetrahedron* **2008,** *64*, 7041–7095.

8. Lautens, M.; Klute, W.; Tam, W. Transition Metal-Mediated Cycloaddition Reactions. *Chem. Rev.* **1996,** *96*, 49–92.

9. Ma, S.; Mandalapu, D.; Wang, S.; Zhang, Q. Biosynthesis of Cyclopropane in Natural Products. *Nat. Prod. Rep.* **2022,** *39*, 926–945.

10. Simmons, H. E.; Smith, R. D. A New Synthesis of Cyclopropanes from Olefins. *J. Am. Chem. Soc.* **1958,** *80*, 5323–5324.

11. Simmons, H. E.; Smith, R. D. A New Synthesis of Cyclopropanes. *J. Am. Chem. Soc.* **1959,** *81*, 4256–4264.

12. Doyle, M. P.; Forbes, D. C. Recent Advances in Asymmetric Catalytic Metal Carbene Transformations. *Chem. Rev.* **1998,** *98*, 911–936.

13. Davies, H. M. L.; Antoulinakis, E. G. Intermolecular Metal-Catalyzed Carbenoid Cyclopropanations. *Org. React.* **2004,** *57*, 1–326.

14. Doyle, M. P. Electrophilic Metal Carbenes as Reaction Intermediates in Catalytic Reactions. *Acc. Chem. Res.* **1986,** *19*, 348–356.

15. Williams, J. M. J. *Preparation of Alkenes: A Practical Approach*. Oxford University Press: Oxford, UK, 1996.

16. Nevesely, T.; Wienhold, M.; Molloy, J. J.; Gilmour, R. Advances in the *E–Z* Isomerization of Alkenes Using Small Molecule Photocatalysts. *Chem. Rev.* **2022,** *122*, 2650–2694.

17. Chandgude, A. L.; Fasan, R. Highly Diastereo- and Enantioselective Synthesis of Nitrile-Substituted Cyclopropanes by Myoglobin-Mediated Carbene Transfer Catalysis. *Angew. Chem., Int. Ed.* **2018,** *57*, 15852–15856.

18. Benkovic, S. J.; Hammes-Schiffer, S. A Perspective on Enzyme Catalysis. *Science* **2003,** *301*, 1196–1202.

19. Chen, K.; Arnold, F. H. Engineering New Catalytic Activities in Enzymes. *Nat. Catal.* **2020,** *3*, 203–213.

20. Yang, Y.; Arnold, F. H. Navigating the Unnatural Reaction Space: Directed Evolution of Heme Proteins for Selective Carbene and Nitrene Transfer. *Acc. Chem. Res.* **2021,** *54*, 1209–1225.

21. Coelho, P. S.; Brustad, E. M.; Kannan, A.; Arnold, F. H. Olefin Cyclopropanation via Carbene Transfer Catalyzed by Engineered Cytochrome P450 Enzymes. *Science* **2012,** *339*, 307–310.

22. Poulos, T. L. Cytochrome P450 Flexibility. *Proc. Natl. Acad. Sci.* **2003,** *100*, 13121–13122.

23. Brandenberg, O. F.; Fasan, R.; Arnold, F. H. Exploiting and Engineering Hemoproteins for Abiological Carbene and Nitrene Transfer Reactions. *Curr. Opin. Biotechnol.* **2017,** *47*, 102–111.

24. Kaur, P.; Tyagi, V. Recent Advances in Iron-Catalyzed Chemical and Enzymatic Carbene-Transfer Reactions. *Adv. Synth. Catal.* **2021,** *363*, 877–905.

25. Zetzsche, L. E.; Narayan, A. R. H. Broadening the Scope of Biocatalytic C–C Bond Formation. *Nat. Rev. Chem.* **2020,** *4*, 334–346.

26. Roelfes, G. Repurposed and Artificial Heme Enzymes for Cyclopropanation Reactions. *J. Inorg. Biochem.* **2021,** *222*, 111523.

27. Key, H. M.; Dydio, P.; Clark, D. S.; Hartwig, J. F. Abiological Catalysis by Artificial Haem Proteins Containing Noble Metals in Place of Iron. *Nature* **2016,** *534*, 534–537.

28. Kulinkovich, O. G. The Chemistry of Cyclopropanols. *Chem. Rev.* **2003,** *103*, 2597–2632.

29. Liu, Q.; You, B.; Xie, G.; Wang, X. Developments in the Construction of Cyclopropanols. *Org. Biomol. Chem.* **2020,** *18*, 191–204.

30. Yang, Y.; Cho, I.; Qi, X.; Liu, P.; Arnold, F. H. An Enzymatic Platform for the Asymmetric Amination of Primary, Secondary and Tertiary C($sp^3$)–H Bonds. *Nat. Chem.* **2019,** *11*, 987–993.

31. Prier, C. K.; Zhang, R. K.; Buller, A. R.; Brinkmann-Chen, S.; Arnold, F. H. Enantioselective, Intermolecular Benzylic C–H Amination Catalysed by an Engineered Iron-Haem Enzyme. *Nat. Chem.* **2017,** *9*, 629–634.

32. Zhang, R. K.; Chen, K.; Huang, X.; Wohlschlager, L.; Renata, H.; Arnold, F. H. Enzymatic Assembly of Carbon–Carbon Bonds via Iron-Catalysed $sp^3$ C–H Functionalization. *Nature* **2019,** *565*, 67–72.

33. Kille, S.; Acevedo-Rocha, C. G.; Parra, L. P.; Zhang, Z. G.; Opperman, D. J.; Reetz, M. T.; Acevedo, J. P. Reducing Codon Redundancy and Screening Effort of Combinatorial Protein Libraries Created by Saturation Mutagenesis. *ACS Synth. Biol.* **2013,** *2*, 83–92.

34. Reetz, M. T.; Carballeira, J. D. Iterative Saturation Mutagenesis (ISM) for Rapid Directed Evolution of Functional Enzymes. *Nat. Protoc.* **2007,** *2*, 891–903.

35. Mykhailiuk, P. K.; Koenigs, R. M. Diazoacetonitrile ($N_2$CHCN): A Long Forgotten but Valuable Reagent for Organic Synthesis. *Eur. J. Chem.* **2020,** *26*, 89–101.

36. Wei, Y.; Tinoco, A.; Steck, V.; Fasan, R.; Zhang, Y. Cyclopropanations via Heme Carbenes: Basic Mechanism and Effects of Carbene Substituent, Protein Axial Ligand, and Porphyrin Substitution. *J. Am. Chem. Soc.* **2018,** *140*, 1649–1662.

37. Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988,** *37*, 785–789.

38. Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993,** *98*, 5648–5652.

39. Barone, V.; Cossi, M. Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model. *J. Phys. Chem. A* **1998,** *102*, 1995–2001.

40. Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005,** *7*, 3297–3305.

41. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate *ab initio* Parametrization of Density Functional Dispersion Correction (DFT-D) For the 94 Elements H-Pu. *J. Chem. Phys.* **2010,** *132*, 154104.

42. Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011,** *32*, 1456–1465.

43. Sharon, D. A.; Mallick, D.; Wang, B.; Shaik, S. Computation Sheds Insight into Iron Porphyrin Carbenes' Electronic Structure, Formation, and N–H Insertion Reactivity. *J. Am. Chem. Soc.* **2016,** *138*, 9597–9610.

44. Meunier, B.; Visser, S. P. d.; Shaik, S. Mechanism of Oxidation Reactions Catalyzed by Cytochrome P450 Enzymes. *Chem. Rev.* **2004,** *104*, 3947–3980.

45. Shaik, S.; Hirao, H.; Kumar, D. Reactivity of High-Valent Iron–Oxo Species in Enzymes and Synthetic Reagents: A Tale of Many States. *Acc. Chem. Res.* **2007**, *40*, 532–542.

46. Dzik, W. I.; Xu, X.; Zhang, X. P.; Reek, J. N. H.; de Bruin, B. 'Carbene Radicals' in Co$^{II}$(por)-Catalyzed Olefin Cyclopropanation. *J. Am. Chem. Soc.* **2010,** *132*, 10891–10902.

47. Lu, H.; Dzik, W. I.; Xu, X.; Wojtas, L.; de Bruin, B.; Zhang, X. P. Experimental Evidence for Cobalt(III)-Carbene Radicals: Key Intermediates in Cobalt(II)-Based Metalloradical Cyclopropanation. *J. Am. Chem. Soc.* **2011,** *133*, 8518–8521.

48. Carminati, D. M.; Fasan, R. Stereoselective Cyclopropanation of Electron-Deficient Olefins with a Cofactor Redesigned Carbene Transferase Featuring Radical Reactivity. *ACS Catal.* **2019,** *9*, 9683–9697.

References for Section 2.7:

1. Yang, Y.; Cho, I.; Qi, X.; Liu, P.; Arnold, F. H. An Enzymatic Platform for the Asymmetric Amination of Primary, Secondary and Tertiary C($sp^3$)–H Bonds. *Nat. Chem.* **2019,** *11*, 987–993.

2. Gibson, D. G.; Young, L.; Chuang, R. Y.; Venter, J. C.; Hutchison, C. A.; Smith, H. O. Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases. *Nat. Methods* **2009,** *6*, 343–345.

3. Kille, S.; Acevedo-Rocha, C. G.; Parra, L. P.; Zhang, Z. G.; Opperman, D. J.; Reetz, M. T.; Acevedo, J. P. Reducing Codon Redundancy and Screening Effort of Combinatorial Protein Libraries Created by Saturation Mutagenesis. *ACS Synth. Biol.* **2013,** *2*, 83–92.

4. A. Berry, E.; L.Trumpower, B. Simultaneous Determination of Hemes a, b, and c from Pyridine Hemochrome Spectra. *Anal. Biochem.* **1987,** *161*, 1-15.

5. Dong, C.; Liu, D.-S.; Zhang, L.; Hu, X.-P. Rh-Catalyzed Asymmetric Hydrogenation of *α*-Aryl-*β*-Alkylvinyl Esters With Chiral Ferrocenylphosphine-Phosphoramidite Ligand. *Tetrahedron Lett.* **2021,** *65*, 152763.

6. Basdevant, B.; Legault, C. Y. Study of the Reactivity of [Hydroxy(Tosyloxy)Iodo]Benzene Toward Enol Esters to Access *alpha*-Tosyloxy Ketones. *J. Org. Chem.* **2015,** *80*, 6897–6902.

7. Singh, K.; Staig, S. J.; Weaver, J. D. Facile Synthesis of *Z*-Alkenes via Uphill Catalysis. *J. Am. Chem. Soc.* **2014,** *136*, 5275–5278.

8. Hock, K. J.; Knorrscheidt, A.; Hommelsheim, R.; Ho, J.; Weissenborn, M. J.; Koenigs, R. M. Tryptamine Synthesis by Iron Porphyrin Catalyzed C–H Functionalization of Indoles with Diazoacetonitrile. *Angew. Chem., Int. Ed.* **2019,** *58*, 3630–3634.

9.  Gomez, J. E.; Guo, W.; Gaspa, S.; Kleij, A. W. Copper-Catalyzed Synthesis of *gamma*-Amino Acids Featuring Quaternary Stereocenters. *Angew. Chem., Int. Ed.* **2017,** *56*, 15035–15038.

10. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16 Rev. A.03*, Wallingford, CT, 2016.

11. Grimme, S. Exploration of Chemical Compound, Conformer, and Reaction Space with Meta-Dynamics Simulations Based on Tight-Binding Quantum Chemical Calculations. *J. Chem. Theory. Comput.* **2019,** *15*, 2847–2862.

12. Pracht, P.; Bohle, F.; Grimme, S. Automated Exploration of the Low-Energy Chemical Space With Fast Quantum Chemical Methods. *Phys. Chem. Chem. Phys.* **2020,** *22*, 7169–7192.

13. Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula Into a Functional of the Electron Density. *Phys. Rev. B* **1988,** *37*, 785–789.

14. Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993,** *98*, 5648-5652.

15. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate ab initio Parametrization of Density Functional Dispersion Correction (DFT-D) For the 94 Elements H-Pu. *J. Chem. Phys.* **2010,** *132*, 154104.

16. Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011,** *32*, 1456–1465.

17. Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to RN: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005,** *7*, 3297–3305.

18. Barone, V.; Cossi, M. Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model. *J. Phys. Chem. A* **1998,** *102*, 1995–2001.

19. Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. Energies, Structures, and Electronic Properties of Molecules in Solution With the C-PCM Solvation Model. *J. Comput. Chem.* **2003,** *24*, 669-681.

20. Luchini, G.; Alegre-Requena, J. V.; Funes-Ardoiz, I.; Paton, R. S. GoodVibes: Automated Thermochemistry for Heterogeneous Computational Chemistry Data. *F1000Research* **2020,** *9*.

21. Grimme, S. Supramolecular Binding Thermodynamics by Dispersion-Corrected Density Functional Theory. *Eur. J. Chem.* **2012,** *18*, 9955–9964.

22. Li, Y.-P.; Gomes, J.; Mallikarjun Sharada, S.; Bell, A. T.; Head-Gordon, M. Improved Force-Field Parameters for QM/MM Simulations of the Energies of Adsorption for Molecules in Zeolites and a Free Rotor Correction to the Rigid Rotor Harmonic Oscillator Model for Adsorption Enthalpies. *J. Phys. Chem. C .* **2015,** *119*, 1840–1850.

23. Legault, C. Y. *CYLview20.* Université de Sherbrooke, 2020.

24. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A. P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T. SWISS-MODEL: Homology Modelling of Protein Structures and Complexes. *Nucleic Acids Res.* **2018,** *46*, 296–303.

25. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Zidek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction With Alphafold. *Nature* **2021,** *596*, 583–589.

26. Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009,** *30*, 2785–2791.

27. Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking With a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2010,** *31*, 455–461.

28. D. A. Case, R. M. B., D. S. Cerutti, T. E. Cheatham, I., T. A. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T. S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K. M. Merz, G. Monard, H. Nguyen, H. T. Nguyen, I. Omelyan, A. Onufriev, D. R. Roe, A. Roitberg, C. Sagui, C. L. Simmerling, W. M. Botello-Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, L. Xiao & Kollman, P. A. *AMBER 2016.* University of California, San Francisco, 2016.

29. Case, D. A., Aktulga, H. M., Belfon, K., Ben-Shalom, I. Y., Brozell, S. R., Cerutti, D. S., Cheatham, T. E., III, Cisneros, G. A., Cruzeiro, V. W. D., Darden, T. A., Duke, R. E., Giambasu, G., Gilson, M. K., Gohlke, H., Goetz, A. W., Harris, R., Izadi, S., Izmailov, S. A., Jin, C., Kasavajhala, K., Kaymak, M. C., King, E., Kovalenko, A., Kurtzman, T., Lee, T. S., LeGrand, S., Li, P., Lin, C., Liu, J., Luchko, T., Luo, R., Machado, M., Man, V., Manathunga, M., Merz, K. M., Miao, Y., Mikhailovskii, O., Monard, G., Nguyen, H., O'Hearn, K. A., Onufriev, A., Pan, F., Pantano, S., Qi, R., Rahnamoun, A., Roe, D. R., Roitberg, A., Sagui, C., Schott-Verdugo, S., Shen, J., Simmerling, C. L., Skrynnikov, N. R., Smith, J., Swails, J.,

Walker, R. C., Wang, J., Wei, H., Wolf, R. M., Wu, X., Xue, Y., York, D. M., Zhao, S. & Kollman, P. A. *Amber 2021.* University of California, San Francisco, 2021.

30. *H++ Server* Virginia Tech, http://newbiophysics.cs.vt.edu/H++/index.php.

31. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory. Comput.* **2015,** *11*, 3696–3713.

32. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983,** *79*, 926–935.

33. Bayly, C. I.; Cieplak, P.; Wendy D, C.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993,** *97*, 10269–10280.

34. Besler, B. H.; Merz, K. M.; Kollman, P. A. Atomic Charges Derived From Semiempirical Methods. *J. Comput. Chem.* **1990,** *11*, 431–439.

35. Singh, U. C.; Kollman, P. A. An Approach to Computing Electrostatic Charges for Molecules. *J. Comput. Chem.* **1984,** *5*, 129–145.

36. Li, P.; Merz, K. M., Jr. MCPB.py: A Python Based Metal Center Parameter Builder. *J. Chem. Inf. Model* **2016,** *56*, 599–604.

37. M. J. Frisch, G. W. T., H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar,

J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J. Fox. Gaussian 09, Revision D.01. Gaussian, Inc., 2013.

38. Salomon-Ferrer, R.; Gotz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory. Comput.* **2013,** *9*, 3878–3888.

39. Shaw, D. E.; Grossman, J. P.; Bank, J. A.; Batson, B.; Butts, J. A.; Chao, J. C.; Deneroff, M. M.; Dror, R. O.; Even, A.; Fenton, C. H.; Forte, A.; Gagliardo, J.; Gill, G.; Greskamp, B.; Ho, C. R.; Ierardi, D. J.; Iserovich, L.; Kuskin, J. S.; Larson, R. H.; Layman, T.; Lee, L.-S.; Lerer, A. K.; Li, C.; Killebrew, D.; Mackenzie, K. M.; Mok, S. Y.-H.; Moraes, M. A.; Mueller, R.; Nociolo, L. J.; Peticolas, J. L.; Quan, T.; Ramot, D.; Salmon, J. K.; Scarpazza, D. P.; Schafer, U. B.; Siddique, N.; Snyder, C. W.; Spengler, J.; Tang, P. T. P.; Theobald, M.; Toma, H.; Towles, B.; Vitale, B.; Wang, S. C.; Young, C. Anton 2: Raising the Bar for Performance and Programmability in a Special-Purpose Molecular Dynamics Supercomputer. *IEEE.* **2014**, 41–53.

40. *The PyMOL Molecular Graphics System, Version 1.8.x.* Schrödinger, LLC.

41. Sheldrick, G. M. Phase Annealing in SHELX-90: Direct Methods for Larger Structures. *Acta Cryst.* **1990,** *A46*, 467–473.

42. Sheldrick, G. M. Crystal Structure Refinement With SHELXL. *Acta Cryst.* **2015,** *C71*, 3–8.

43. Müller, P. Practical Suggestions for Better Crystal Structures. *Crystallogr. Rev.* **2009,** *15*, 57–83.

*C h a p t e r   3*

# BIOCATALYTIC, STEREOCONVERGENT ALKYLATION OF (*Z/E*)-

# TRISUBSTITUTED SILYL ENOL ETHERS

Content in this chapter is adapted from published work:

Mao, R.[†]; Taylor, D. M.[†]; Wackelin, D. J.[†]; Rogge, T.; Wu, S. J.; Sicinski, K. S.; Arnold, F. H. Biocatalytic, Stereoconvergent Alkylation of (*Z/E*)-Trisubstituted Silyl Enol Ethers. *Manuscript submitted for publication*. **2023**.

R.M. conceptualized and designed the overall project under the guidance of F.H.A.. R.M. and D.J.W. carried out the initial screening of heme proteins. D.M.T. performed the directed evolution experiments, with support from S.J.W.. R.M. and D.J.W. investigated the substrate scope and reaction mechanism. K.M.S. purified and obtained both (Z)-1a and (E)-1a. T.R. carried out the computational studies with K.N.H providing guidance. R.M. and F.H.A wrote the manuscript with input from all authors. R.M., D.M.T., and D.J.W. contributed equally to this work.

**Abstract**

Alkene functionalization has garnered significant attention due to the versatile reactivity of C=C bonds. A major challenge is the selective conversion of Z/E mixtures of alkenes into chiral products. Researchers have devised various biocatalytic strategies to transform isomeric alkenes into stereopure compounds; while selective, the enzymes often specifically convert one alkene isomer, thereby limiting overall yield. To increase the overall yield, scientists have introduced additional driving forces to interconvert alkene isomers, often at the cost of increasing energy consumption and chemical waste. Developing an enzyme for stereoconvergent alkene functionalization would offer an ideal solution, although such catalysts are rarely reported. Here we present engineered hemoproteins derived from a bacterial cytochrome P450 that efficiently catalyze α-carbonyl alkylation of isomeric silyl enol ethers, producing stereopure products. Through screening and directed evolution, we generated P450BM3 variant P411-SCA-5188, which catalyzes stereoconvergent carbene transfer in Escherichia coli with high efficiency and stereoselectivity to various Z/E mixtures of silyl enol ethers. In contrast to established stereospecific transformations that leave one isomer unreacted, P411-SCA-5188 converts both isomers to a stereopure product. This biocatalytic approach simplifies the synthesis of chiral α-branched ketones by eliminating the need for stoichiometric chiral auxiliaries, strongly basic alkali-metal enolates, and harsh conditions, delivering products with high efficiency and excellent chemo- and stereoselectivities.

## 3.1 Introduction

The use of enzymes to produce valuable fine chemicals, pharmaceuticals, and agrochemicals is at the forefront of the drive toward mild and ecologically sustainable synthesis techniques.[1-6] This can be largely attributed to the unparalleled chemo-, regio-, and stereoselectivities achievable through biotransformations.[5] Alkene biocatalytic functionalization, in particular, has garnered interest due to its potential to produce valuable chemicals from low-cost resources.[7] Enzyme-catalyzed alkene functionalizations offer advantages over their small molecule-catalyzed counterparts that include superior efficiency, unmatched selectivity, and the absence of toxic metals or reagents.[7-11] Nevertheless, enzymes often convert one alkene isomer while leaving the other unaltered (see a representative example in 0A[12] among others[13]) or they catalyze the formation of distinct products based on the geometric properties of the alkenes.[14] Under certain conditions, enzymes may even display comparable conversion rates for both (*Z*)- and (*E*)-alkenes, resulting in the synthesis of identical products with inverse configurations (see a representative example in 0B[15] among others[16]). Because the synthesis of geometrically pure alkenes can be very challenging, separation of the alkene isomers is often necessary to achieve high efficiency and stereoselectivity in biocatalytic processes.[17,18] To overcome this issue, in 2019 the Hartwig and Zhao groups pioneered a cooperative chemoenzymatic strategy for stereoconvergent reduction of *Z/E*-mixed alkenes that merges a photocatalytic cycle with an ene-reductase (0C).[19] Under blue light irradiation, the unreactive *Z* isomer could be converted to the active *E* isomer, which allowed an overall stereoconvergent reduction. This

cooperative catalytic system is efficient and selective, but also requires additional inputs, such as blue-light irradiation and an expensive photocatalyst.

Arguably the most desirable transformation would use a stereoconvergent catalyst to directly convert a mixture of *Z/E* alkenes into a single product, with a theoretical yield of 100%. Although stereoconvergent catalysts have been studied extensively for small-molecule catalysis, stereoconvergent enzymes are seldom reported.[13,20-23] A stereoconvergent enzyme's active site must accommodate each stereoisomer, and in a unique pose, in order to have convergent selectivity toward a single product. In 2019, members of our group identified an engineered cytochrome P411 (a serine-ligated cytochrome P450 variant), designated P411$_{Diane3}$, that catalyzes the enantioconvergent intramolecular amination of racemic tertiary benzylic C($sp^3$)–H bonds with sulfamoyl azides.[24] Very recently, the group also reported a diastereomer-differentiating enzyme, P411-INC-5186, which catalyzes the transformation of (*E*)-enol acetates to *α*-branched ketones with high levels of enantioselectivity, while simultaneously cyclopropanating the (*Z*)-enol acetates with excellent activities and selectivities (0D).[25] Although different products were generated from (*Z*)- and (*E*)-enol acetates, the *β*-carbons of the products of the (*Z*)- and (*E*)-enol acetates yielded an (*S*)-configured stereocenter. Inspired by this, we wondered whether this diastereomer-differentiating reaction could be directed toward a stereoconvergent biotransformation (0E).

We focused on the synthesis of *α*-branched ketones, a class of compounds prevalent in both natural and synthetic bioactive molecules.[26] From a synthetic perspective, stereoselective

alkylation of alkali-metal enolates is widely employed for constructing $\alpha$-branched ketone moieties.[26-28] Extensive studies have been conducted on controlling the stereochemistry of the $\alpha$-carbon using various stoichiometric chiral auxiliaries and chiral catalysts.[28-30] The strong basicity of alkali-metal enolates, however, presents challenges in handling and limits the choice of functional groups. Additionally, the formation of distinct metal enolate isomers often complicates the stereochemical outcome of the subsequent reaction, as the product configuration is influenced by the geometry of the alkali-metal enolates. Furthermore, inherent difficulties in alkali-metal enolate alkylation, such as competitive *C*- and *O*-alkylation, overalkylation, and racemization, are non-trivial concerns.[28]

In this study, we showcase the adaptation and evolution of a biocatalytic platform initially designed for diastereomer-differentiating reactions to efficiently catalyze stereoconvergent alkylation. The evolved biocatalyst enables the enantioselective conversion of (*Z/E*)-mixed trisubstituted silyl enol ethers into chiral $\alpha$-branched ketones. The biocatalytic method broadens the limited scope of enzymatic stereoconvergent transformations and complements traditional synthetic techniques, owing to its mild reaction conditions and excellent efficiency and selectivity.
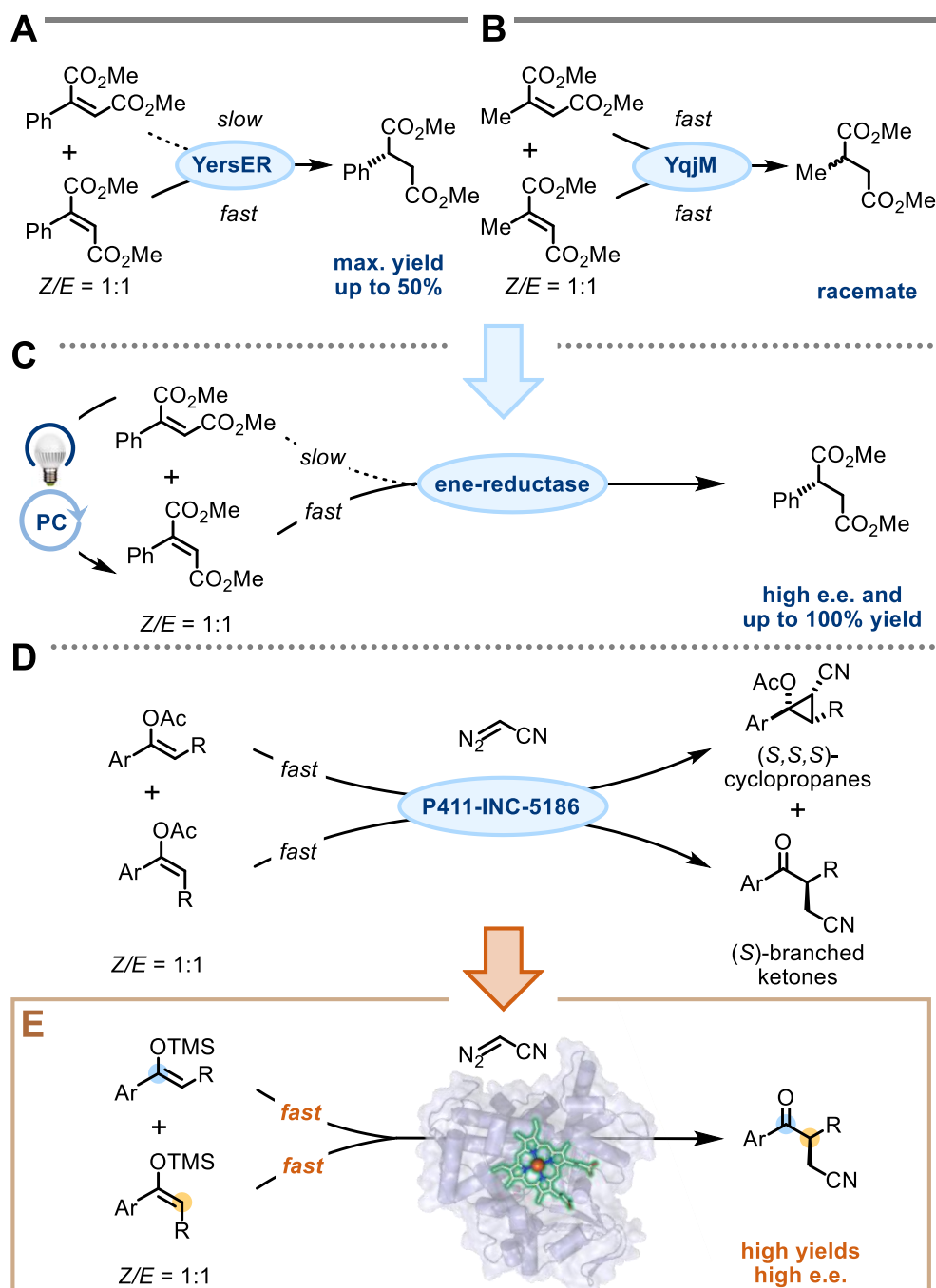
**Figure 3-1. Existing stereospecific transformations of (Z/E) alkenes**. (A) Stereospecific

biotransformation of *Z/E* alkene mixtures, exemplified by the asymmetric bioreduction by YersER

from *Yersinia bercovieri*,[12] an ene-reductase from the Old Yellow Enzyme family, requires isomerically pure alkenes to achieve high yields. (B) Substrate-controlled biotransformation of *Z/E* alkene mixtures, exemplified by the asymmetric bioreduction by YqjM from *Bacillus subtilis*,[15] an ene-reductase from the Old Yellow Enzyme family, converts *Z* and *E* isomers into products exhibiting opposite configurations. (C) A cooperative chemoenzymatic system developed by the Hartwig and Zhao groups enables the effective stereoconvergent reduction of *Z/E* mixtures of alkenes. This method uses blue-light irradiation and a photocatalyst to convert the unreactive *Z* isomer to the active *E* isomer.[19] (D) A diastereomer-differentiating biotransformation method: carbene transferase P411-INC-5186 catalyzes the cyclopropanation of (*Z*)-enol acetates while simultaneously catalyzing (*E*)-enol acetates to *α*-branched ketones, both products containing an (*S*)-configured stereocenter. (E) This work: biocatalytic stereoconvergent alkylation of (*Z/E*)-trisubstituted silyl enol ethers to produce chiral *α*-branched ketones with high yields in a stereoconvergent manner. The structural illustration is adapted from Protein Data Bank (PDB) ID 5UCW (cytochrome P450$_{BM3}$ variant **E10**).[31] EWG, electron-withdrawing groups; R, organic groups; Ac, acetyl group; TMS, trimethylsilyl group.

## 3.2 Initial screening and directed evolution of stereoconvergent alkyl transferase P411-SCA-5188

We commenced this investigation by focusing on the coupling between a 1:1 *Z/E* mixture of butyrophenone-derived silyl enol ether **1a** and diazoacetonitrile **2a** (0). Silyl enol ethers serve as mild, stable enolate equivalents in comparison to alkali-metal enolates, offering ease of synthesis and handling. Compared to enol acetates previously used in P411-INC-5186-catalyzed cyclopropanation reactions,[25] silyl-protecting groups demonstrate increased propensity for hydrolysis, thereby promoting ketone product formation. Additionally, the

inherent geometric information embedded in silyl enol ethers facilitates the development

of the stereoconvergent capacity of the enzyme. Lastly, diazoacetonitrile **2a** was selected due

to its potent electron-withdrawing nature, minimal steric interference, and the importance of

the nitrile functional group.[32-33] We tested whether variants in the P411-INC-5186 directed

evolution lineage could produce $\alpha$-branched ketone **3a** through the coupling of **1a** and **2a**.

Encouragingly, an evaluation of P411-INC-5186's precursors (Figure 3-2A, P411-INC-5183

to P411-INC-5186) indicated that all variants are capable of catalyzing the transfer of

diazoacetonitrile **2a** to silyl enol ether **1a**, with an activity trend matching that of the

cyclopropanation of enol acetates.[25] P411-INC-5186 demonstrated the highest activity and

selectivity (200 TTN; 42% yield and 97% e.e. of **3a**, Figure 3-2A). When ethyl diazoacetate

(EDA) was used as a carbene precursor, no alkylated product was detected (see

Supplementary Table 1). Although P411-INC-5186 can catalyze this new biotransformation

with remarkable enantioselectivity, its activity required further improvement. Thus, we

subjected P411-INC-5186 to sequential rounds of site-saturation mutagenesis (SSM) and

screening, focusing on key active-site residues previously shown to affect carbene and

nitrene transfer reactions.[34] Mutagenesis and screening introduced mutations V437A and

L177M (Figures 3-2A, 3-2B and 3-2C), yielding variant P411-SCA-5188, which catalyzes

the formation of **3a** with 290 TTN, 57% yield, and 99% e.e. (Figure 3-2A). Notably, the yield

of alkylated product **3a** (57%) exceeded the proportion of either isomer in the 1:1 *Z/E* mixture

of silyl enol ether **1a** (*Z*- and *E*-isomers constituting 50% each), strongly suggesting a

stereoconvergent process. This biocatalytic stereoconvergence stands out from other

reported olefin functionalizations, which predominantly adhere to stereospecific

pathways.[12,14-16]

We next fine-tuned the reaction conditions to improve the yield (Figure 3-2A; see Supplementary Tables 5–8 for details). After examining various concentrations of *E. coli* harboring P411-SCA-5188 ($OD_{600}$ values), concentrations of **1a** and **2a**, and pH values, we found that conducting the biotransformation using $OD_{600} = 90$, with 3 mM of **1a** and 37 mM of **2a** in M9-N aqueous buffer (pH 8) at room temperature afforded 85% yield with 99% e.e. (Figure 3-2A).
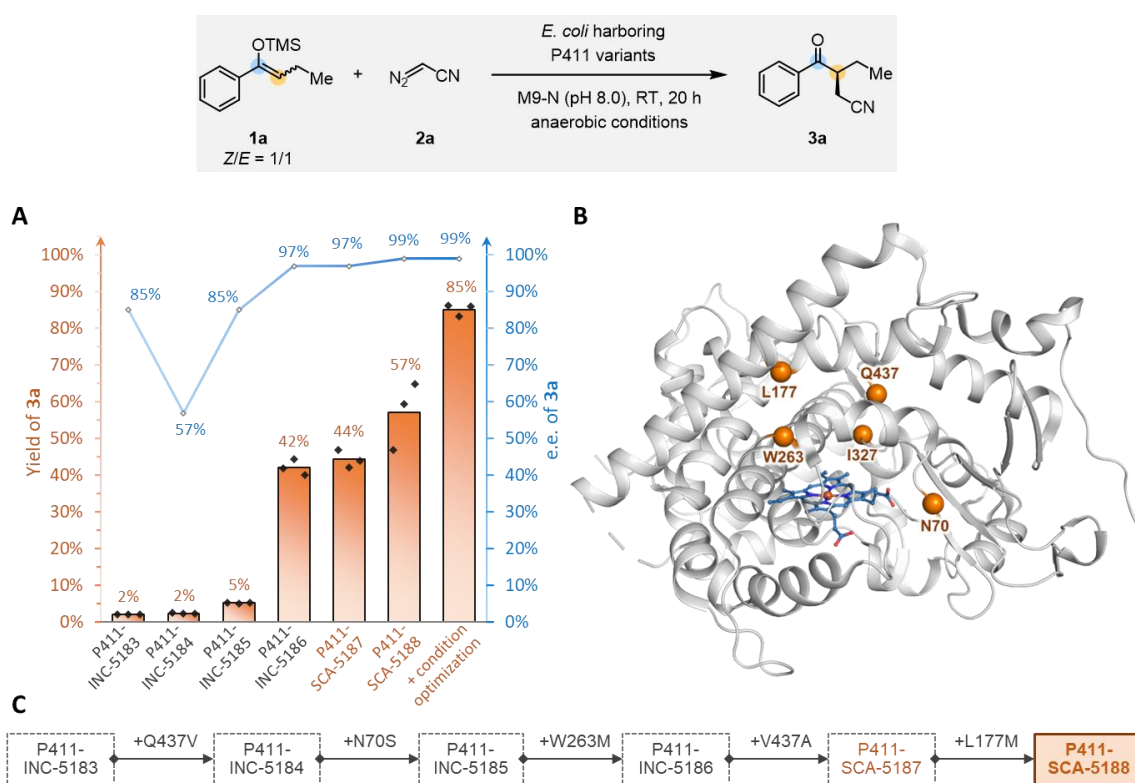


**Figure 3-2. Directed evolution for stereoconvergent alkylation.** Reaction conditions: 3 mM **1a**, 37 mM **2a**, *E. coli* whole cells harboring P411 variants ($OD_{600} = 30$) in M9-N aqueous buffer (pH 8.0), 10% v/v EtOH (co-solvent), room temperature, anaerobic conditions, 20 h. Optimized reaction conditions: 3 mM **1a**, 37 mM **2a**, *E. coli* whole cells harboring P411 variants ($OD_{600} = 90$) in M9-N aqueous buffer (pH 8.0), 10% v/v EtOH (co-solvent), room temperature, anaerobic conditions, 20 h.

(A) Assessment of the enzyme lineage encompassing P411-INC-5186 and directed evolution culminating in P411-SCA-5188. Yields were calculated from HPLC calibration curves and the average of triplicate experiments (n = 3). (B) The mutated residues (N70, L177, W263, I327, and Q437) that enhance activity or enantioselectivity are highlighted in the active site of closely related P411 variant **E10** (PDB ID: 5UCW).[31] (C) Summary of beneficial mutations for stereoconvergent alkylation.

## 3.3 Experimental studies of stereoconvergent alkyl transferase P411-SCA-5188

To confirm that this biotransformation is in fact stereoconvergent, we obtained the isolated **1a** stereoisomers, (*Z*)-**1a** (99% stereopurity) and (*E*)-**1a** (94% stereopurity) and used them individually as substrates for P411-SCA-5188-catalyzed *α*-carbonyl alkylation with diazoacetonitrile **2a** (Figure 3-3A). Under standard conditions, P411-SCA-5188 converted pure (*Z*)-**1a** into **3a** with an *S* configuration in 49% yield (99 TTN) and 99% e.e. and transformed (*E*)-**1a** into **3a** with the same *S* configuration in 91% yield (230 TTN) and 99% e.e. (Figures 3-3A and 3-3B). Subsequently, we examined how catalytic efficiency against pure (*Z*)-**1a** and (*E*)-**1a** changed across the lineage (see Supplementary Tables 3 and 4 for more detail). Variants P411-INC-5183 to P411-INC-5185 exclusively catalyzed the transformation of (*Z*)-**1a** to **3a**, while leaving (*E*)-**1a** untouched. These results are consistent with the cyclopropanation of (*Z/E*)-trisubstituted enol acetates recently reported by the Arnold lab.[25] Upon introducing the W263M mutation, variants P411-INC-5186, P411-SCA-5187 and P411-SCA-5188 converted both (*Z*)-**1a** and (*E*)-**1a** into **3a**, accompanied by increases in activity

Since silyl enol ethers may degrade in the presence of water, forming butyrophenone **4a** after hydrolysis, we investigated whether butyrophenone **4a** could be converted to the $\alpha$-branched ketone product **3a**. However, butyrophenone **4a** proved unreactive under standard conditions (Figure 3-3C), eliminating a potential conversion pathway from **4a** to **3a**. We then examined whether (*Z*)-**1a** and (*E*)-**1a** can possibly interconvert during the enzymatic process. Under standard conditions, but without **2a** and with a reduced reaction time of two hours to ensure sufficient residual **1a** for pre- and post-reaction stereopurity comparison, we observed no change in the stereopurity of *Z* or *E* isomers (Figures 3-3D and 3-3E).

Based on these findings and our previous computational studies on P411-INC-5186-catalyzed diastereomer-differentiating transformations,[25] we hypothesized the polar stepwise pathway for this alkylation depicted in Figure 3-3F. In this mechanism, diazoacetonitrile **2a** initially reacts with a hemoprotein to form electrophilic iron carbene **II**. The lone pair conjugation of the $\beta$-carbon of silyl enol ether **1a** endows it with nucleophilic characteristics, which facilitates the nucleophilic attack of (*Z*)-**1a** or (*E*)-**1a** onto **II**, resulting in the formation of intermediate (*Z*)-**III** or (*E*)-**III**. Crucially, during the attack of the $\beta$-carbon of either (*Z*)- or (*E*)-**1a** on **II**, the enzyme strictly controls the orientation of the ethyl group on the $\beta$-carbon of **1a** due to its proximity to the nucleophilic site and its more significant steric clash with electrophilic metal carbene **II** and surrounding residues. Consequently, regardless of whether (*Z*)- or (*E*)-**1a** is used as the substrate, the $\beta$-carbon on (*Z*)-**III** or (*E*)-**III** possesses the same *S* configuration, ensuring enantioconvergence at this carbon. Both intermediates (*Z*)-**III** or

(*E*)-**III** are then enthalpically favored to undergo hydrolysis due to the unstable nature of these intermediates, yielding **3a**. Because the configuration of **3a** is determined in the initial nucleophilic attack step, the overall reaction is thus stereoconvergent.
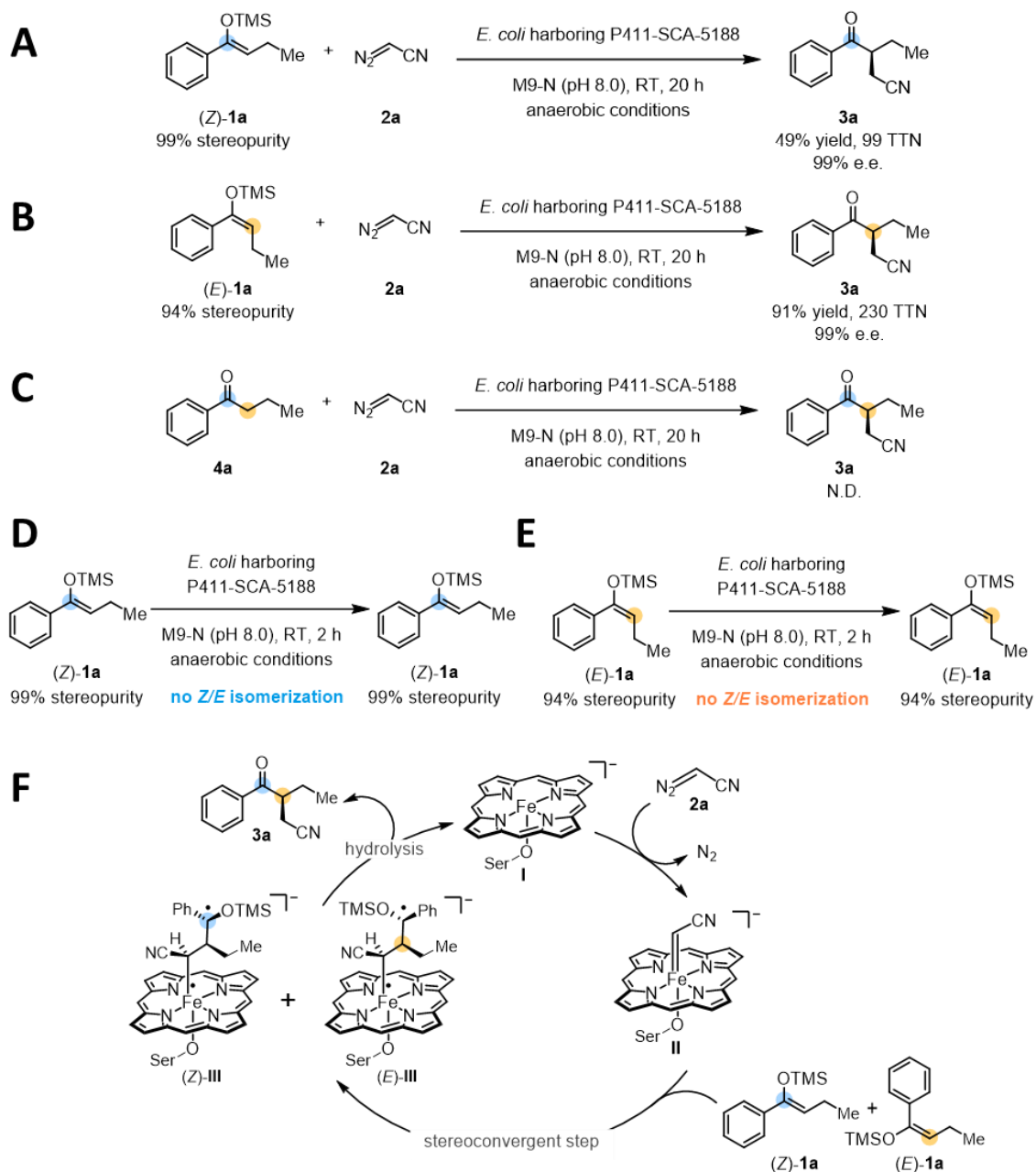
**Figure 3-3. Mechanistic studies of stereoconvergent enzymatic alkylation.** (A) and (B) Alkyl transfer reactions catalyzed by P411-SCA-5188 using (*Z*)- or (*E*)-**1a**. Reaction conditions: 3 mM (*Z*)-**1a** or (*E*)-**1a**, 37 mM **2a**, *E. coli* whole cells harboring P411-SCA-5188 (OD$_{600}$ = 90) in M9-N aqueous buffer (pH 8.0), 10% v/v EtOH (co-solvent), room temperature, anaerobic conditions, 20 h. (C) Alkyl transfer reactions catalyzed by P411-SCA-5188 using butyrophenone **4a**. Reaction conditions: 3 mM butyrophenone **4a**, 37 mM **2a**, *E. coli* whole cells harboring P411-SCA-5188 (OD$_{600}$ = 90) in M9-N aqueous buffer (pH 8.0), 10% v/v EtOH (co-solvent), room temperature, anaerobic conditions, 20 h. (D) and (E) Stereoretention of (*Z*)- or (*E*)-**1a**. Reaction conditions: 3 mM (*Z*)- or (*E*)-**1a**, *E. coli* whole cells harboring P411-SCA-5188 (OD$_{600}$ = 90) in M9-N aqueous buffer (pH 8.0), 10% v/v EtOH (co-solvent), room temperature, anaerobic conditions, 2 h. (F) Proposed mechanistic pathway of P411-SCA-5188-catalyzed stereoconvergent alkylation. Yields were calculated from HPLC calibration curves and the average of triplicate experiments (n = 3). N.D. = not detected; TMS, trimethylsilyl group.

## 3.4 Molecular dynamics simulations

To examine the mechanistic hypothesis further, molecular dynamics (MD) simulations were performed on intermediates (*E*)-**III** and (*Z*)-**III** in the active site of the enzyme. For (*E*)-**III**, both the cyano- and the ethyl-groups were found to orient towards the outward-facing side of the active site, with the CH$_3$-group of the ethyl-substituent angled in the direction of residue P329 (Figure 3-4A). In addition, the phenyl substituent is placed in a sterically accessible area at the top of the active site, defined by surrounding residues L75, M185, A264 and P329 (see also Supplementary Figures 3-5A and 3-6A for more details). The relatively bulky TMS group is preferentially oriented towards a non-congested area located at the backside of the active site in the vicinity of G268 and V328. Additionally, a rather short

distance between the silyl ether oxygen and V328 indicates the presence of stabilizing C–H--O hydrogen bond interactions. With intermediate (*Z*)-**III**, a comparable preferred orientation was observed, with the cyano and ethyl substituents again angled towards the front of the active site, allowing for the TMS and the phenyl substituent to be placed in the sterically accessible areas at the top and back of the active site, respectively (Figure 3-4B and Supplementary Figures 3-5B and 3-6B). In contrast, when the diastereomeric intermediate (*E*)-**IIIa**, which upon hydrolysis results in the formation of the minor enantiomer, was simulated, the large phenyl and TMS substituents are forced into less optimal positions (Figure 3-4C and Supplementary Figures 3-5C and 3-6C). In the preferred structure, the substrate's phenyl ring is placed in close proximity to A264, T438 and V328, resulting in a displacement of V328 and P329 and likely leading to a destabilization of (*E*)-**IIIa** (Supplementary Figures 3-5C and 3-6C). The significantly worse fit of (*E*)-**IIIa** in the active site is in good agreement with the experimental results and our mechanistic proposal. For the diastereomer of (*Z*)-**III**, (*Z*)-**IIIa**, the TMS group of the substrate was preferentially oriented towards the outward-facing side of the active site, with short distances between the TMS substituent and L75, T260 and A264 (Figure 3-4D and Supplementary Figures 3-5D and 3-6D). These short distances indicate unfavorable steric interactions and thus destabilization of this intermediate, reasonably resulting in the high enantioselectivity observed experimentally (Figure 3-3).
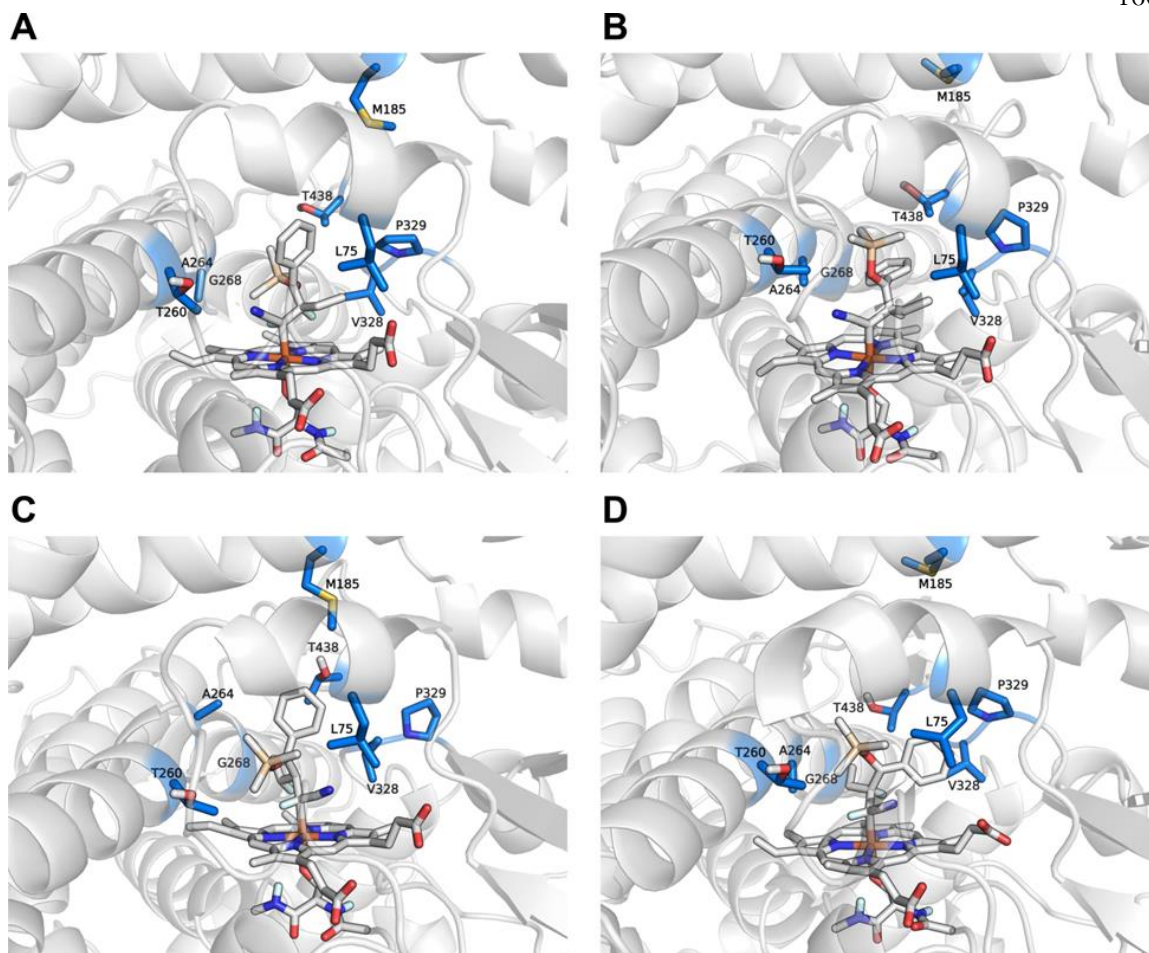
**Figure 3-4. Molecular dynamics simulations.** Representative structures of the most populated clusters obtained from MD simulations on (A) intermediate (*E*)-**III** and (B) intermediate (*Z*)-**III** and the corresponding diastereomers (C) (*E*)-**IIIa** and (D) (*Z*)-**IIIa**. Non-relevant, non-polar hydrogen atoms are omitted for clarity.

## 3.5 Substrate scope study

We next investigated the activity of P411-SCA-5188 on an array of *Z/E* mixtures of silyl enol ethers. Generally, P411-SCA-5188 catalyzes the alkylation of diverse *Z/E* mixtures of *α*-aryl, *β*-alkyl-substituted silyl enol ethers **1** with diazoacetonitrile **2a**, producing the desired

products **3** in moderate to high yields and with exceptional enantioselectivities (up to 91%

yield and >99% e.e., Figure 3-5), regardless of the stereopurity of the silyl enol ethers

employed. Substrates featuring various substituents on the *α*-aryl group exhibit excellent

compatibility with this biotransformation. Electron-donating, -neutral, and -withdrawing

substituents on the aromatic ring were all well-tolerated, generating *α*-branched ketones with

consistently high enantioselectivity levels (**3a**–**3m**, Figure 3-5). *Ortho-*, *para-* and *meta-*

substituted *α*-aryl silyl enol ethers reacted efficiently with diazoacetonitrile **2a**, yielding the

corresponding *α*-branched ketones (**3b**–**3i** and **3m**, Figure 3-5). The introduction of a bulky

group at the *para*-position also resulted in the *α*-branched ketone product with great

enantioselectivity (95% e.e.), albeit with a reduced yield (**3d**, Figure 3-5). Substrates

containing a halogen functional group, such as fluoro- (**1e** and **1f**, Figure 3-5), chloro- (**1g**),

and bromo- (**1h**), proved viable for generating alkylated products (**3e**–**3h**, Figure 3-5) with

synthetically useful yields and excellent enantioselectivities. The incorporation of an

electron-donating group also facilitated the synthesis of the *α*-branched ketone, yielding the

desired product with over 99% e.e. (**3i**, Figure 3-5). A structural variation made by replacing

the aryl ring with thiophene (**1j**, Figure 3-5) is also well-tolerated by the biocatalytic

stereoconvergent transformation, generating **3j** in 88% yield and with 99% e.e. (Figure 3-5).

Moreover, *β*-substituents with diverse alkyl chains (**1k**–**1m**, Figure 3-5) could be converted

to the corresponding *α*-branched ketones in satisfactory yields and with high

enantioselectivity. Notably, lengthening the *β*-alkyl chain appeared to have a detrimental

effect on activity (**3k**, Figure 3-5), likely due to steric clash of the alkyl chain with

surrounding residues. Conversely, shortening the *β*-alkyl chain of the silyl enol ethers from

propyl (**1k**, Figure 3-5) or ethyl (**1a**, Figure 3-5) to methyl (**1l**, Figure 3-5) enhanced both

activity and selectivity, resulting in a 91% yield and 99% e.e. for product **3l** (Figure 3-5).

Intrigued by this observation, we also examined the impact of the bulkiness of the silyl-protecting groups: reactivity was completely abolished when the trimethylsilyl group (TMS; **1a**, Figure 3-5) was replaced with a triethylsilyl group (TES; **1n**, Figure 3-5), a triisopropylsilyl group (TIPS; **1o**, Figure 3-5), or a *tert*-butyldimethylsilyl group (TBDMS; **1p**, Figure 3-5). We attributed this to the enhanced nucleophilicity and smaller steric hindrance provided by a TMS group, which allows the silyl enol ether to more effectively approach and interact with its reaction partner in the enzyme pocket. Notably, the synthesis of trimethylsilyl enol ethers is more cost-effective compared to other silyl enol ethers.
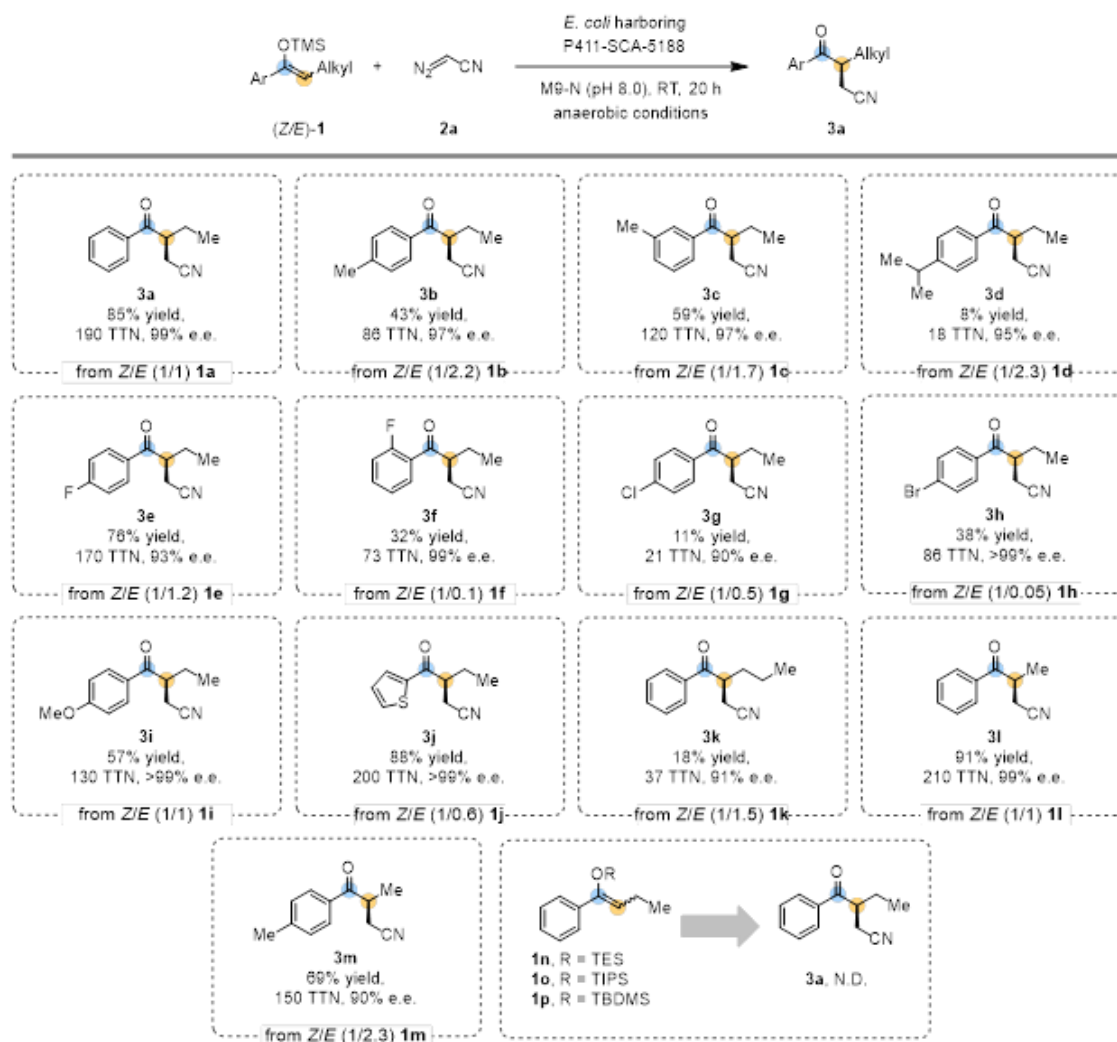
**Figure 3-5. Substrate scope study.** Reaction conditions: 3 mM **1**, 37 mM **2a**, *E. coli* whole cells harboring P411-SCA-5188 (OD$_{600}$ = 90) in M9-N aqueous buffer (pH 8.0), 10% v/v EtOH (co-solvent), room temperature, anaerobic conditions, 20 h. Yields were calculated from HPLC calibration curves and the average of triplicate experiments (n = 3). Ar, aryl groups; Alkyl, alkyl groups; R, organic groups; TMS, trimethylsilyl group; TES, triethylsilyl group; TIPS, triisopropylsilyl group; TBDMS, *tert*-butyldimethylsilyl group; N.D. = not detected.

# 3.6 Summary and conclusions

We have developed an enzyme, $P450_{BM3}$ variant P411-SCA-5188, which catalyzes stereoconvergent alkylation of isomeric silyl enol ethers, yielding high-value chiral $\alpha$-branched ketones. Compared to state-of-the-art methods for synthesizing chiral $\alpha$-branched ketones from alkali-metal enolates, this work presents a streamlined, biocatalytic approach that operates under mild conditions. This strategy complements stereospecific biocatalytic alkene functionalization methods and more complex cooperative chemoenzymatic alkene functionalization approaches. We anticipate these findings will inspire further development of stereoconvergent enzymes to broaden the scope of biocatalysis and unlock new biocatalytic strategies to address synthetic challenges.

# 3.7 Supplementary information for Chapter 3

## Supplementary Tables

**Table 3-S1**. Initial activity screening with engineered P411s.[a]
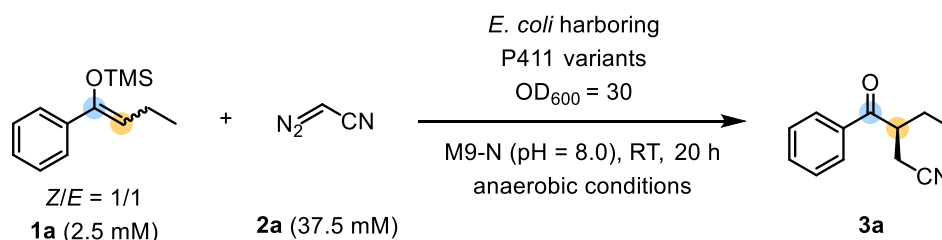


Four P411 variants, which were previously employed for the cyclopropanation of enol acetates,[1] were evaluated. Among these, P411-INC-5186, a truncated P411 variant devoid of the FAD domain, demonstrated the highest activity and enantioselectivity for synthesis of α-

branched ketone product **3a**.

| Entry | Variant | Yield of 3a[a] | Standard deviation of yield | TTN of 3a | Standard deviation of TTN |
|---|---|---|---|---|---|
| 1 | P411-INC-5183 (P411-INC-5182 I327P, Y263W) | 2.1% | 0.1% | 12 | 0.5 |
| 2 | P411-INC-5184 (P411-INC-5183 Q437V) | 2.4% | 0.2% | 17 | 1.1 |
| 3 | P411-INC-5185 (P411-INC-5184 N70S) | 5.2% | 0.1% | 45 | 0.7 |
| 4 | P411-INC-5186 (P411-INC-5185 W263M) | 42.1% | 2.1% | 202 | 10.2 |
| 5 | [b] hemin (20 μM) | 13% | N/A | 20 | N/A |
| 6 | [c] hemin (20 μM) + $Na_2S_2O_4$ | N.D. | N/A | N.D. | N/A |
| 7 | [c] hemin (20 μM) + $Na_2S_2O_4$ + BSA (20 μM) | N.D. | N/A | N.D. | N/A |
| 9 | [d] cellular background | N.D. | N/A | N.D. | N/A |
| 10 | [e] replace **2a** with ethyl diazoacetate (EDA) | N.D. | - | - | - |

[a] Experiments were performed using whole *E. coli* cells according to the protocol described

in **Sections 3.4** and **3.5**. The yields of **3a** were calculated based on comparing the signal

integration ratio of the products and an added internal standard compound (1,2,3-trimethoxylbenzene). All reactions were performed in duplicate, with the reported yields representing the average of two experiments. The signal calibration curves correlating the products and the internal standard compound can be found in **Section 8**. To enhance systematic management of different enzyme variants within the Arnold lab, we recently implemented a new nomenclature system. Variants are named as follows: family name-chemistry abbreviation-entry code. All enzyme variants in this study follow this nomenclature.

[b] Negative control experiments (with free hemin) were performed without the addition of $Na_2S_2O_4$. Under this condition, the resting oxidation state of hemin in aqueous buffer should be Fe(III).

[c] Negative control experiments using free hemin under reduced conditions (Fe(II)) were performed using an excess amount of $Na_2S_2O_4$ (20 mM).

[d] Cellular background control experiments were performed in whole-cell format, using *E. coli* (*E. cloni* BL21(DE3)) cells harboring an engineered tryptophan synthase *β*-subunit (Tm9D8*). The gene of this enzyme was also cloned into the pET22b(+) vector (Novagen) between restriction sites *Nde*I and *Xho*I. The protein expression protocol for this experiment followed the standard P450 and P411 expression conditions as described in **Sections 3.6**, **3.7**, and **3.8**.

[e] When using ethyl diazoacetate (EDA) as a carbene precursor, no corresponding alkylated product was detected.

N.D. – no product was detected.

**Table 3-S2**. Directed evolution of P411-SCA-5188 for stereoconvergent alkylation.[a]
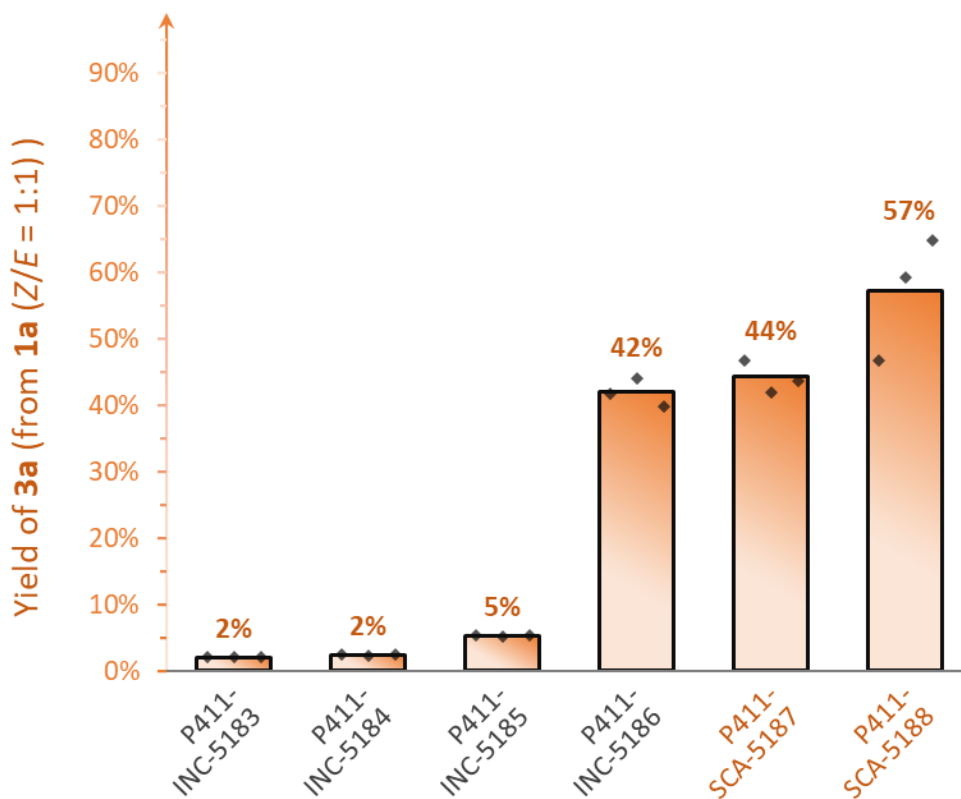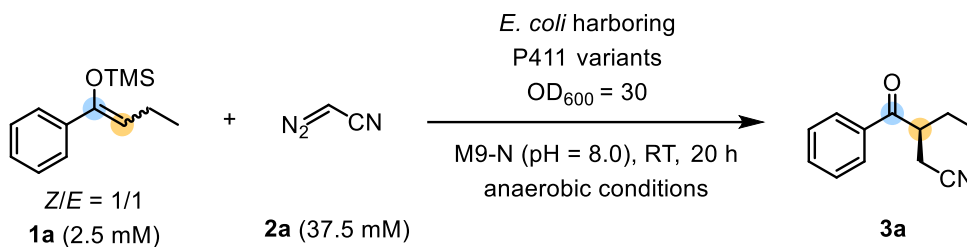


| Round # | Variant | Yield of 3a | Standard deviation of yield | TTN of 3a | Standard deviation of TTN |
|---------|---------|-------------|------------------------------|-----------|----------------------------|
| 1 | P411-INC-5186 | 42% | 2.1% | 200 | 10 |
| 2 | P411-SCA-5187 (P411-INC-5186 V437A) | 44% | 2.4% | 250 | 14 |

| 3 | P411-SCA-5188 (P411-SCA-5187 L177M) | 57% | 9.2% | 290 | 47 |
|---|---|---|---|---|---|

[a] Experiments were performed using a suspension of *E. coli* cells harboring enzyme variants prepared according to the protocol described in **Sections 3.6**, **3.7**, and **3.8**. Yields were calculated from HPLC calibration curves and the average of triplicate experiments (n = 3). TTN is defined as the amount of indicated product divided by heme protein as measured by the hemochrome assay (see **Section 3.7** for more details).

**Table 3-S3**. Activity of all lineage variants against (*Z*)-**1a**.[a]
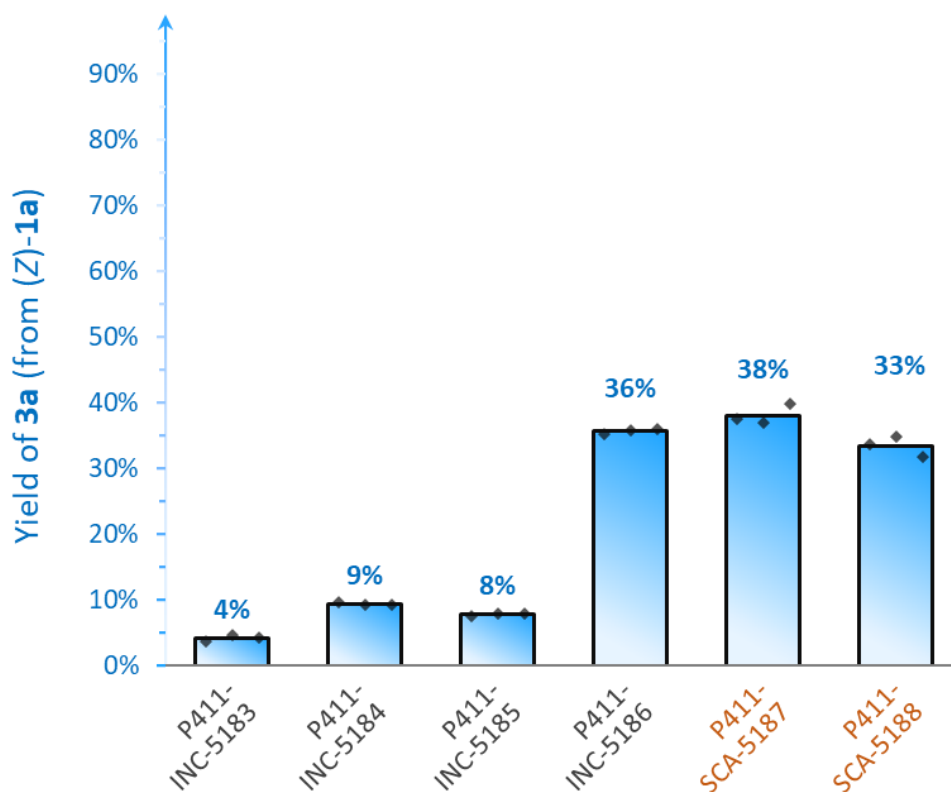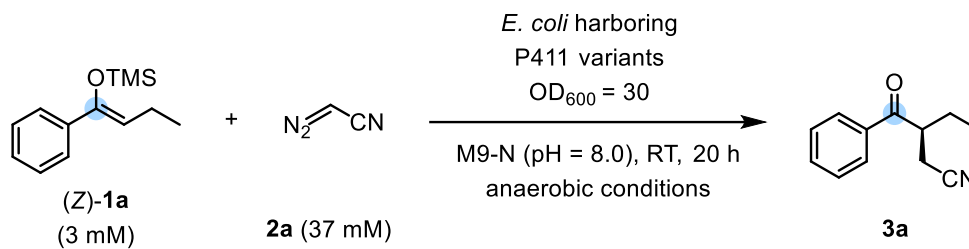


| Round # | Variant | Yield of 3a | Standard deviation of yield | TTN of 3a | Standard deviation of TTN |
|---------|---------|-------------|------------------------------|-----------|-----------------------------|
| 1 | P411-INC-5183 | 4.1% | 0.5% | 36 | 4 |

| 2 | P411-INC-5184 | 9.4% | 0.3% | 79 | 2 |
| 3 | P411-INC-5185 | 7.8% | 0.2% | 95 | 3 |
| 4 | P411-INC-5186 | 36% | 0.4% | 234 | 2 |
| 5 | P411-SCA-5187 | 38% | 1.6% | 290 | 12 |
| 6 | P411-SCA-5188 | 33% | 1.5% | 263 | 12 |

[a] Experiments were performed using a suspension of *E. coli* cells harboring enzyme variants prepared according to the protocol described in **Sections 3.6**, **3.7**, and **3.8**. Yields were calculated from HPLC calibration curves and the average of triplicate experiments (n = 3). TTN is defined as the amount of indicated product divided by heme protein as measured by the hemochrome assay (see **Section 3.7** for more details).

**Table 3-S4**. Activity of all lineage variants against (*E*)-**1a**.[a]



| Round # | Variant | Yield of 3a | Standard deviation of yield | TTN of 3a | Standard deviation of TTN |
|---------|---------|-------------|-----------------------------|-----------|---------------------------|
| 1 | P411-INC-5183 | N.D. | - | - | - |

| 2 | P411-INC-5184 | N.D. | - | - | - |
|---|---------------|------|---|---|---|
| 3 | P411-INC-5185 | N.D. | - | - | - |
| 4 | P411-INC-5186 | 85% | 4.2% | 560 | 28 |
| 5 | P411-SCA-5187 | 76% | 0.5% | 580 | 4 |
| 6 | P411-SCA-5188 | 89% | 2.3% | 700 | 18 |

[a] Experiments were performed using a suspension of *E. coli* cells harboring enzyme variants prepared according to the protocol described in **Sections 3.6**, **3.7**, and **3.8**. Yields were calculated from HPLC calibration curves and the average of triplicate experiments (n = 3). TTN is defined as the amount of indicated product divided by heme protein as measured by the hemochrome assay (see **Section 3.7** for more details).

N.D. – no product was detected.

**Table 3-S5**. Optimization of pH values (M9-N minimal media) of P411-SCA-5188-catalyzed alkylation reactions.[a]
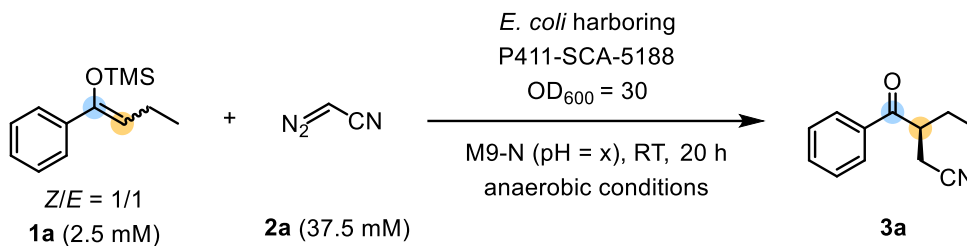


| Entry | Buffer condition | Yield of 3a | Standard deviation of yield | TTN of 3a | Standard deviation of TTN |
|---|---|---|---|---|---|
| 1 | M9-N (pH = 6.0) | 48% | 1.1% | 230 | 5 |
| 2 | M9-N (pH = 7.0) | 49% | 0.9% | 234 | 4 |
| 3 | M9-N (pH = 8.0) | 57% | 9.2% | 294 | 47 |
| 4 | M9-N (pH = 9.0) | 47% | 1.5% | 223 | 7 |

[a] Experiments were performed using a suspension of *E. coli* cells harboring P411-SCA-5188 prepared according to the protocol described in **Sections 3.6**, **3.7**, and **3.8**. Reactions were performed in triplicate. Yields reported are the average of three experiments.

**Table 3-S6**. Optimization of **1a/2a** ratios of P411-SCA-5188-catalyzed alkylation reactions.[a]



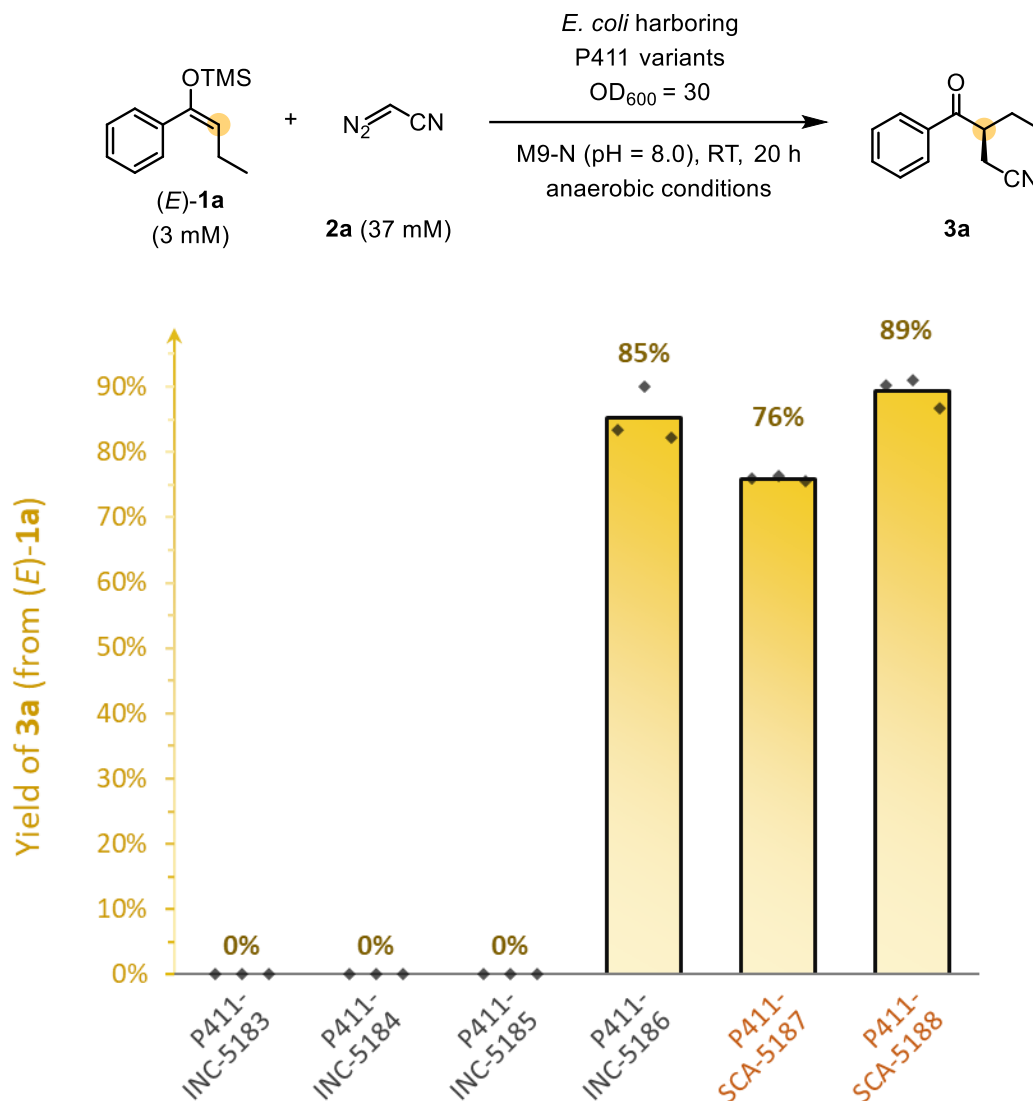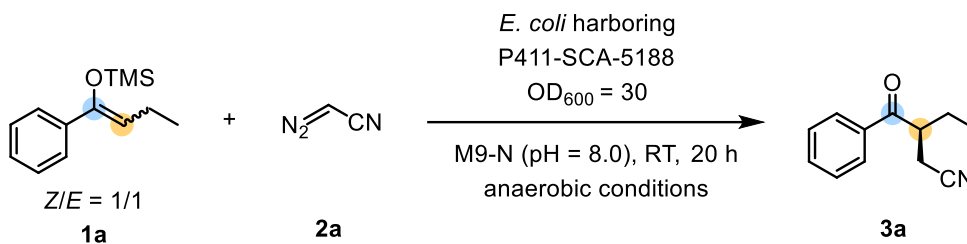| Entry | Concentration of 1a | Concentration of 2a | Yield of 3a | Standard deviation of yield | TTN of 3a | Standard deviation of TTN |
|-------|--------------------|--------------------|-------------|----------------------------|-----------|---------------------------|
| 1 | 2.5 mM | 37.5 mM | 57.1% | 9.2% | 294 | 47 |
| 2 | 3 mM | 37 mM | 60.5% | 1.5% | 414 | 10 |
| 3 | 5 mM | 35 mM | 29.7% | 1.3% | 283 | 12 |
| 4 | 10 mM | 30 mM | 19.2% | 0.7% | 367 | 13 |

[a] Experiments were performed using a suspension of *E. coli* cells harboring P411-SCA-5188 prepared according to the protocol described in **Sections 3.6**, **3.7**, and **3.8**. Reactions were performed in triplicate. Yields reported are the average of three experiments.

**Table 3-S7**. Optimization of $OD_{600}$ of P411-SCA-5188-catalyzed alkylation reactions.[a]



$Z/E = 1/1$
**1a** (2.5 mM)   **2a** (37.5 mM)   **3a**
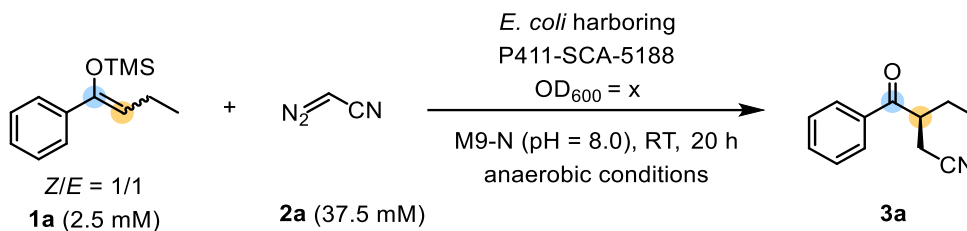
| Entry | $OD_{600}$ of cell suspension | Yield of 3a | Standard deviation of yield | TTN of 3a | Standard deviation of TTN |
|-------|------------------------------|-------------|-----------------------------|-----------|---------------------------|
| 1 | 10 | 12.8% | 1.6% | 183 | 23 |
| 2 | 30 | 57.1% | 0.9% | 294 | 47 |
| 3 | 45 | 66.9% | 2% | 239 | 7 |
| 4 | 90 | 79.4% | 17.6% | 136 | 30 |

[a] Experiments were performed using a suspension of *E. coli* cells harboring P411-SCA-5188 prepared according to the protocol described in **Sections 3.6**, **3.7**, and **3.8**. Reactions were performed in triplicate. Yields reported are the average of three experiments.

**Table 3-S8**. Optimization of *Z/E* ratios of P411-SCA-5188-catalyzed alkylation reactions.[a]



E. coli harboring
P411-SCA-5188
OD_600 = 30

M9-N (pH = 8.0), RT, 20 h
anaerobic conditions

*Z/E* = x/y
**1a** (3 mM)

**2a** (37 mM)

**3a**
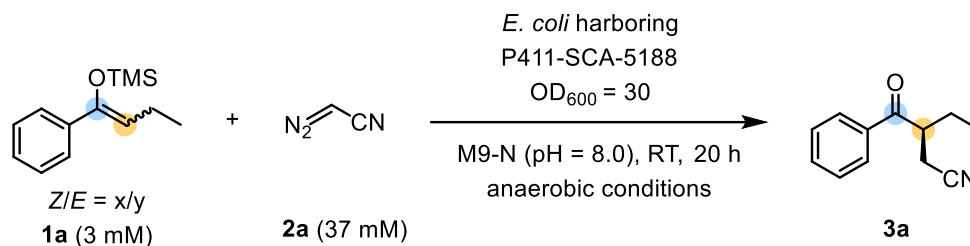
| Entry | *Z/E* (x/y) | Yield of 3a | Standard deviation of yield | TTN of 3a | Standard deviation of TTN |
|---|---|---|---|---|---|
| 1 | 1/1 | 85.2% | 1.7% | 195 | 4 |
| 2 | 0/1 | 90.8% | 4.1% | 227 | 10 |
| 3 | 1/0 | 49.4% | 0.7% | 99 | 1 |

[a] Experiments were performed using a suspension of *E. coli* cells harboring P411-SCA-5188 prepared according to the protocol described in **Sections 3.6**, **3.7**, and **3.8**. Reactions were performed in triplicate. Yields reported are the average of three experiments.

# Supplementary Figures



**Figure 3-S1. Quantifying Stereopurity of (*E*)-1a using GC-FID.** GC-FID Column: HP-5ms Ultra Inert Column (Part Number:19091S-413UIJ&W, HP-5ms Ultra Inert GC Column, 30 m, 0.32 mm, 0.25 μm, 7-inch cage). GC-FID method: Hold at 100 ºC for 2 min, ramp to 300 ºC at 50 ºC / min, and hold at 300 ºC for 2 min.
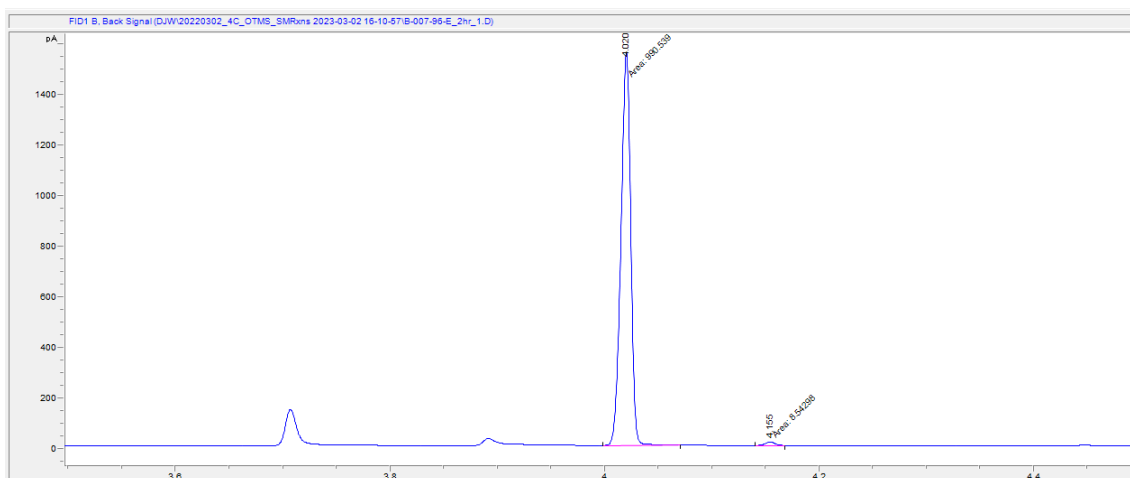


**Figure 3-S2. Quantification of Stereopurity of (*E*)-1a using GC-FID after 2-hour P411-SCA-5188-catalyzed Reaction.** GC-FID Column: HP-5ms Ultra Inert Column (Part Number:19091S-413UIJ&W, HP-5ms Ultra Inert GC Column, 30 m, 0.32 mm, 0.25 μm, 7-inch cage). GC-FID method: Hold at 100 ºC for 2 min, ramp to 300 ºC at 50 ºC / min, and hold at 300 ºC for 2 min. Reaction conditions: 3 mM (*E*)-**1a**, *E. coli* whole cells harboring P411-SCA-5188 (OD$_{600}$ = 90) M9-N aqueous buffer (pH 8.0), 10% v/v EtOH (co-solvent); room temperature, anaerobic conditions, 2 h.

**Figure 3-S3. Quantifying Stereopurity of (*Z*)-1a using GC-FID.** GC-FID Column: HP-5ms Ultra Inert Column (Part Number:19091S-413UIJ&W, HP-5ms Ultra Inert GC Column, 30 m, 0.32 mm, 0.25 µm, 7-inch cage). GC-FID method: Hold at 100 ºC for 2 min, ramp to 300 ºC at 50 ºC / min, and hold at 300 ºC for 2 min.



**Figure 3-S4. Quantification of Stereopurity of (*Z*)-1a using GC-FID after 2-hour P411-SCA-5188-catalyzed Reaction.** GC-FID Column: HP-5ms Ultra Inert Column (Part Number:19091S-413UIJ&W, HP-5ms Ultra Inert GC Column, 30 m, 0.32 mm, 0.25 µm, 7-inch cage). GC-FID method: Hold at 100 ºC for 2 min, ramp to 300 ºC at 50 ºC / min, and hold at 300 ºC for 2 min. Reaction conditions: 3 mM (*Z*)-**1a**, *E. coli* whole cells harboring P411-SCA-5188 (OD$_{600}$ = 90) M9-N aqueous buffer (pH 8.0), 10% v/v EtOH (co-solvent), room temperature, anaerobic conditions, 2 h.

**Figure 3-S5. Structures from Molecular Dynamics Simulations.** Representative structures of the most populated cluster of each independent replica for (A) (*E*)-**III,** (B) (*Z*)-**III,** (C) (*E*)-**IIIa**, and (D) (*Z*)-**IIIa**. Non-relevant, non-polar hydrogen atoms are omitted for clarity.
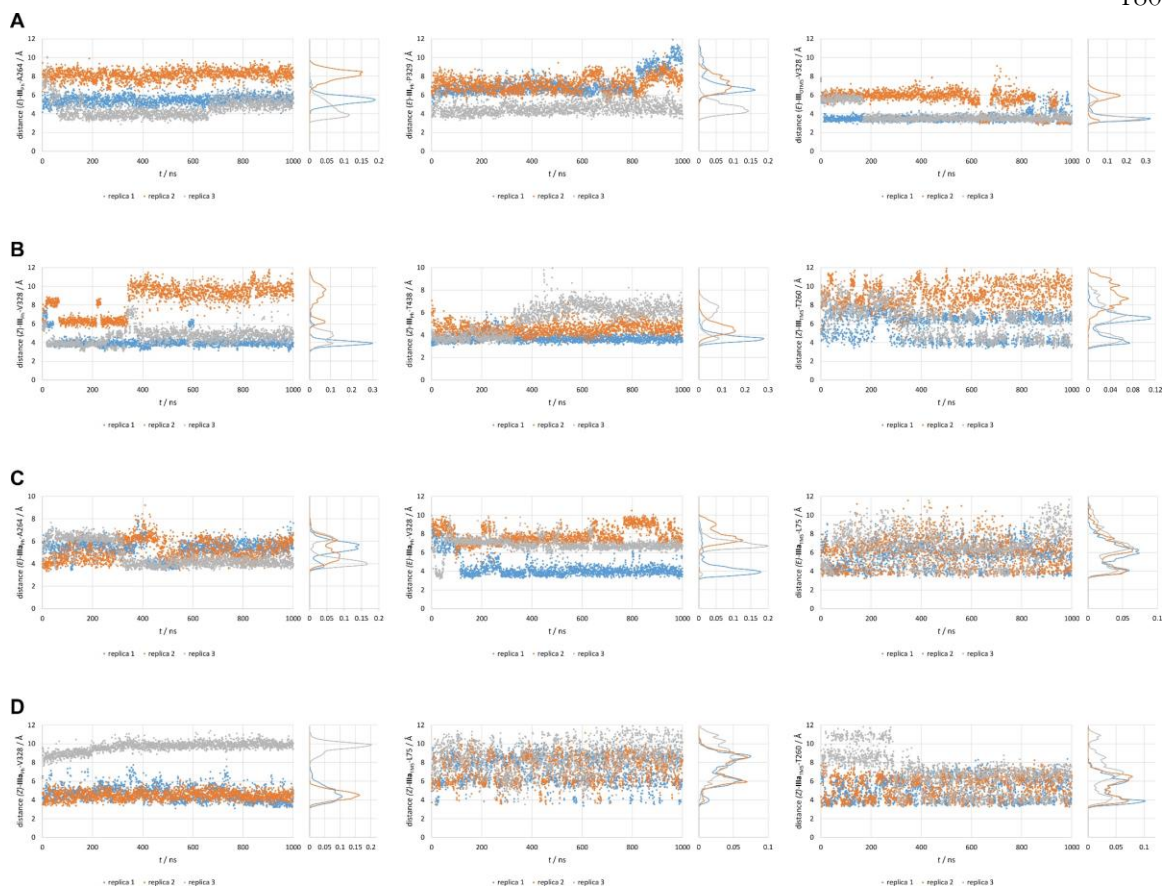
**Figure 3-S6. Distances to Key Residues from MD Simulations Data.** Distances between the ligand and selected key residues over time and probability density plots for (A) (*E*)-**III,** (B) (*Z*)-**III,** (C) (*E*)-**IIIa**, and (D) (*Z*)-**IIIa**.

## Experimental Methods

### General

Unless otherwise noted, all chemicals and reagents were obtained from commercial suppliers (Sigma-Aldrich, VWR, Alfa Aesar, Acros, Combi Blocks, etc.) and used without further purification. Silica gel chromatography was carried out using AMD Silica Gel 60, 230–400 mesh. $^{1}$H and $^{13}$C NMR spectra were recorded on a Bruker Prodigy 400 MHz instrument

(400 MHz for $^1$H and 101 MHz for $^{13}$C NMR) or a Varian 300 MHz Spectrometer (300

MHz for $^1$H NMR). Chemical shifts ($\delta$) are reported in ppm downfield from

tetramethylsilane, using the solvent resonance as the internal standard ($^1$H NMR: $\delta = 7.26$,

$^{13}$C NMR: $\delta = 77.36$ for CDCl$_3$). Data for $^1$H NMR are reported as follows: chemical shift ($\delta$

ppm), multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, p = pentet, sext = sextet,

m = multiplet, dd = doublet of doublets, dt = doublet of triplets, ddd = doublet of doublet of

doublets), coupling constant (Hz), integration. Sonication was performed using a Qsonica

Q500 sonicator. High-resolution mass spectra were obtained at the California Institute of

Technology Mass Spectral Facility. Samples were analyzed by field ionization (FI) using a

JEOL AccuTOF GC-Alpha (JMS-T2000GC) mass spectrometer interfaced with an Agilent

8890 GC system. Ions detected by FI are radical cations.

*Escherichia coli* cells were grown using Luria-Bertani medium or HyperBroth (AthenaES)

with 100 μg/mL ampicillin (LB$_{amp}$ or HB$_{amp}$). Primer sequences are available upon request.

T5 exonuclease, Phusion polymerase, and *Taq* ligase were purchased from New England

Biolabs (NEB, Ipswich, MA). M9-N minimal media (abbreviated as M9-N buffer, pH = 6.0,

7.0, 8.0, and 9.0) were used as buffering systems for whole cells, lysates, and purified

proteins, unless otherwise specified. M9-N buffer was used without a carbon source, it

contains 47.7 mM Na$_2$HPO$_4$, 22.0 mM KH$_2$PO$_4$, 8.6 mM NaCl, 2.0 mM MgSO$_4$, and 0.1

mM CaCl$_2$.

**Chromatography**

Chemical reactions were monitored using thin layer chromatography (Merck 60 silica gel

plates) and a UV lamp for visualization. Reverse-phase high-performance liquid

chromatography-mass spectroscopy (HPLC) for analysis were carried out using Agilent

1200 series instruments, with an Agilent C18 column (InfinityLab Poroshell 120 EC-C18,

4.6 x 50 mm, 2.7 µm; Part Number: 699975-902T). Water and acetonitrile containing 0.1%

acetic acid were used as eluents. Chiral analyses were conducted using either an Agilent 1200

series HPLC instrument with *n*-hexane and isopropanol as the mobile phase or an SFC using

supercritical $CO_2$ and isopropanol as the mobile phase. Enantiomers were separated using

one of the following chiral columns: Chiralpak IA (4.6 mm × 25 cm), Chiralpak IB (4.6 mm

× 25 cm), Chiralpak IC (4.6 mm × 25 cm), Chiralpak AD-H, or Chiralpak OD-H.

**Cloning and site-saturation mutagenesis**

The genes encoding all enzymes described in this study were cloned using Gibson assembly[2]

into vector pET22b(+) (Novagen) between restriction sites *Nde*I and *Xho*I in frame with a *C*-

terminal 6×His-tag. Site-saturation mutagenesis was performed using the "22c-trick"[3] or

"NNK" as degenerate codons. The PCR products were digested with *Dpn*I, gel purified,

and ligated using Gibson Mix™.[2] Without further purification after the Gibson step, 1 µL of

the Gibson product was used to transform 50 µL of electrocompetent *E. coli* BL21 *E. cloni*®

(Lucigen) cells.

**Expression of P450 and P411 variants in 96-well plates**

Single colonies from LB$_{amp}$ agar plates were picked using sterile toothpicks and cultured in

deep-well 96-well plates containing LB$_{amp}$ (400 µL/well) at 37 °C, 80% humidity, and 220

rpm shaking overnight. Subsequently, HB$_{amp}$ (1080 µL/well) in a deep-well plate was

inoculated with an aliquot (120 µL/well) of these overnight cultures and allowed to shake for

3 hours at 37 °C, 80% humidity, and 220 rpm. The plates were then cooled on ice for 30

minutes, and the cultures were induced with 0.5 mM isopropyl $\beta$-D-1-thiogalactopyranoside (IPTG) and 1.0 mM 5-aminolevulinic acid (ALA) (final concentrations). Expression was then conducted at 20 °C, 220 rpm for 18–20 hours.

**Plate reaction screening in whole-cell format**

*E. coli* cells harboring P411 variants in deep-well 96-well plates were pelleted (3,500 g, 5 min, 4 °C) and resuspended in M9-N buffer (360 μL, pH 8.0) by gentle vortexing. The 96-well plates were then transferred to an anaerobic chamber. To deep-well plates of cell suspensions were added the silyl enol ether substrate **1a** (3 μL/well, 400 mM in EtOH) and diazoacetonitrile **2a** (37 μL/well, 400 mM in EtOH). During the addition, the stock solution of diazoacetonitrile **2a** was kept on an ice bath to minimize evaporation. The plates were sealed with aluminum sealing foil immediately after the addition and shaken in the anaerobic chamber at room temperature and 600 rpm. After 20 hours, the seals were removed and ethanol (400 μL/well) was added. The plates were tightly sealed with silicone mats, vigorously vortexed, and centrifuged (4,000 x g, 5 min) to precipitate proteins and cell debris. The supernatant (200 μL/well) was filtered through an AcroPrep 96-well filter plate (0.2 μm) into a shallow 96-well plate for LC-MS analysis to determine the yield.

**Flask expression of P411 variants**

*E. coli* (*E. cloni* BL21(DE3)) cells carrying plasmid encoding the appropriate P411 variant were grown overnight in 5 mL Luria-Bertani medium supplemented with 0.1 mg/mL ampicillin (LB$_{amp}$). The preculture (1 mL) was used to inoculate 50 mL of Hyperbroth medium supplemented with 0.1 mg/mL ampicillin (HB$_{amp}$) in a 125-mL Erlenmeyer flask.

This culture was incubated at 37 °C and 230 rpm for 2.5 hours. The culture was then cooled on ice for 30 min and induced with 0.5 mM IPTG and 1.0 mM ALA (final concentrations). Expression was conducted at 20 °C, 150 rpm for 16–18 hours. Subsequently, the *E. coli* cells were pelleted by centrifugation (4,000 g, 4 min, and 4 °C). Media were removed, and the pellets were resuspended to an optical density at 600 nm ($OD_{600}$) of 30 in M9-N minimal medium with pH adjusted to 8.0. Aliquots of the cell suspension (3–4 mL) were used to determine protein concentration after lysis by sonication.

**Hemochrome assay for the determination of hemoprotein concentration**

Protein concentration in the cell was determined by performing hemochrome assay on the cell lysate.[4] Lysate was obtained by sonication (6 minutes, 1 second on, 2 seconds off, 35% amplitude, on wet ice). The cell debris was removed by centrifugation (14,000 g, 10 minutes, 4 °C). To a cuvette, 500 μL of the lysate and 500 μL of solution I [0.2 M NaOH, 40% (v/v) pyridine, 0.5 mM $K_3Fe(CN)_6$] were added. The UV-Vis spectrum (380–620 nm) of the oxidized state Fe(III) was recorded immediately. Sodium dithionite (10 μL of 0.5 M solution in water) was added, and the UV-Vis spectrum of the reduced state Fe(II) was recorded immediately. The protein concentration was calculated using the extinction coefficient and dilution factor (2× dilution in volume): $\varepsilon\_[557_{reduced} - 540_{oxidized}] = 23.98$ mM$^{-1}$cm$^{-1}$. Please note that TTN values are lower bounds since the hemochrome assay detects the levels of heme, not necessarily concentration of the enzyme; however, heme concentration closely approximates enzyme concentration.

**Biotransformation using whole *E. coli* cells**

A suspension of *E. coli* (*E. cloni* BL21(DE3)) cells expressing the appropriate hemoprotein variant in M9-N (pH = 8.0) buffer (typically $OD_{600}$ = 30) were transferred to a reaction vial. The headspace of the reaction vial was purged with a stream of argon for at least 15 minutes. Enzymatic reactions were then set up in an anaerobic chamber. To a 2-mL vial were added degassed suspension of *E. coli* expressing P411 variant (typically $OD_{600}$ = 30, 360 μL), silyl enol ether substrates (typically 3 μL of 400 mM stock solution in EtOH), and the diazoacetonitrile solution (typically 37 μL of 400 mM stock solution in EtOH). The concentration of the diazo solution was measured by $^1$H-NMR and was kept in an ice bath and added last. The final volume of the biotransformation was set to be 400 μL, with 10% vol EtOH. The reaction vials were then capped and shaken in the anaerobic chamber at room temperature and 600 rpm for 20 hours. After the completion of the reaction, 400 μL ethanol containing 2 mM 1,2,3-trimethoxybenzene internal standard were added to the vial. The resulting mixture was transferred to a 1.7-mL microcentrifuge tube, vigorously vortexed, and centrifuged (14,000 × g, 5 min, 4 °C). A sample of the supernatant (0.2 mL) was transferred to a vial with an insert for LC-MS analysis.

To determine the enantiomeric excess (e.e.) of the α-branched ketones **3**, the remaining supernatants of three parallel analytical reactions were combined and transferred to a 2-dram vial and the solvent was removed by blowing air. The α-branched ketones **3** were extracted from the remaining residues with a solution of hexane and EtOAc mixture (1:1) and subjected to normal-phase HPLC to determine the enantiomeric excess (e.e.).

# Nucleotide and Amino Acid Sequences

The genes encoding the heme proteins shown below were cloned using Gibson assembly[2] into vector pET-22b(+) (Novagen) between restriction sites *Nde*I and *Xho*I in frame with a *C*-terminal 6×His-tag.

**Table 3-S9**. Mutations of the P411 variants in this study.

| Name | Mutations relative to the wild type P450$_{BM3}$ | Mutations relative to P411-INC-5182 |
|---|---|---|
| P411-INC-5182 | A74G, V78L, A82L, F87A, P142S, T175I, M177L, A184V, S226R, H236Q, E252G, I263Y, T268G, A290V, T327I, A328V, L353V, I366V, C400S, T436L, L437Q, E442K, ΔFAD domain | - |
| P411-INC-5183 | A74G, V78L, A82L, F87A, P142S, T175I, M177L, A184V, S226R, H236Q, E252G, I263W, T268G, A290V, T327P, A328V, L353V, I366V, C400S, T436L, L437Q, E442K, ΔFAD domain | Y263W, I327P |
| P411-INC-5184 | A74G, V78L, A82L, F87A, P142S, T175I, M177L, A184V, S226R, H236Q, E252G, I263W, T268G, A290V, T327P, A328V, L353V, I366V, C400S, T436L, L437V, E442K, ΔFAD domain | Y263W, I327P, Q437V |
| P411-INC-5185 | N70S, A74G, V78L, A82L, F87A, P142S, T175I, M177L, A184V, S226R, H236Q, E252G, I263W, T268G, A290V, T327P, A328V, L353V, I366V, C400S, T436L, L437V, E442K, ΔFAD domain | N70S, Y263W, I327P, Q437V |

| P411-INC-5186 | N70S, A74G, V78L, A82L, F87A, P142S, T175I, M177L, A184V, S226R, H236Q, E252G, I263M, T268G, A290V, T327P, A328V, L353V, I366V, C400S, T436L, L437V, E442K, ΔFAD domain | N70S, Y263M, I327P, Q437V |
|---|---|---|
| P411-SCA-5187 | N70S, A74G, V78L, A82L, F87A, P142S, T175I, M177L, A184V, S226R, H236Q, E252G, I263M, T268G, A290V, T327P, A328V, L353V, I366V, C400S, T436L, L437A, E442K, ΔFAD domain | N70S, Y263M, I327P, Q437A |
| P411-SCA-5188 | N70S, A74G, V78L, A82L, F87A, P142S, T175I, A184V, S226R, H236Q, E252G, I263M, T268G, A290V, T327P, A328V, L353V, I366V, C400S, T436L, L437A, E442K, ΔFAD domain | N70S, L177M, Y263M, I327P, Q437A |

DNA and amino acid sequences of P411-INC-5186:

ATGACAATTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATTTACCGTTATTAA

ACACAGATAAACCGGTTCAAGCTTTGATGAAAATTGCGGATGAATTAGGAGAAATCTTTAAATT

CGAGGCGCCTGGTCGTGTAACGCGCTACTTATCAAGTCAGCGTCTAATTAAAGAAGCATGCGAT

GAATCACGCTTTGATAAATCTTTAAGTCAAGGTCTGAAATTTCTGCGTGATTTTCTTGGAGACGG

GTTAGCCACAAGCTGGACGCATGAAAAAAATTGGAAAAAAGCGCATAATATCTTACTTCCAAG

CTTTAGTCAGCAGGCAATGAAAGGCTATCATGCGATGATGGTCGATATCGCCGTGCAGCTTGTT

CAAAAGTGGGAGCGTCTAAATGCAGATGAGCATATTGAAGTATCGGAAGACATGACACGTTTA

ACGCTTGATACAATTGGTCTTTGCGGCTTTAACTATCGCTTTAACAGCTTTTACCGAGATCAGCC

TCATCCATTTATTATAAGTCTGGTCCGTGCACTGGATGAAGTAATGAACAAGCTGCAGCGAGCA

AATCCAGACGACCCAGCTTATGATGAAAACAAGCGCCAGTTTCAAGAAGATATCAAGGTGATG

AACGACCTAGTAGATAAAATTATTGCAGATCGCAAAGCAAGGGGTGAACAAAGCGATGATTTA

TTAACGCAGATGCTAAACGGAAAAGATCCAGAAACGGGTGAGCCGCTTGATGACGGGAACATT

CGCTATCAAATTATTACATTCTTAATGGCGGGACACGAAGGTACAAGTGGTCTTTTATCATTTGC

GCTGTATTTCTTAGTGAAAAATCCACATGTATTACAAAAAGTAGCAGAAGAAGCAGCACGAGT

TCTAGTAGATCCTGTTCCAAGCTACAAACAAGTCAAACAGCTTAAATATGTCGGCATGGTCTTA

AACGAAGCGCTGCGCTTATGGCCACCGGTTCCTGCGTTTTCCCTATATGCAAAAGAAGATACGG

TGCTTGGAGGAGAATATCCTTTAGAAAAAGGCGACGAAGTAATGGTTCTGATTCCTCAGCTTCA

CCGTGATAAAACAGTTTGGGGAGACGATGTGGAGGAGTTCCGTCCAGAGCGTTTTGAAAATCC

AAGTGCGATTCCGCAGCATGCGTTTAAACCGTTTGGAAACGGTCAGCGTGCGTCTATCGGTCAG

CAGTTCGCTCTTCATGAAGCAACGCTGGTACTTGGTATGATGCTAAAACACTTTGACTTTGAAG

ATCATACAAACTACGAGCTCGATATTAAAGAACTGGTGACGTTAAAACCTAAAGGCTTTGTGGT

AAAAGCAAAATCGAAAAAAATTCCGCTTGGCGGTATTCCTTCACCTAGCACTGAACAGTCTGCT

AAAAAAGTACGCAAAAAGGCAGAAAACGCTCATAATACGCCGCTGCTTGTGCTATACGGTTCA

AATATGGGTACCGCTGAAGGAACGGCGCGTGATTTAGCAGATATTGCAATGAGCAAAGGATTT

GCACCGCAGGTCGCAACGCTTGATTCACACGCCGGAAATCTTCCGCGCGAAGGAGCTGTATTAA

TTGTAACGGCGTCTTATAACGGTCATCCGCCTGATAACGCAAAGCAATTTGTCGACTGGTTAGA

CCAAGCGTCTGCTGATGAAGTAAAAGGCGTTCGCTACTCCGTATTTGGATGCGGCGATAAAAAC

TGGGCTACTACGTATCAAAAAGTGCCTGCTTTTATCGATGAAACGCTTGCCGCTAAAGGGGCAG

AAAACATCGCTGACCGCGGTGAAGCAGATGCAAGCGACGACTTTGAAGGCACATATGAAGAAT

GGCGTGAACATATGTGGAGTGACGTAGCAGCCTACTTTAACCTCGACATTGAAAACAGTGAAG

ATAATAAATCTACTCTTTCACTTCAATTTGTCGACAGCGCCGCGGATATGCCGCTTGCGAAAAT

GCACGGTGCGTTTTCAACGCTCGAGCACCACCACCACCACCACTGA

MTIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQRLIKEACDESRF

DKSLSQGLKFLRDFLGDGLATSWTHEKNWKKAHNILLPSFSQQAMKGYHAMMVDIAVQLVQKWE

RLNADEHIEVSEDMTRLTLDTIGLCGFNYRFNSFYRDQPHPFIISLVRALDEVMNKLQRANPDDPAY

DENKRQFQEDIKVMNDLVDKIIADRKARGEQSDDLLTQMLNGKDPETGEPLDDGNIRYQIITFLMA

GHEGTSGLLSFALYFLVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRLWPPVP

AFSLYAKEDTVLGGEYPLEKGDEVMVLIPQLHRDKTVWGDDVEEFRPERFENPSAIPQHAFKPFGN

GQRASIGQQFALHEATLVLGMMLKHFDFEDHTNYELDIKELVTLKPKGFVVKAKSKKIPLGGIPSPS

TEQSAKKVRKKAENAHNTPLLVLYGSNMGTAEGTARDLADIAMSKGFAPQVATLDSHAGNLPREG

AVLIVTASYNGHPPDNAKQFVDWLDQASADEVKGVRYSVFGCGDKNWATTYQKVPAFIDETLAA

KGAENIADRGEADASDDFEGTYEEWREHMWSDVAAYFNLDIENSEDNKSTLSLQFVDSAADMPLA

KMHGAFSTLEHHHHHH*

DNA and amino acid sequences of P411-SCA-5188:

ATGACAATTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATTTACCGTTATTAA

ACACAGATAAACCGGTTCAAGCTTTGATGAAAATTGCGGATGAATTAGGAGAAATCTTTAAATT

CGAGGCGCCTGGTCGTGTAACGCGCTACTTATCAAGTCAGCGTCTAATTAAAGAAGCATGCGAT

GAATCACGCTTTGATAAATCTTTAAGTCAAGGTCTGAAATTTCTGCGTGATTTTCTTGGAGACGG

GTTAGCCACAAGCTGGACGCATGAAAAAAATTGGAAAAAAGCGCATAATATCTTACTTCCAAG

CTTTAGTCAGCAGGCAATGAAAGGCTATCATGCGATGATGGTCGATATCGCCGTGCAGCTTGTT

CAAAAGTGGGAGCGTCTAAATGCAGATGAGCATATTGAAGTATCGGAAGACATGACACGTTTA

ACGCTTGATACAATTGGTCTTTGCGGCTTTAACTATCGCTTTAACAGCTTTTACCGAGATCAGCC

TCATCCATTTATTATAAGTATGGTCCGTGCACTGGATGAAGTAATGAACAAGCTGCAGCGAGCA

AATCCAGACGACCCAGCTTATGATGAAAACAAGCGCCAGTTTCAAGAAGATATCAAGGTGATG

AACGACCTAGTAGATAAAATTATTGCAGATCGCAAAGCAAGGGGTGAACAAAGCGATGATTTA

TTAACGCAGATGCTAAACGGAAAAGATCCAGAAACGGGTGAGCCGCTTGATGACGGGAACATT

CGCTATCAAATTATTACATTCTTAATGGCGGGACACGAAGGTACAAGTGGTCTTTTATCATTTGC

GCTGTATTTCTTAGTGAAAAATCCACATGTATTACAAAAAGTAGCAGAAGAAGCAGCACGAGT

TCTAGTAGATCCTGTTCCAAGCTACAAACAAGTCAAACAGCTTAAATATGTCGGCATGGTCTTA

AACGAAGCGCTGCGCTTATGGCCACCGGTTCCTGCGTTTTCCCTATATGCAAAAGAAGATACGG

TGCTTGGAGGAGAATATCCTTTAGAAAAAGGCGACGAAGTAATGGTTCTGATTCCTCAGCTTCA

CCGTGATAAAACAGTTTGGGGAGACGATGTGGAGGAGTTCCGTCCAGAGCGTTTTGAAAATCC

AAGTGCGATTCCGCAGCATGCGTTTAAACCGTTTGGAAACGGTCAGCGTGCGTCTATCGGTCAG

CAGTTCGCTCTTCATGAAGCAACGCTGGTACTTGGTATGATGCTAAAACACTTTGACTTTGAAG

ATCATACAAACTACGAGCTCGATATTAAAGAACTGGCGACGTTAAAACCTAAAGGCTTTGTGGT

AAAAGCAAAATCGAAAAAAATTCCGCTTGGCGGTATTCCTTCACCTAGCACTGAACAGTCTGCT

AAAAAAGTACGCAAAAAGGCAGAAAACGCTCATAATACGCCGCTGCTTGTGCTATACGGTTCA

AATATGGGTACCGCTGAAGGAACGGCGCGTGATTTAGCAGATATTGCAATGAGCAAAGGATTT

GCACCGCAGGTCGCAACGCTTGATTCACACGCCGGAAATCTTCCGCGCGAAGGAGCTGTATTAA

TTGTAACGGCGTCTTATAACGGTCATCCGCCTGATAACGCAAAGCAATTTGTCGACTGGTTAGA

CCAAGCGTCTGCTGATGAAGTAAAAGGCGTTCGCTACTCCGTATTTGGATGCGGCGATAAAAAC

TGGGCTACTACGTATCAAAAAGTGCCTGCTTTTATCGATGAAACGCTTGCCGCTAAAGGGGCAG

AAAACATCGCTGACCGCGGTGAAGCAGATGCAAGCGACGACTTTGAAGGCACATATGAAGAAT

GGCGTGAACATATGTGGAGTGACGTAGCAGCCTACTTTAACCTCGACATTGAAAACAGTGAAG

ATAATAAATCTACTCTTTCACTTCAATTTGTCGACAGCGCCGCGGATATGCCGCTTGCGAAAAT

GCACGGTGCGTTTTCAACGCTCGAGCACCACCACCACCACCACTGA


MTIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQRLIKEACDESRF

DKSLSQGLKFLRDFLGDGLATSWTHEKNWKKAHNILLPSFSQQAMKGYHAMMVDIAVQLVQKWE

RLNADEHIEVSEDMTRLTLDTIGLCGFNYRFNSFYRDQPHPFIISMVRALDEVMNKLQRANPDDPAY

DENKRQFQEDIKVMNDLVDKIIADRKARGEQSDDLLTQMLNGKDPETGEPLDDGNIRYQIITFLMA

GHEGTSGLLSFALYFLVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRLWPPVP

AFSLYAKEDTVLGGEYPLEKGDEVMVLIPQLHRDKTVWGDDVEEFRPERFENPSAIPQHAFKPFGN

GQRASIGQQFALHEATLVLGMMLKHFDFEDHTNYELDIKELATLKPKGFVVKAKSKKIPLGGIPSPS

TEQSAKKVRKKAENAHNTPLLVLYGSNMGTAEGTARDLADIAMSKGFAPQVATLDSHAGNLPREG

AVLIVTASYNGHPPDNAKQFVDWLDQASADEVKGVRYSVFGCGDKNWATTYQKVPAFIDETLAA

KGAENIADRGEADASDDFEGTYEEWREHMWSDVAAYFNLDIENSEDNKSTLSLQFVDSAADMPLA

KMHGAFSTLEHHHHHH*

## Substrate Syntheses and Characterizations

### General procedure A: preparation of mixtures of silyl enol ethers 1
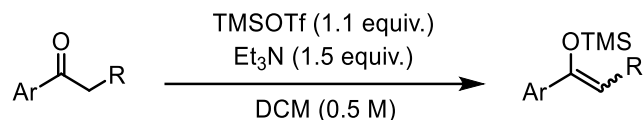


Step 1: Silyl enol ethers were prepared using a modified procedure reported by Melchiorre and coworkers.[5] In an oven-dried round-bottom flask kept under nitrogen, ketone (2.5 mmol) was dissolved in dry DCM (5 mL), and the resulting solution was cooled in an ice bath. Triethylamine (1.5 equiv., 3.75 mmol, 523 µL) was added followed by trimethylsilyl trifluoromethanesulfonate (TMSOTf, 1.1 equiv., 2.75 mmol, 499 µL). The mixture was stirred for two hours at this temperature and then was warmed up to room temperature and stirred overnight. Next, the reaction mixture was poured into water/DCM in an extraction funnel. The phases were separated, and the water phase was extracted twice with DCM, the combined organic fractions dried ($Na_2SO_4$) and concentrated to dryness. The residue was purified by flash column chromatography (hexanes/EtOAc 90/10 containing 1% triethylamine) to afford the desired product.
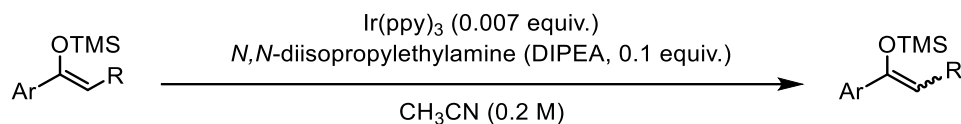


Step 2: Isomerization of **1** was done via a known procedure,[6] but with slight modifications as follows. A vial fitted with a rubber septum was charged with $Ir(ppy)_3$ (0.007 equiv., 1.88 mM stock solution of catalyst in $CH_3CN$), **1** (1 equiv.), diisopropylethylamine

(DIPEA, 0.1 equiv.), and the reaction mixture was degassed via Argon (Ar) gas bubbling

for 5–10 min and then left under positive Ar pressure by removing the exit needle. The vial

was placed in a light bath (Blue LEDs were purchased from Kessil Co., Ltd. (40 W max.,

product No. A160WE)) for 5 hours. The crude product was purified by normal phase

chromatography. The *Z/E* ratio of **1** was measured and determined by ^{1}H NMR (see 0 for

details).
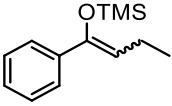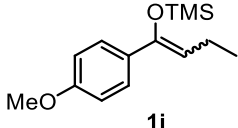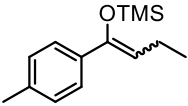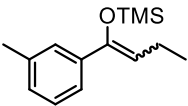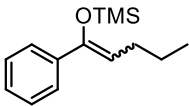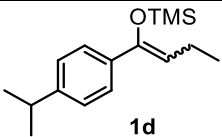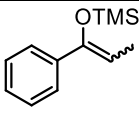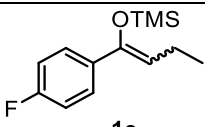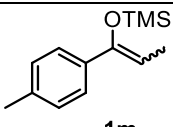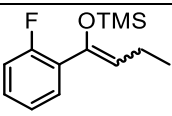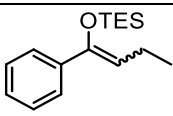
### General procedure B: preparation of mixtures of silyl enol ethers 1n, 1o, and 1p
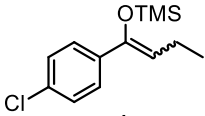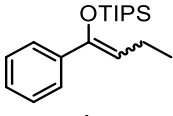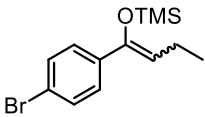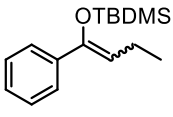


Step 1: Silyl enol ethers **1n**, **1o,** and **1p** were prepared using a modified procedure
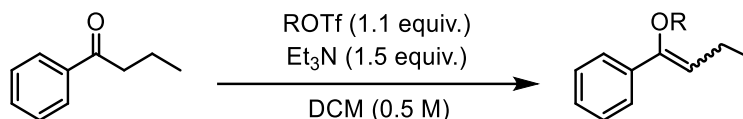
reported by Melchiorre and coworkers.[5] In an oven-dried round-bottom flask kept under

nitrogen, butyrophenone (2.5 mmol) was dissolved in dry DCM (5 mL) and the resulting

solution was cooled in an ice bath. Triethylamine (1.5 equiv., 3.75 mmol, 523 µL) was added

followed by ROTf (1.1 equiv., 2.75 mmol, triethylsilyl trifluoromethanesulfonate (TESOTf),

triisopropylsilyl      trifluoromethanesulfonate      (TIPSOTf),      or      *tert*-butyldimethylsilyl

trifluoromethanesulfonate (TBDMSOTf) for the synthesis of **1n**, **1o** or **1p**, respectively). The

mixture was stirred for two hours at this temperature and then was warmed up to room

temperature and stirred overnight. Next, the reaction mixture was poured into water/DCM in

an extraction funnel. The phases were separated, and the water phase was extracted twice

with DCM, the combined organic fractions dried ($Na_2SO_4$) and concentrated to dryness. The

residue was purified by flash column chromatography (hexanes/EtOAc 90/10 containing 1%

triethylamine) to afford the desired product. The *Z/E* ratio of **1n**, **1o**, **1p** was measured

and determined by ^{1}H NMR (see 0 for details).

**Table 3-S10**. *Z/E* ratios of silyl enol ethers **1**.[a]

| 1 | Z/E | 1 | Z/E |
|---|-----|---|-----|
|  **1a** | 1/1 [b] |  **1i** | 1/1 |
|  **1b** | 1/2.2 |  **1j** | 1/0.6 |
|  **1c** | 1/1.7 |  **1k** | 1/1.5 |
|  **1d** | 1/2.3 |  **1l** | 1/1 |
|  **1e** | 1/1.2 |  **1m** | 1/2.3 |
|  **1f** | 1/0.1 [c] |  **1n** | 1/1.3 |

| | | | |
|---|---|---|---|
| OTMS structure (4-Cl) **1g** | 1/0.5 | OTIPS structure **1o** | 1/4 |
| OTMS structure (4-Br) **1h** | 1/0.05[c] | OTBDMS structure **1p** | 1/1.5 |

[a] Silyl enol ethers **1** were prepared according to the protocol described in General procedure A:. *Z/E*- ratios were determined by ^{1}H NMR.

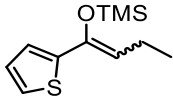[b] Chiral chromatographic separation of the isomers of **1a** was conducted using a Thar analytical SFC system using a Mettler-Toledo column compartment and an Agilent 1200 series G1315B diode-array detector using isopropanol/$CO_2$ and two 4.6 x 250 mm Chiralcel® AD-H columns.

[c] Silyl enol ether **1f** and **1h** were prepared according to the protocol described in General procedure A:, step 1. *Z/E*- ratios were determined by ^{1}H NMR.

**(Z)-trimethyl((1-phenylbut-1-en-1-yl)oxy)silane ((Z)-1a)**



(Z)-**1a**
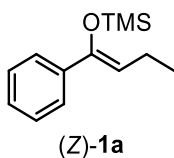
**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.49 – 7.43 (m, 2H), 7.29 (tt, *J* = 6.6, 1.0 Hz, 2H), 7.25 – 7.19 (m, 1H), 5.24 (t, *J* = 7.1 Hz, 1H), 2.22 (p, *J* = 7.4 Hz, 2H), 1.04 (t, *J* = 7.5 Hz, 3H), 0.13 (s, 9H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 148.49 , 139.35 , 128.14 , 127.43 , 125.43 , 113.47 , 19.69 , 14.40 , 0.66 .

**(E)-trimethyl((1-phenylbut-1-en-1-yl)oxy)silane ((E)-1a)**



(E)-**1a**
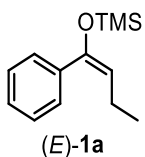
**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.39 – 7.26 (m, 5H), 5.02 (t, *J* = 7.6 Hz, 1H), 2.11 (p, *J* = 7.5 Hz, 2H), 0.99 (t, *J* = 7.4 Hz, 3H), 0.13 (s, 9H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 148.79 , 137.94 , 128.42 , 127.92 , 127.78 , 113.35 , 21.16 , 15.50 , 0.42 .

**trimethyl((1-(*p*-tolyl)but-1-en-1-yl)oxy)silane (1b)**
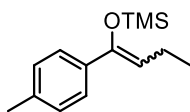


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.35 – 7.09 (m, 4H), 5.28 – 4.90 (m, 1H), 2.34 (d, *J* = 8.1 Hz, 3H), 2.25 – 2.06 (m, 2H), 1.01 (dt, *J* = 20.3, 7.4 Hz, 3H), 0.13 (d, *J* = 1.9 Hz, 9H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 148.79 , 148.48 , 137.51 , 137.14 , 136.54 , 135.07 , 128.83 , 128.60 , 128.31 , 125.35 , 112.95 , 112.64 , 21.20 , 19.65 , 15.52 , 14.45 , 0.68 , 0.44 .

**trimethyl((1-(*m*-tolyl)but-1-en-1-yl)oxy)silane (1c)**
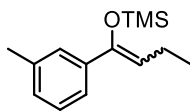


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.43 – 6.97 (m, 4H), 5.11 (dt, *J* = 91.8, 7.4 Hz, 1H), 2.35 (d, *J* = 6.1 Hz, 3H), 2.16 (dp, *J* = 41.8, 7.5 Hz, 2H), 1.01 (dt, *J* = 21.8, 7.5 Hz, 3H), 0.13 (s, 9H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 148.49 , 148.12 , 138.80 , 137.41 , 137.14 , 136.98 , 128.60 , 128.07 , 127.73 , 127.58 , 127.32 , 125.69 , 125.12 , 122.15 , 112.81 , 112.68 , 21.18 , 20.74 , 19.22 , 15.06 , 13.96 , 0.22 , 0.00 .

**((1-(4-isopropylphenyl)but-1-en-1-yl)oxy)trimethylsilane (1d)**
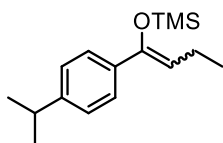


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.24 – 6.99 (m, 4H), 4.96 (dt, *J* = 87.2, 7.4 Hz, 1H), 2.88 – 2.69 (m, 1H), 2.04 (dp, *J* = 29.9, 7.5 Hz, 2H), 1.12 (dd, *J* = 6.9, 4.4 Hz, 9H), 0.92 – 0.83 (m, 3H), 0.01 (s, 9H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 148.84 , 148.50 , 148.44 , 148.15 , 136.82 , 135.36 , 128.45 , 128.32 , 126.76 , 126.16 , 125.93 , 125.35 , 112.79 , 112.66 , 77.36 , 34.02 , 33.90 , 24.10 , 23.85 , 21.23 , 19.67 , 15.57 , 14.46 , 0.70 , 0.45 .

**((1-(4-fluorophenyl)but-1-en-1-yl)oxy)trimethylsilane (1e)**
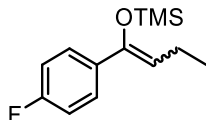


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.39 – 7.31 (m, 2H), 7.11 – 6.88 (m, 2H), 5.08 (dt, *J* = 61.2, 7.4 Hz, 1H), 2.32 – 1.97 (m, 2H), 1.01 (dt, *J* = 18.7, 7.5 Hz, 3H), 0.12 (d, *J* = 2.6 Hz, 9H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 163.53 (d, *J* = 8.4 Hz), 161.08 (d, *J* = 9.0 Hz), 147.76 (d, *J* = 19.9 Hz), 135.57 (d, *J* = 3.3 Hz), 134.03 (d, *J* = 3.3 Hz), 130.10 (d, *J* = 8.0 Hz), 127.08 (d, *J* = 7.9 Hz), 115.03 (d, *J* = 12.9 Hz), 114.81 (d, *J* = 12.8 Hz), 113.24 (d, *J* = 2.1 Hz), 77.36 , 21.17 , 19.67 , 15.45 , 14.38 , 0.64 , 0.40 .

## ((1-(2-fluorophenyl)but-1-en-1-yl)oxy)trimethylsilane (1f)
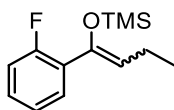


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.49 – 7.28 (m, 1H), 7.20 (dddd, *J* = 8.1, 7.0, 4.9, 1.9 Hz, 1H), 7.13 – 6.97 (m, 2H), 5.26 – 5.08 (m, 1H), 2.40 – 1.81 (m, 2H), 1.04 (t, *J* = 7.5 Hz, 3H), 0.09 (s, 9H).

## ((1-(4-chlorophenyl)but-1-en-1-yl)oxy)trimethylsilane (1g)
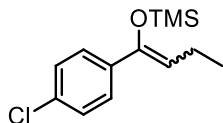


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.40 – 7.25 (m, 4H), 5.12 (dt, *J* = 78.9, 7.4 Hz, 1H), 2.26 – 2.00 (m, 2H), 1.01 (dt, *J* = 18.6, 7.5 Hz, 3H), 0.13 (d, *J* = 1.0 Hz, 9H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 147.70 , 147.53 , 137.90 , 136.43 , 133.47 , 133.09 , 129.69 , 129.61 , 129.01 , 128.32 , 128.17 , 126.67 , 113.99 , 113.86 , 77.36 , 21.17 , 19.71 , 15.42 , 14.31 , 0.66 , 0.40 .

## ((1-(4-bromophenyl)but-1-en-1-yl)oxy)trimethylsilane (1h)
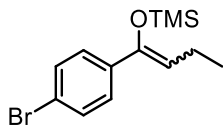


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.40 – 7.25 (m, 4H), 5.12 (dt, *J* = 78.9, 7.4 Hz, 1H), 2.26 – 2.00 (m, 2H), 1.01 (dt, *J* = 18.6, 7.5 Hz, 3H), 0.13 (d, *J* = 1.0 Hz, 9H).

**((1-(4-methoxyphenyl)but-1-en-1-yl)oxy)trimethylsilane (1i)**
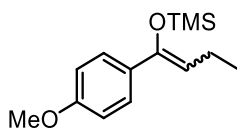


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.35 (dd, *J* = 31.3, 8.9 Hz, 2H), 6.84 (dd, *J* = 14.1, 8.8 Hz, 2H), 5.03 (dt, *J* = 62.7, 7.3 Hz, 1H), 3.81 (d, *J* = 4.5 Hz, 3H), 2.15 (dp, *J* = 35.7, 7.5 Hz, 2H), 1.01 (dt, *J* = 18.0, 7.5 Hz, 3H), 0.12 (s, 9H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 159.12 , 148.47 , 148.20 , 132.10 , 130.51 , 129.65 , 126.70 , 113.48 , 113.27 , 112.49 , 111.84 , 77.36 , 55.39 , 55.37 , 21.22 , 19.64 , 15.54 , 14.51 , 0.67 , 0.43 .

**trimethyl((1-(thiophen-2-yl)but-1-en-1-yl)oxy)silane (1j)**
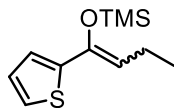


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.26 – 6.88 (m, 3H), 5.11 (dt, *J* = 91.3, 7.3 Hz, 1H), 2.24 (dp, *J* = 52.8, 7.4 Hz, 2H), 1.09 – 0.98 (m, 3H), 0.21 (s, 9H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 143.83 , 143.49 , 143.10 , 141.32 , 127.16 , 126.81 , 125.66 , 125.04 , 123.77 , 122.90 , 113.48 , 112.72 , 21.12 , 19.59 , 15.04 , 14.28 , 0.67 , 0.31 .

**trimethyl((1-phenylpent-1-en-1-yl)oxy)silane (1k)**
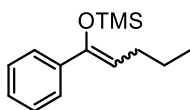


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.49 – 7.19 (m, 5H), 5.15 (dt, J = 89.8, 7.4 Hz, 1H), 2.13 (dq, J = 38.5, 7.4 Hz, 2H), 1.43 (dq, J = 21.9, 7.4 Hz, 2H), 0.93 (dt, J = 32.1, 7.4 Hz, 3H), 0.13 (d, J = 1.4 Hz, 9H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 148.80 , 148.62 , 139.00 , 137.58 , 128.07 , 127.71 , 127.48 , 127.30 , 127.00 , 125.03 , 111.17 , 111.09 , 76.94 , 29.39 , 28.04 , 23.61 , 22.67 , 13.77 , 13.48 , 0.29 , 0.01 .

**trimethyl((1-phenylprop-1-en-1-yl)oxy)silane (1l)**
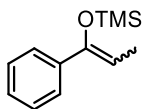


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.53 – 7.13 (m, 5H), 5.22 (dq, *J* = 91.1, 7.1 Hz, 1H), 1.72 (dd, *J* = 13.7, 7.1 Hz, 3H), 0.13 (d, *J* = 5.8 Hz, 9H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 149.97 , 149.74 , 139.31 , 137.73 , 128.52 , 128.15 , 127.92 , 127.69 , 127.41 , 125.32 , 105.50 , 105.13 , 77.36 , 13.28 , 11.83 , 0.72 , 0.42 .

**trimethyl((1-(*p*-tolyl)prop-1-en-1-yl)oxy)silane (1m)**
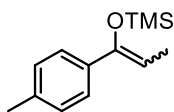


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.46 – 7.01 (m, 4H), 5.36 – 4.97 (m, 1H), 2.35 (d, *J* = 9.4 Hz, 3H), 1.77 – 1.65 (m, 3H), 0.13 (d, *J* = 7.0 Hz, 9H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 149.52 , 149.31 , 136.96 , 136.66 , 136.07 , 134.41 , 128.92 , 128.40 , 128.16 , 127.96 , 124.79 , 104.23 , 104.17 , 76.92 , 20.97 , 20.83 , 12.86 , 11.34 , 0.30 , -0.00 .

**triethyl((1-phenylbut-1-en-1-yl)oxy)silane (1n)**
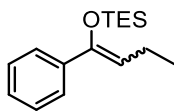


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.41 – 7.18 (m, 6H), 5.00 (dt, *J* = 39.4, 7.4 Hz, 1H), 2.10 (dp, *J* = 48.6, 7.4 Hz, 2H), 0.98 – 0.89 (m, 3H), 0.88 – 0.84 (m, 9H), 0.56 – 0.42 (m, 6H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 148.79 , 139.69 , 137.90 , 128.23 , 127.93 , 127.71 , 127.59 , 127.30 , 125.50 , 113.24 , 112.66 , 77.22 , 21.06 , 19.41 , 15.36 , 14.27 , 6.82 , 6.72 , 6.43 , 5.32 , 5.01 .

**triisopropyl((1-phenylbut-1-en-1-yl)oxy)silane (1o)**
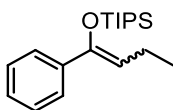


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.47 – 7.10 (m, 5H), 4.90 (dt, *J* = 18.4, 7.3 Hz, 1H), 2.25 – 1.96 (m, 2H), 1.11 – 0.86 (m, 26H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 149.73 , 140.38 , 132.87 , 128.26 , 127.85 , 127.62 , 127.29 , 126.00 , 113.16 , 111.70 , 19.49 , 18.04 , 17.93 , 17.72 , 14.23 , 13.51 , 12.68 .

**_tert_-butyldimethyl((1-phenylbut-1-en-1-yl)oxy)silane (1p)**
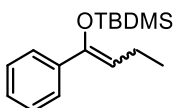


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.32 (d, *J* = 8.6 Hz, 1H), 7.25 – 7.21 (m, 1H), 6.80 – 6.72 (m, 2H), 5.22 – 4.81 (m, 1H), 2.14 (dt, *J* = 34.2, 7.5 Hz, 2H), 0.98 (d, *J* = 2.7 Hz, 12H), 0.19 (d, *J* = 4.9 Hz, 6H), 0.11 (s, 9H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 154.86 , 148.14 , 147.92 , 130.65 , 129.20 , 126.28 , 119.29 , 119.06 , 112.05 , 111.46 , 76.95 , 25.42 , 20.80 , 19.21 , 17.95 , 15.12 , 14.09 , 0.24 , 0.00 , -4.66 , -4.69 .

**General procedure C: preparation of diazoacetonitrile 2a:**
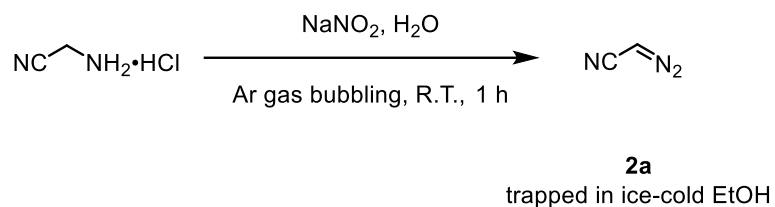
$$NC\diagup NH_2\cdot HCl \xrightarrow[\text{Ar gas bubbling, R.T., 1 h}]{\text{NaNO}_2,\ \text{H}_2\text{O}} NC\diagup N_2$$

**2a**
trapped in ice-cold EtOH

The stock solution of diazo compound **2** was prepared using a modified procedure reported by Hock et al.[7] To a vigorously stirring solution of 1.3 g (14 mmol) of aminoacetonitrile hydrochloride in 1 mL of water a solution of 0.96 g NaNO$_2$ in 2 mL of water was added slowly (1 hour) at room temperature using a syringe pump. The rapidly evolving yellow gas was carefully bubbled (carrier argon gas was used) via PTFE tubing into 7 mL of ethanol placed in a 7.5-mL sealed vial which was chilled on an ice bath. Upon the completion of addition, the aqueous mixture was further bubbled for 20 min to maximize the stock solution concentration and then the vial containing the diazo ethanol solution was carefully removed from the ice bath and stored at -20 °C. The substrate may deactivate/polymerize within two days to yield a black, insoluble mixture, therefore it needed to be prepared freshly. The concentration of the stock solution was measured by $^1$H NMR with non-deuterated DMSO as internal standard. The diazo solution concentration was then adjusted to 400 mM for biotransformations unless noted.

**CAUTION:** Diazo compounds are toxic and potentially explosive and should be handled with care in a well-ventilated hood. Do not attempt to concentrate or isolate.

## Synthesis of Reference Compounds

**General procedure D: preparation of α-branched ketones 3:**
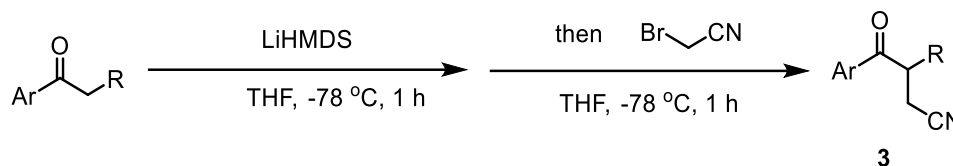


α-Branched ketones were prepared by using a modified procedure reported by Kleij et al.[8] A solution of lithium bis(trimethylsilyl)amide (LiHMDS, 1.0 M in THF, 12.0 mmol, 1.2 equiv.) in THF (10 mL) was cooled down to -78 ℃. A solution of ketone (10 mmol, 1.0 equiv.) in THF (10 mL) was added dropwise, and the resultant solution was stirred at -78 ℃ for 1 h. After that, a solution of bromoacetonitrile (12.0 mmol, 1.2 equiv) in THF (10 mL) was added dropwise and the reaction mixture was stirred at -78 ℃ for 1 h, quenched with hydrochloric acid (1 N), and extracted with Et₂O (3 × 20 mL). The combined organic layers were dried over anhydrous Na₂SO₄, filtered, and evaporated under reduced pressure. The crude product was then purified by flash chromatography affording product α-branched ketone **3**.

**3-benzoylpentanenitrile (3a)**
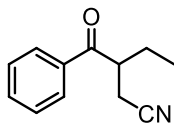


**¹H NMR** (400 MHz, Chloroform-*d*) δ 8.03 – 7.87 (m, 2H), 7.67 – 7.55 (m, 1H), 7.55 – 7.44 (m, 2H), 3.77 (ddd, *J* = 13.0, 6.9, 6.0 Hz, 1H), 2.74 (dd, *J* = 16.9, 6.8 Hz, 1H), 2.64 (dd, *J* = 16.9, 7.2 Hz, 1H), 1.99 – 1.88 (m, 1H), 1.85 – 1.70 (m, 1H), 0.91 (t, *J* = 7.5 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 199.91 , 135.63 , 133.86 , 129.04 , 128.48 , 118.75 , 43.95 , 25.25 , 18.00 , 10.49 .

**Physical State**: colorless oil.

**HRMS** (ESI): calcd for $C_{12}H_{14}NO$ [M+H]⁺ 188.1075; found 188.1080.

**3-(4-methylbenzoyl)pentanenitrile (3b)**



**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.92 – 7.76 (m, 2H), 7.39 – 7.26 (m, 2H), 3.74 (ddd, *J* = 13.1, 7.1, 6.1 Hz, 1H), 2.72 (dd, *J* = 16.9, 6.7 Hz, 1H), 2.63 (dd, *J* = 16.9, 7.4 Hz, 1H), 2.43 (s, 3H), 1.97 – 1.87 (m, 1H), 1.84 – 1.71 (m, 1H), 0.90 (t, *J* = 7.5 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 199.52 , 144.91 , 133.19 , 129.75 , 128.66 , 118.86, 43.83 , 25.41 , 21.84 , 18.13 , 10.55 .

**Physical State**: colorless oil.

**HRMS** (ESI): calcd for $C_{13}H_{16}NO$ [M+H]⁺ 202.1232; found 202.1215.

**3-(3-methylbenzoyl)pentanenitrile (3c)**



**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.78 – 7.68 (m, 2H), 7.46 – 7.34 (m, 2H), 3.80 – 3.72 (m, 1H), 2.73 (dd, *J* = 16.9, 6.7 Hz, 1H), 2.63 (dd, *J* = 17.0, 7.3 Hz, 1H), 1.99 – 1.85 (m, 1H), 1.83 – 1.71 (m, 1H), 0.90 (t, *J* = 7.5 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 200.14 , 138.96 , 135.73 , 134.66 , 129.00 , 128.89, 125.69 , 118.80 , 43.96 , 25.27 , 21.51 , 18.05 , 10.48 .

**Physical State**: colorless oil.
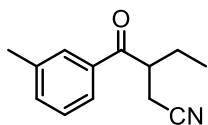
**HRMS** (ESI): calcd for $C_{13}H_{16}NO$ [M+H]⁺ 202.1232; found 202.1238.

**3-(4-isopropylbenzoyl)pentanenitrile (3d)**
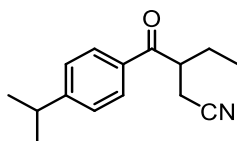


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.95 – 7.83 (m, 2H), 7.35 (d, *J* = 8.2 Hz, 2H), 3.76 (dd, *J* = 12.9, 6.3 Hz, 1H), 2.98 (p, *J* = 6.9 Hz, 1H), 2.79 – 2.57 (m, 2H), 1.93 (dtd, *J* = 15.0, 7.5, 5.9 Hz, 1H), 1.85 – 1.72 (m, 1H), 1.28 (d, *J* = 6.9 Hz, 6H), 0.91 (t, *J* = 7.5 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 199.36 , 155.45 , 133.36 , 128.68 , 127.04 , 118.74 , 43.71 , 34.31 , 29.72 , 25.25 , 23.62 , 18.01 , 10.44 .

**Physical State**: colorless oil.

**3-(4-fluorobenzoyl)pentanenitrile (3e)**



**¹H NMR** (400 MHz, Chloroform-*d*) δ 8.03 – 7.93 (m, 2H), 7.22 – 7.12 (m, 2H), 3.78 – 3.67 (m, 1H), 2.74 (dd, *J* = 16.9, 6.9 Hz, 1H), 2.64 (dd, *J* = 16.9, 7.1 Hz, 1H), 2.01 – 1.85 (m, 1H), 1.84 – 1.70 (m, 1H), 0.91 (t, *J* = 7.5 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 198.34 , 167.53 , 164.98 , 132.09 (d, *J* = 2.9 Hz), 131.24 (d, *J* = 9.4 Hz), 118.65 , 116.28 (d, *J* = 22.0 Hz), 44.00 , 25.38 , 18.10 , 10.54 .

**¹⁹F NMR** (377 MHz, Chloroform-*d*) δ -103.74 .

**Physical State**: colorless oil.

**HRMS (**ESI): calcd for $C_{12}H_{13}NOF$ [M+H]⁺ 206.0981; found 206.0970.

**3-(2-fluorobenzoyl)pentanenitrile (3f)**



**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.84 (td, *J* = 7.6, 1.9 Hz, 1H), 7.57 (dddd, *J* = 8.3, 7.1, 5.1, 1.9 Hz, 1H), 7.29 – 7.25 (m, 1H), 7.16 (ddd, *J* = 11.4, 8.3, 1.1 Hz, 1H), 3.68 (qd, *J* = 6.7, 5.2 Hz, 1H), 2.77 (ddd, *J* = 16.9, 6.7, 1.0 Hz, 1H), 2.63 (ddd, *J* = 16.9, 6.8, 0.9 Hz, 1H), 1.99 – 1.90 (m, 1H), 1.71 (dddd, *J* = 14.2, 7.6, 6.7, 0.9 Hz, 1H), 0.93 (t, *J* = 7.5 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 198.30 (d, *J* = 4.2 Hz), 161.52 (d, *J* = 253.8 Hz), 135.32 (d, *J* = 9.2 Hz), 131.27 (d, *J* = 2.5 Hz), 125.08 (d, *J* = 3.4 Hz), 124.59 (d, *J* = 12.5

Hz), 118.67 , 116.90 (d, $J$ = 23.9 Hz), 48.51 (d, $J$ = 7.6 Hz), 24.24 (d, $J$ = 1.4 Hz), 17.46

, 10.55 .

**¹⁹F NMR** (377 MHz, Chloroform-$d$) δ -110.81 .

**Physical State**: colorless oil.

**HRMS (**ESI): calcd for $C_{12}H_{13}NOF$ [M+H]⁺ 206.0981; found 206.0964.

**3-(4-chlorobenzoyl)pentanenitrile (3g)**



**¹H NMR** (400 MHz, Chloroform-$d$) δ 7.92 – 7.74 (m, 2H), 7.64 – 7.36 (m, 2H), 3.77 – 3.66

(m, 1H), 2.73 (dd, $J$ = 17.0, 6.9 Hz, 1H), 2.64 (dd, $J$ = 16.9, 7.1 Hz, 1H), 1.98 – 1.84 (m, 1H),

1.80 – 1.70 (m, 1H), 0.91 (t, $J$ = 7.5 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-$d$) δ 198.75 , 140.50 , 133.96 , 129.92 , 129.44 , 118.58 ,

44.05 , 25.31 , 18.02 , 10.51 .

**Physical State**: colorless oil.

**HRMS (**ESI): calcd for $C_{12}H_{13}NOCl$ [M+H]⁺ 222.0686; found 222.0685.

**3-(4-bromobenzoyl)pentanenitrile (3h)**



**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.90 – 7.68 (m, 2H), 7.79 – 7.52 (m, 2H), 3.75 – 3.66 (m, 1H), 2.74 (dd, *J* = 16.9, 6.9 Hz, 1H), 2.64 (dd, *J* = 16.9, 7.1 Hz, 1H), 1.98 – 1.88 (m, 1H), 1.80 – 1.73 (m, 1H), 0.91 (t, *J* = 7.5 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 198.83 , 134.24 , 132.32 , 129.87 , 129.14 , 118.43 , 77.22 , 43.91 , 25.17 , 17.88 , 10.37 .

**Physical State**: colorless oil.

**HRMS** (ESI): calcd for $C_{12}H_{13}NOBr$ [M+H]$^+$ 266.0181; found 266.0200.
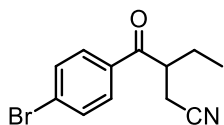
**3-(4-methoxybenzoyl)pentanenitrile (3i)**



**¹H NMR** (400 MHz, Chloroform-*d*) δ 8.11 – 7.89 (m, 2H), 6.97 (dd, *J* = 9.0, 2.4 Hz, 2H), 3.88 (s, 1H), 3.78 – 3.66 (m, 1H), 2.72 (ddd, *J* = 16.9, 6.8, 2.2 Hz, 1H), 2.63 (ddd, *J* = 17.0, 7.3, 2.3 Hz, 1H), 1.98 – 1.86 (m, 1H), 1.85 – 1.71 (m, 1H), 0.91 (td, *J* = 7.5, 2.2 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 198.33 , 164.17 , 130.90 , 128.67 , 118.93 , 114.23 , 55.70 , 43.59 , 25.57 , 18.28 , 10.63 .

**Physical State**: colorless oil.

**HRMS** (ESI): calcd for $C_{13}H_{16}NO_2$ [M+H]$^+$ 218.1181; found 218.1163.
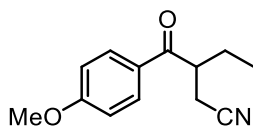
### 3-(thiophene-2-carbonyl)pentanenitrile (3j)



**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.74 (ddd, *J* = 15.2, 4.4, 1.1 Hz, 2H), 7.18 (dd, *J* = 5.0, 3.8 Hz, 1H), 3.62 – 3.50 (m, 1H), 2.64 (dd, *J* = 17.0, 7.4 Hz, 1H), 1.96 (ddd, *J* = 14.3, 7.6, 6.8 Hz, 1H), 1.84 (ddd, *J* = 12.4, 7.5, 6.0 Hz, 1H), 0.95 (t, *J* = 7.5 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 192.53 , 143.05 , 135.07 , 132.60 , 128.47 , 118.34 , 99.98 , 77.22 , 45.58 , 25.89 , 18.48 , 10.79 .

**Physical State**: colorless oil.

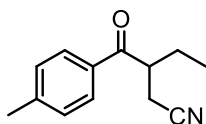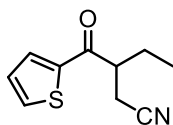**HRMS** (ESI): calcd for $C_{10}H_{12}NOS$ [M+H]⁺ 194.0640; found 194.0646.

### 3-benzoylhexanenitrile (3k)
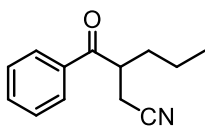


**¹H NMR** (400 MHz, Chloroform-*d*) δ 8.01 – 7.87 (m, 2H), 7.67 – 7.57 (m, 1H), 7.55 – 7.46 (m, 2H), 3.83 – 3.77 (m, 1H), 2.75 (dd, *J* = 16.9, 6.8 Hz, 1H), 2.63 (dd, *J* = 16.9, 6.9 Hz, 1H), 1.91 – 1.77 (m, 1H), 1.72 – 1.60 (m, 1H), 1.41 – 1.26 (m, 2H), 0.91 (t, *J* = 7.3 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 200.12 , 135.69 , 133.91 , 129.09 , 128.53 , 118.76 , 42.78 , 34.47 , 19.78 , 18.65 , 14.09 .

**Physical State**: colorless oil.

**HRMS** (ESI): calcd for $C_{13}H_{16}NO$ [M+H]⁺ 202.1232; found 202.1257.

**3-methyl-4-oxo-4-phenylbutanenitrile (3l)**
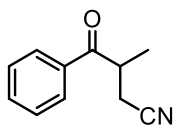


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.91 – 7.83 (m, 2H), 7.59 – 7.50 (m, 1H), 7.48 – 7.38 (m, 2H), 3.81 – 3.67 (m, 1H), 2.68 (dd, *J* = 16.9, 5.8 Hz, 1H), 2.55 (dd, *J* = 16.9, 8.1 Hz, 1H), 1.33 (d, *J* = 7.2 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 199.90 , 134.69 , 133.82 , 128.97 , 128.48 , 118.61 , 77.23 , 38.11 , 20.40 , 18.12 .

**Physical State**: colorless oil.

**HRMS** (ESI): calcd for $C_{11}H_{12}NO$ [M+H]$^+$ 174.0919; found 174.0914.

**3-methyl-4-oxo-4-(p-tolyl)butanenitrile (3m)**



**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.84 (d, *J* = 8.3 Hz, 2H), 7.36 – 7.26 (m, 2H), 3.83 – 3.74 (m, 1H), 2.73 (dd, *J* = 16.9, 5.8 Hz, 1H), 2.61 (dd, *J* = 16.9, 8.2 Hz, 1H), 2.43 (s, 3H), 1.39 (d, *J* = 7.2 Hz, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 199.63 , 144.95 , 132.31 , 129.78 , 128.74 , 118.84 , 38.08 , 21.85 , 20.55 , 18.34 .

**Physical State**: colorless oil.

**HRMS** (ESI): calcd for $C_{12}H_{14}NO$ [M+H]$^+$ 188.1075; found 188.1080.

## HPLC Calibration Curves

Calibration curves of synthesized reference compounds were created for the determination of yield and TTN. For each substrate, four different concentrations of product (1, 2, 4, and 6 mM) with 2 mM internal standard in 375-µL acetonitrile solutions were mixed each with 375 µL water. The mixtures were vortexed and then analyzed by HPLC based on UV absorbance at 210 nm. All data points represent the average of duplicate runs. The calibration curves depict the ratio of product area to internal standard area (x-axis) against product concentration in mM (y-axis). Notes: Pdt = product area, IS = internal standard area, [Pdt] = product concentration in reaction, [PC] = protein concentration in reaction, Avg. TTN = average total turnover number, SD TTN = standard deviation of TTN, Avg. Yield = average yield, SD Yield = standard deviation of yield.

**Note:** Data points represent average of duplicate runs. Pdt, product. IS, internal standard. PC, protein concentration. SD, standard deviation. Avg., average.

**3-benzoylpentanenitrile (3a)**
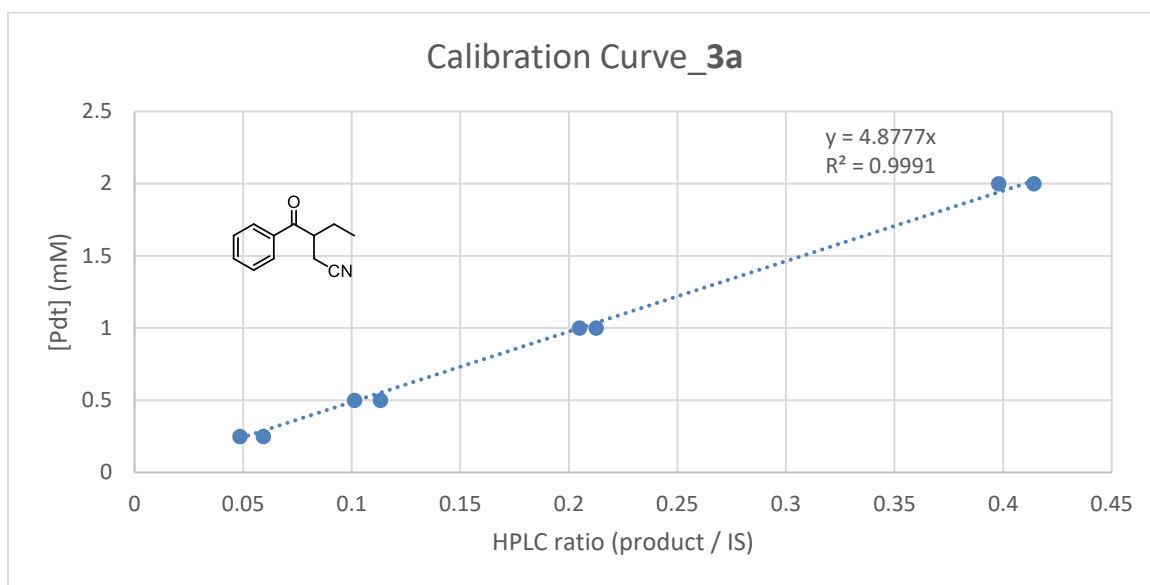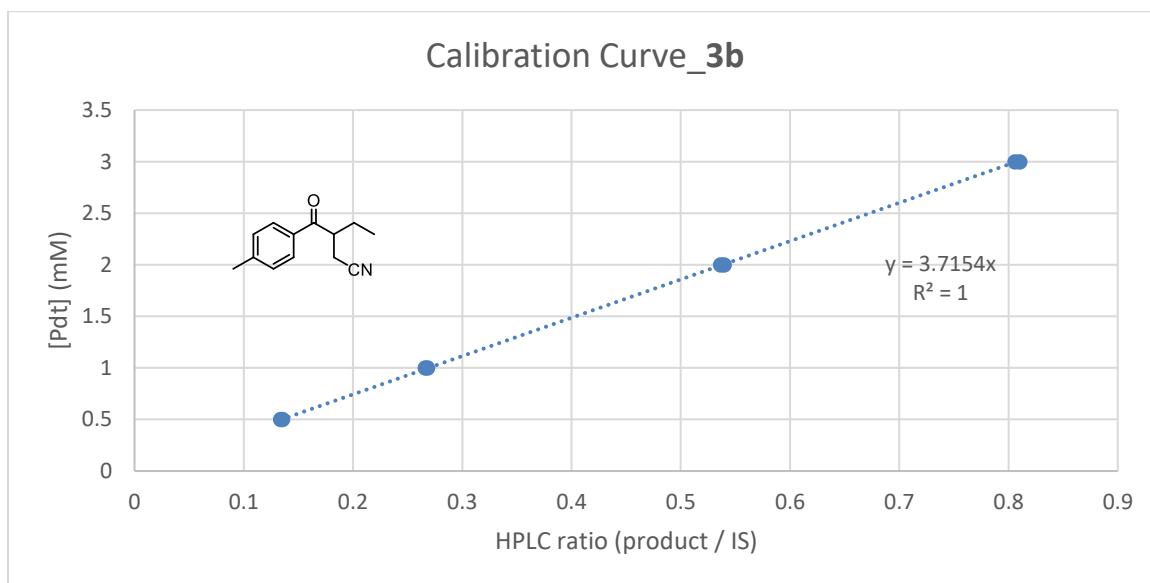


**Calibration Curve_3a**

y = 4.8777x
R² = 0.9991

[Pdt] (mM)

HPLC ratio (product / IS)

**3-(4-methylbenzoyl)pentanenitrile (3b)**



**Calibration Curve_3b**

y = 3.7154x
R² = 1

[Pdt] (mM)

HPLC ratio (product / IS)

**3-(3-methylbenzoyl)pentanenitrile (3c)**



Calibration Curve_**3c**

y = 2.4634x
R² = 0.9998

[Pdt] (mM)

HPLC ratio (product / IS)

**3-(4-isopropylbenzoyl)pentanenitrile (3d)**



Calibration Curve_**3d**

y = 3.5494x
R² = 0.9998

[Pdt] (mM)

HPLC ratio (product / IS)
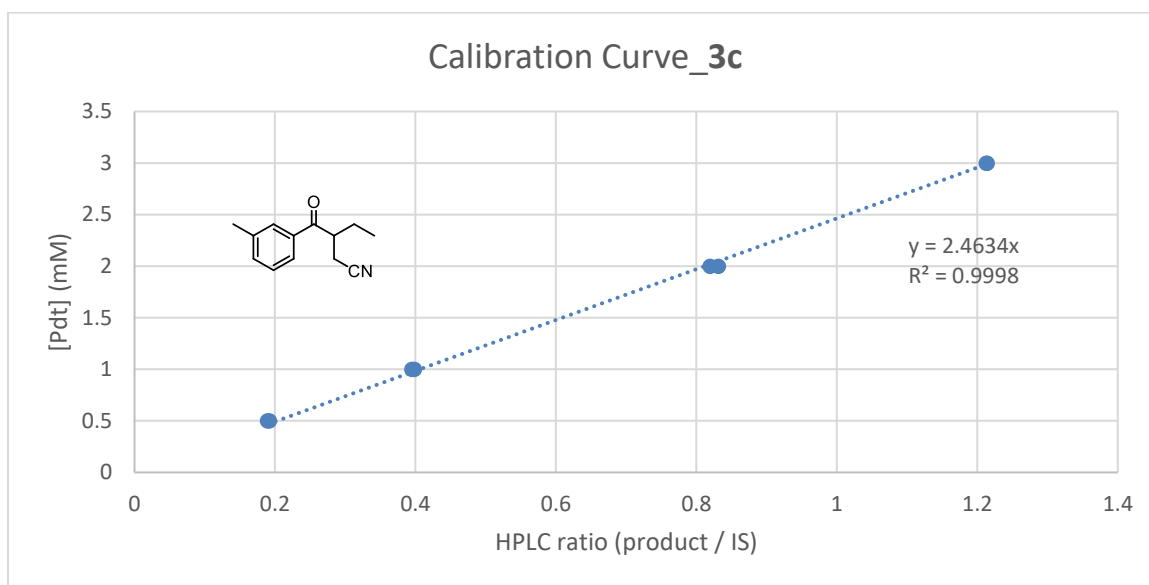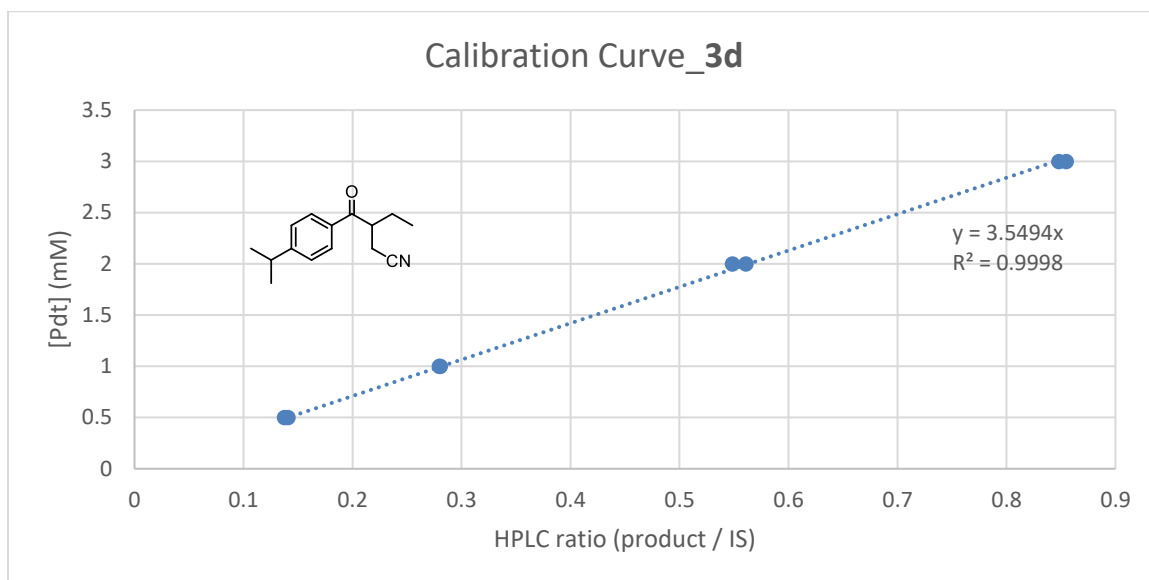
**3-(4-fluorobenzoyl)pentanenitrile (3e)**
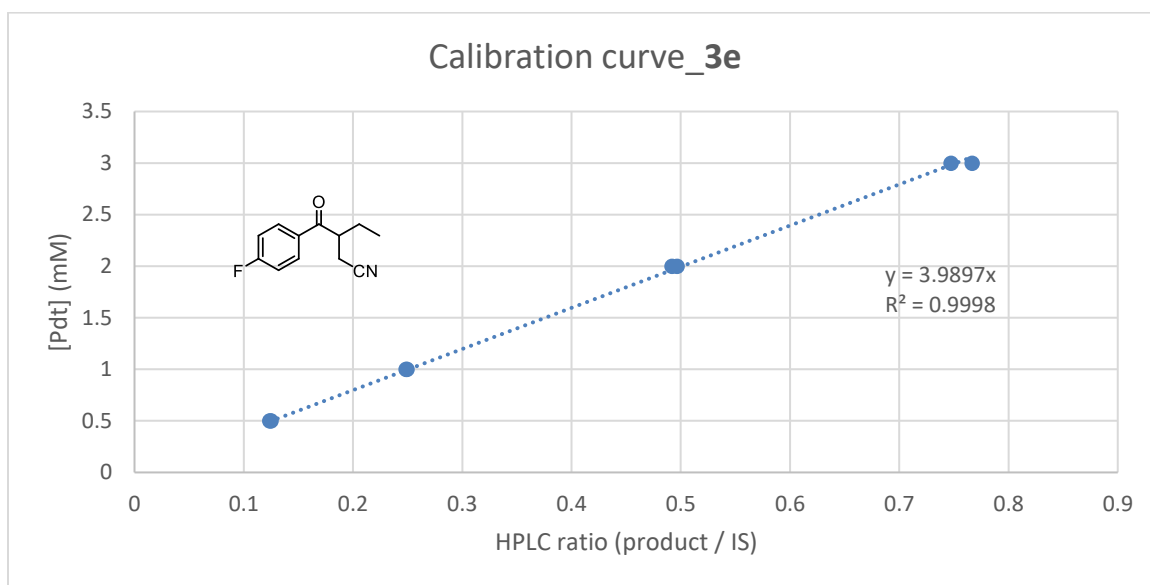


**3-(2-fluorobenzoyl)pentanenitrile (3f)**
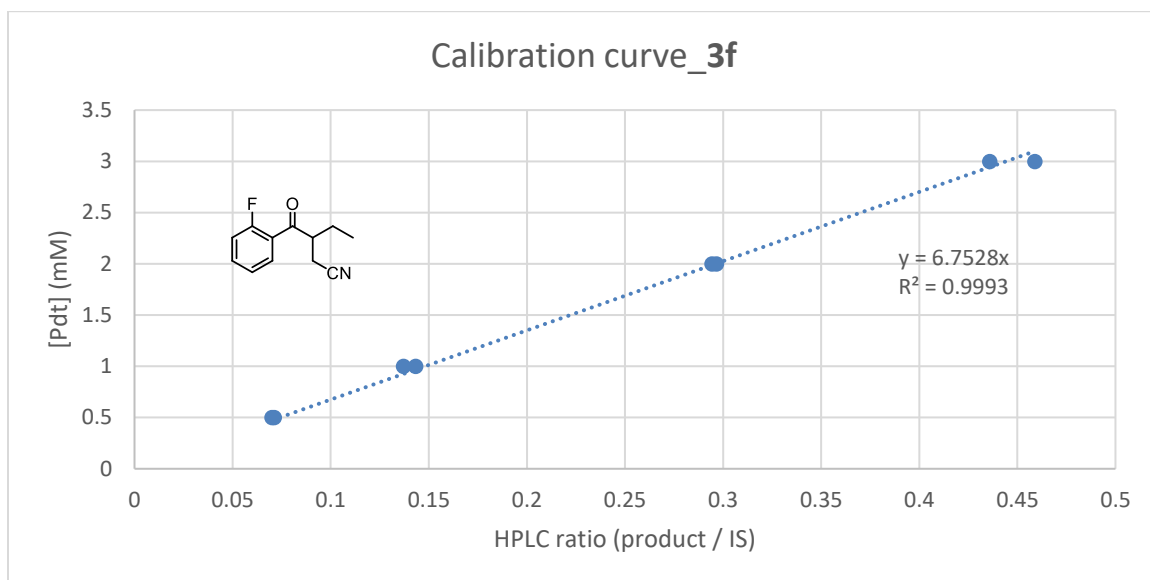
**3-(4-chlorobenzoyl)pentanenitrile (3g)**
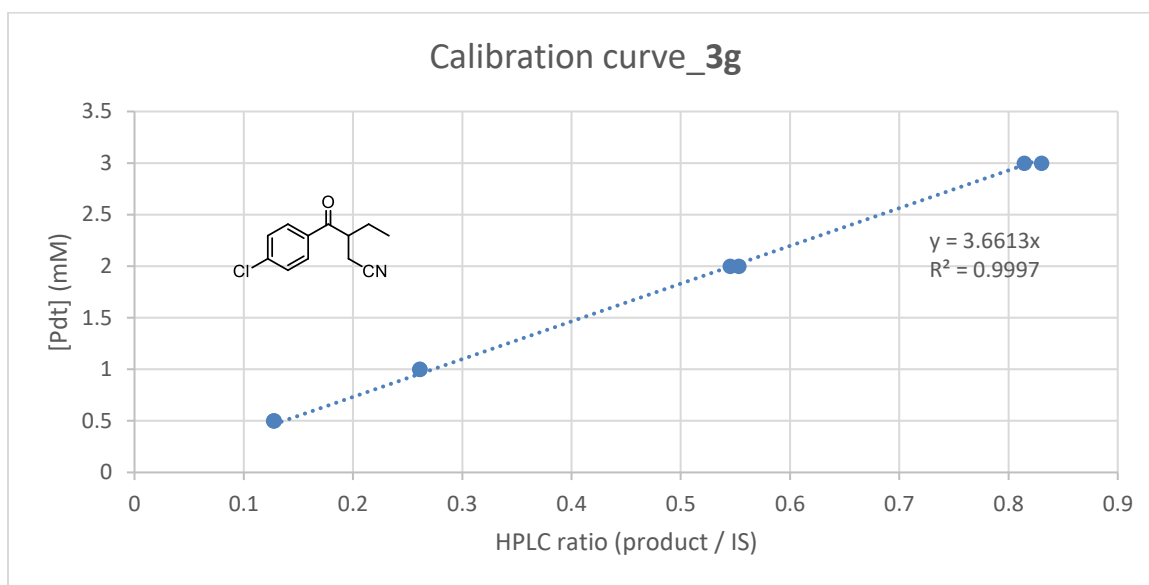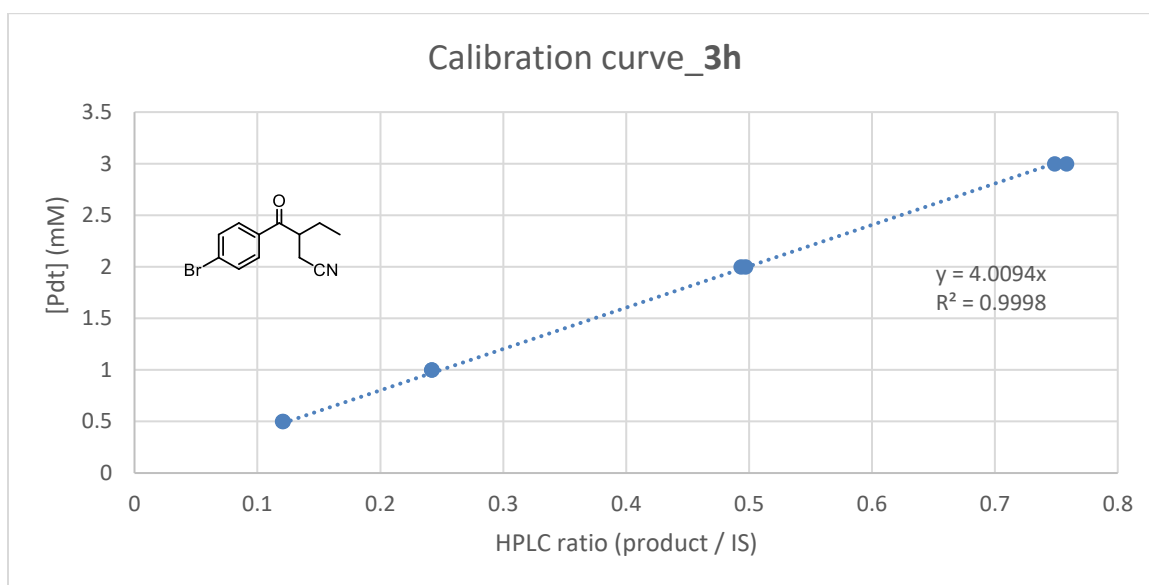


Calibration curve_**3g**

y = 3.6613x
R² = 0.9997

**3-(4-bromobenzoyl)pentanenitrile (3h)**



Calibration curve_**3h**

y = 4.0094x
R² = 0.9998

**3-(4-methoxybenzoyl)pentanenitrile (3i)**



Calibration curve_**3i**

y = 6.647x
R² = 0.9999

**3-(thiophene-2-carbonyl)pentanenitrile (3j)**



Calibration curve_**3j**

y = 31.114x
R² = 0.9995

**3-benzoylhexanenitrile (3k)**



Calibration curve_**3k**

y = 5.2528x
R² = 0.9999

[Pdt] (mM) vs HPLC ratio (product / IS)

**3-methyl-4-oxo-4-phenylbutanenitrile (3l)**



Calibration curve_**3l**

y = 4.9891x
R² = 0.9999

[Pdt] (mM) vs HPLC ratio (product / IS)

**3-methyl-4-oxo-4-(p-tolyl)butanenitrile (3m)**



Calibration curve_**3m**

y = 3.2708x
R² = 0.9996

[Pdt] (mM)

HPLC ratio (product / IS)
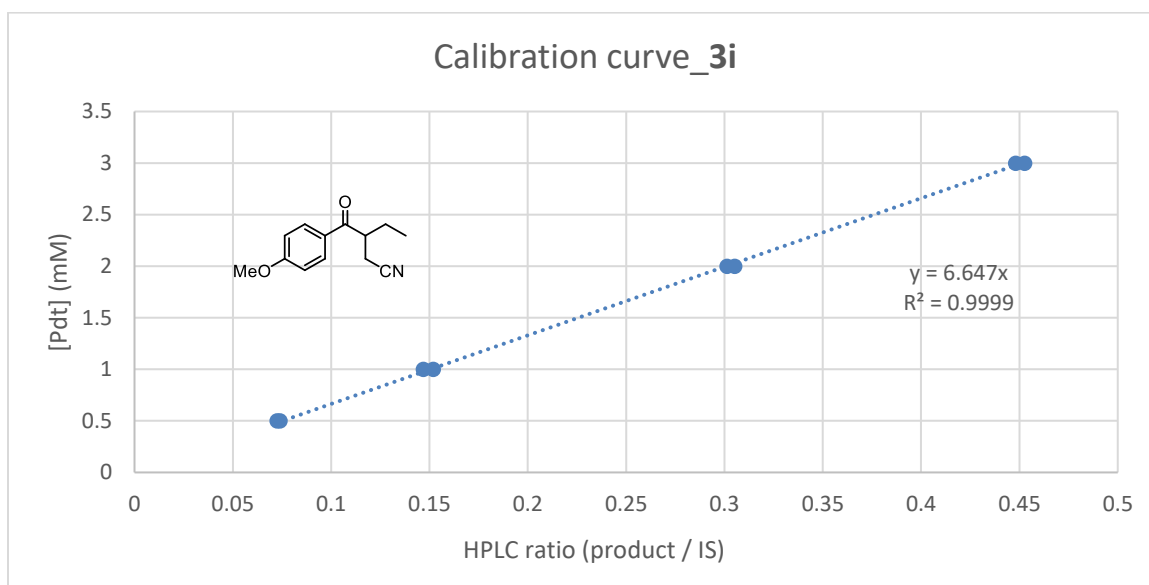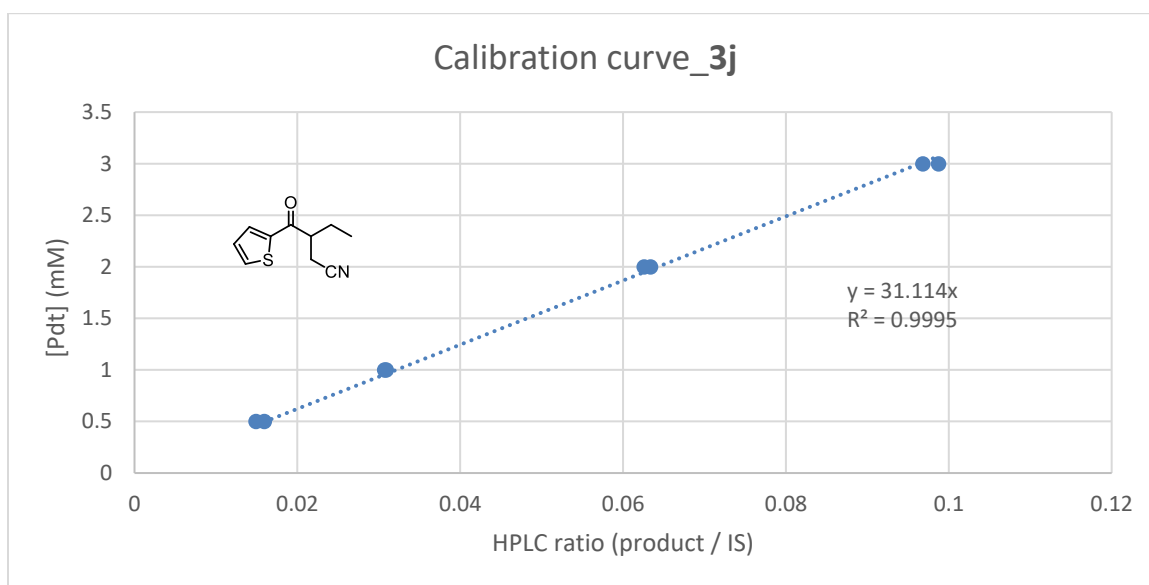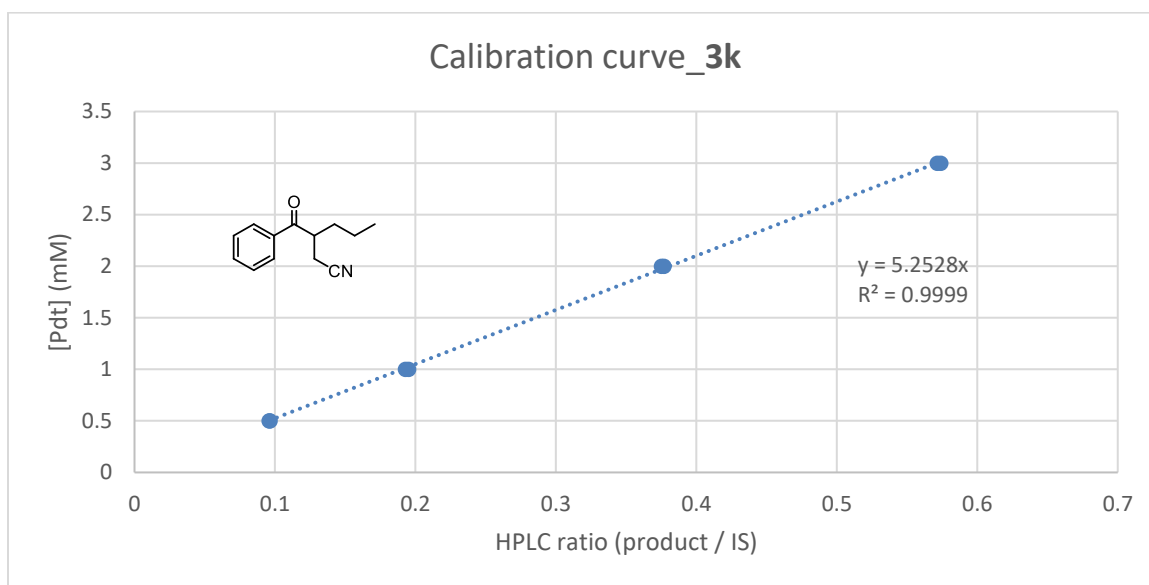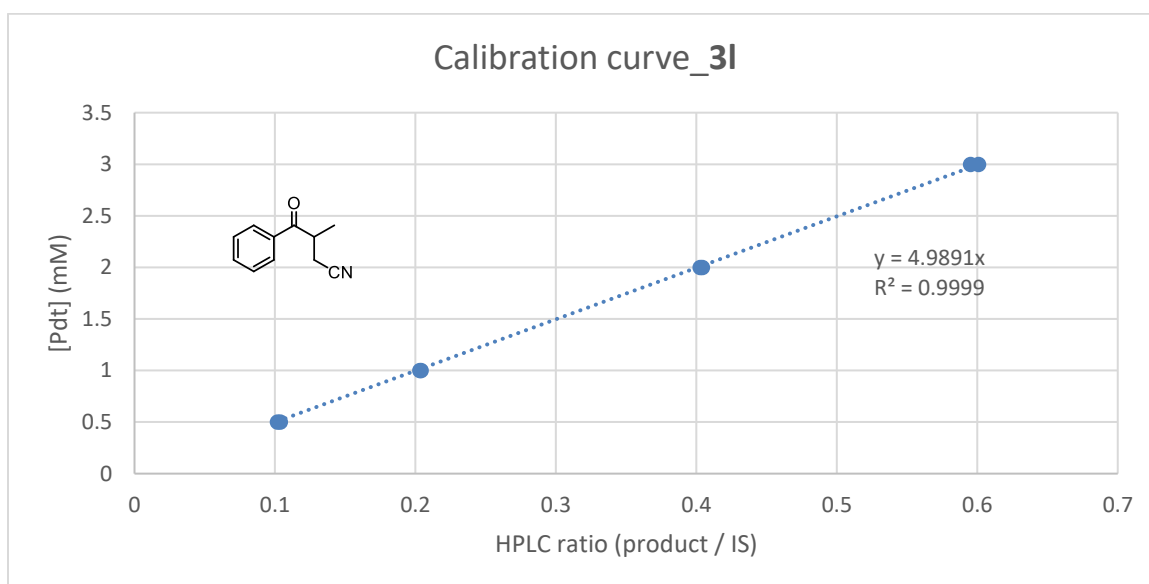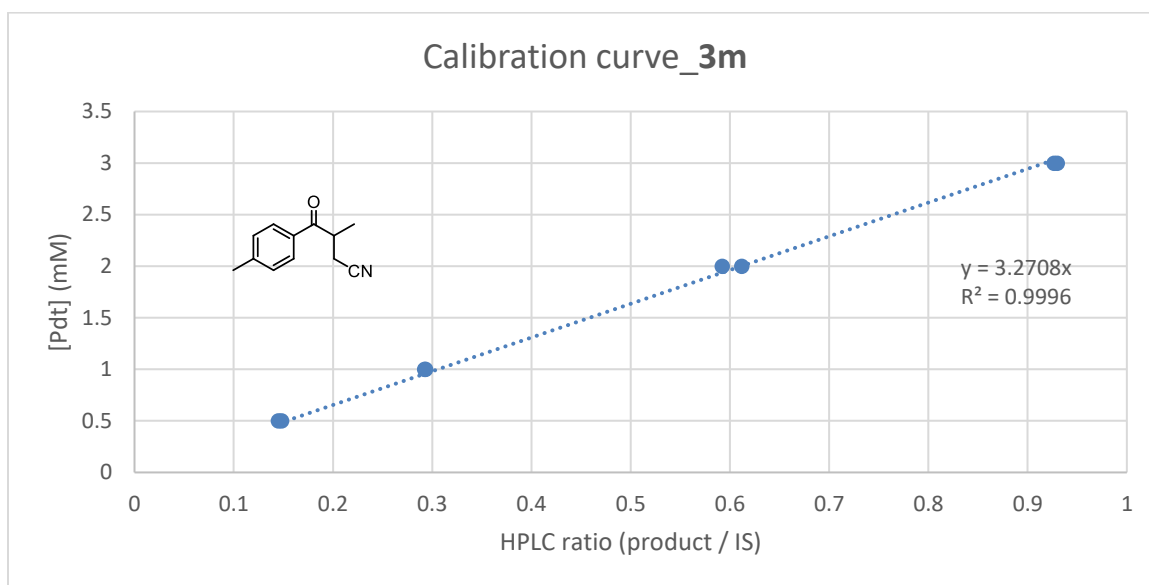
## Enantioselectivity Determinations

**Enantioselectivity determinations of *α*-branched ketone products 3**

The absolute configuration of enzymatic product **3l** was assigned to be *S* by comparing the elution order of two enantiomers with a literature report under same elution conditions using the same column (Chiralpak IA).[1] The other products **3a**–**3m** were assigned by analogy.

**Note:** For some samples, we found the retention time drifts in normal-phase HPLC.

**(*S*)-3-benzoylpentanenitrile (3a)**



**Chiral-phase HPLC conditions:** Chiralpak IB, 15% *i*-PrOH in hexane, 1.0 mL/min, 25

ºC, 254 nm

**Racemic 3a:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 8.397 | 2567.3 | 251.1 | 0.1553 | 50.594 | 0.712 |
| 2 | 10.613 | 2507 | 193.7 | 0.2157 | 49.406 | 0.792 |
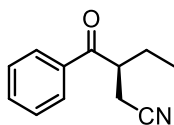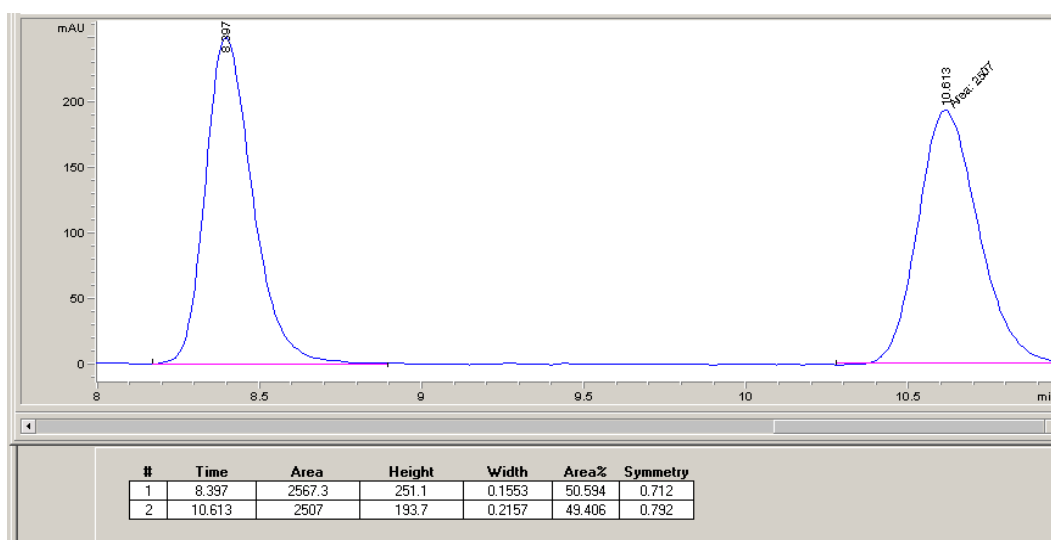
**Enzymatic preparation of 3a with** P411-SCA-5188**:** 99% e.e.
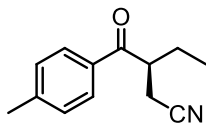


| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 8.426 | 10.8 | 1.7 | 0.1032 | 0.500 | 1.051 |
| 2 | 10.555 | 2157.5 | 174.4 | 0.1924 | 99.500 | 0.791 |

**(*S*)-3-(4-methylbenzoyl)pentanenitrile (3b)**



**Chiral-phase HPLC conditions:** Chiralpak IA, 5% *i*-PrOH in hexane, 1.0 mL/min, 25 ℃,

245 nm

**Racemic 3b:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 10.175 | 2387.8 | 169 | 0.2355 | 49.618 | 0.682 |
| 2 | 10.723 | 2424.5 | 162.4 | 0.2488 | 50.382 | 0.731 |
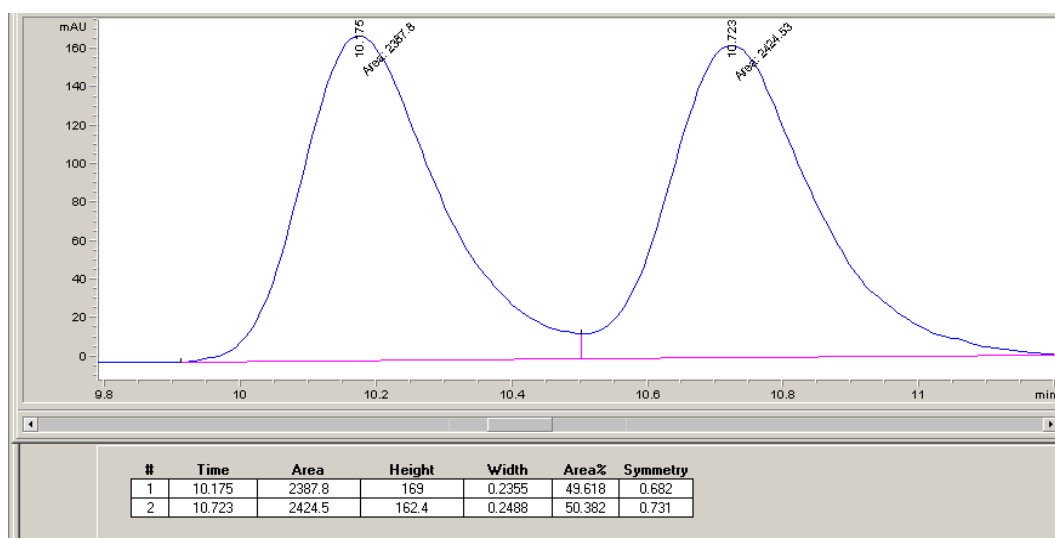
**Enzymatic preparation of 3b with** P411-SCA-5188**:** 97% e.e.
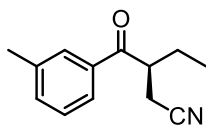


| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 10.187 | 32.3 | 2.9 | 0.1836 | 1.450 | 2.04 |
| 2 | 10.701 | 2198.6 | 153.1 | 0.2394 | 98.550 | 0.714 |

**(*S*)-3-(3-methylbenzoyl)pentanenitrile (3c)**



**Chiral-phase HPLC conditions:** Chiralpak IB, 15% *i*-PrOH in hexane, 1.0 mL/min, 25

ºC, 254 nm

**Racemic 3c:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 6.825 | 8059.9 | 965.7 | 0.1391 | 49.376 | 0.744 |
| 2 | 8.082 | 8263.6 | 849.2 | 0.1622 | 50.624 | 0.74 |
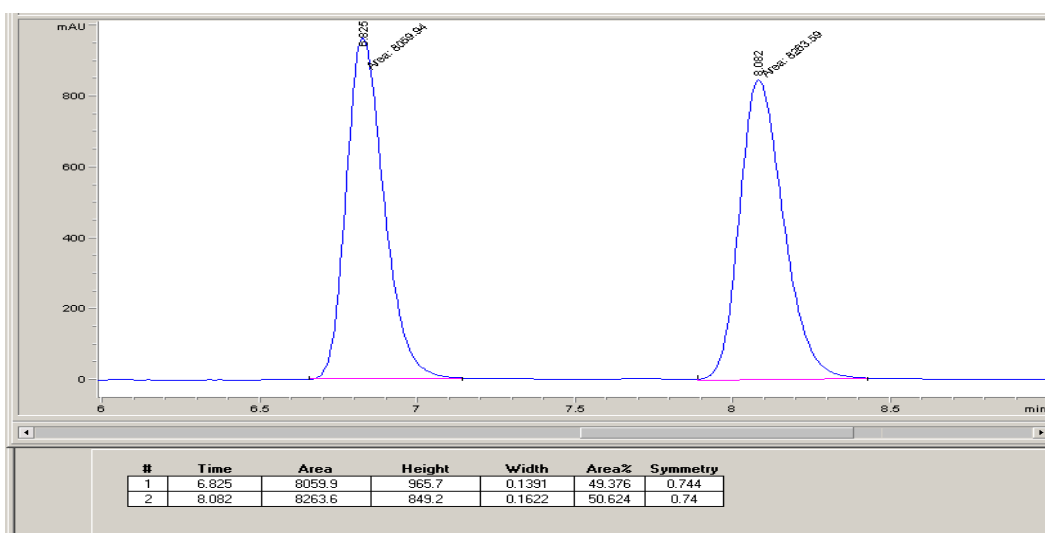
**Enzymatic preparation of 3c with** P411-SCA-5188**:** 97% e.e.
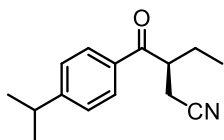


| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 6.851 | 80.5 | 12.9 | 0.1036 | 1.389 | 1.138 |
| 2 | 8.08 | 5712.8 | 596.3 | 0.1597 | 98.611 | 0.788 |

**(*S*)-3-(4-isopropylbenzoyl)pentanenitrile (3d)**



**Chiral-phase HPLC conditions:** Chiralpak IA, 5% *i*-PrOH in hexane, 1.0 mL/min, 25 ºC,

254 nm

**Racemic 3d:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 8.802 | 4778.7 | 378.9 | 0.2102 | 50.088 | 0.71 |
| 2 | 10.639 | 4761.9 | 312.4 | 0.254 | 49.912 | 0.665 |
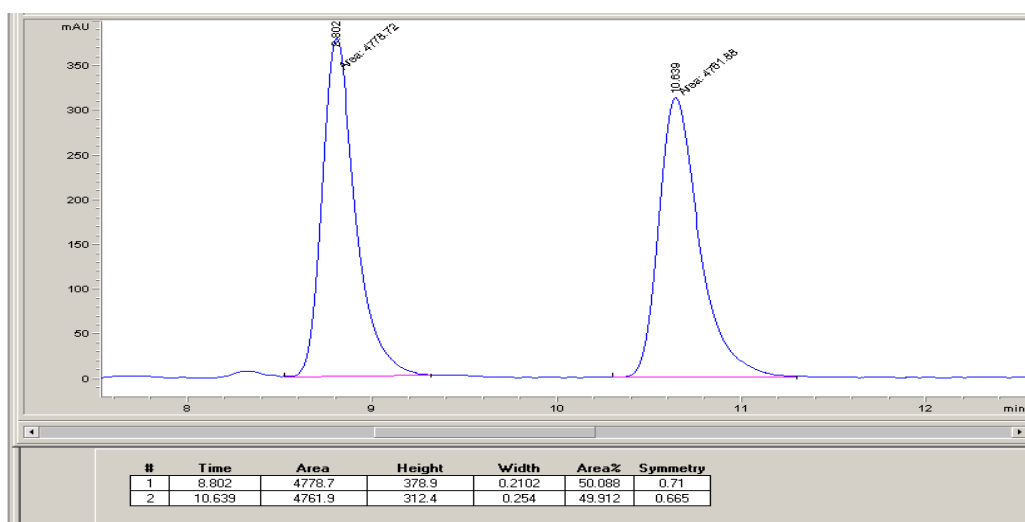
**Enzymatic preparation of 3d with** P411-SCA-5188**:** 95% e.e.
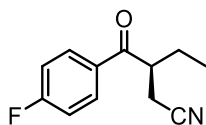


| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 8.829 | 33.1 | 3.3 | 0.1683 | 2.585 | 1.217 |
| 2 | 10.624 | 1247.6 | 81.3 | 0.2557 | 97.415 | 0.721 |

**(*S*)-3-(4-fluorobenzoyl)pentanenitrile (3e)**



**Chiral-phase HPLC conditions:** Chiralpak IA, 5% *i*-PrOH in hexane, 1.0 mL/min, 25 ºC,

254 nm

**Racemic 3e:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 11.127 | 3664 | 237.8 | 0.2568 | 49.197 | 0.668 |
| 2 | 11.769 | 3783.6 | 223.4 | 0.2823 | 50.803 | 0.65 |
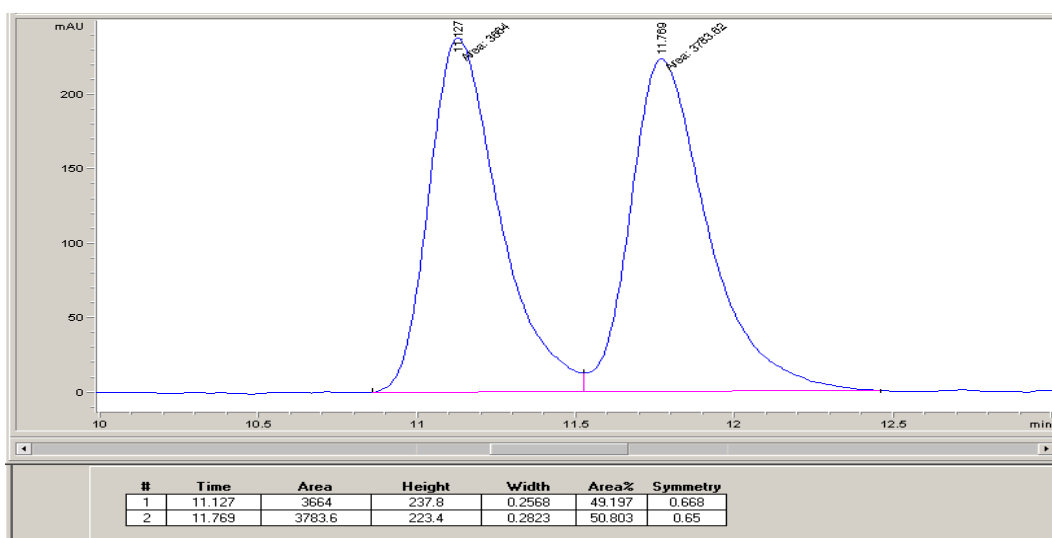
**Enzymatic preparation of 3e with** P411-SCA-5188**:** 93% e.e.
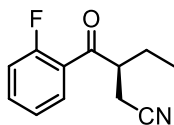


| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 11.099 | 66.2 | 5.2 | 0.2138 | 3.422 | 0.703 |
| 2 | 11.735 | 1867.4 | 114.2 | 0.2725 | 96.578 | 0.669 |

**(*S*)-3-(2-fluorobenzoyl)pentanenitrile (3f)**



**Chiral-phase HPLC conditions:** Chiralpak IA, 5% *i*-PrOH in hexane, 1.0 mL/min, 25 ℃,

245 nm

**Racemic 3f:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 8.305 | 294.3 | 26.8 | 0.1831 | 50.552 | 0.747 |
| 2 | 9.088 | 287.8 | 24.5 | 0.1955 | 49.448 | 0.769 |
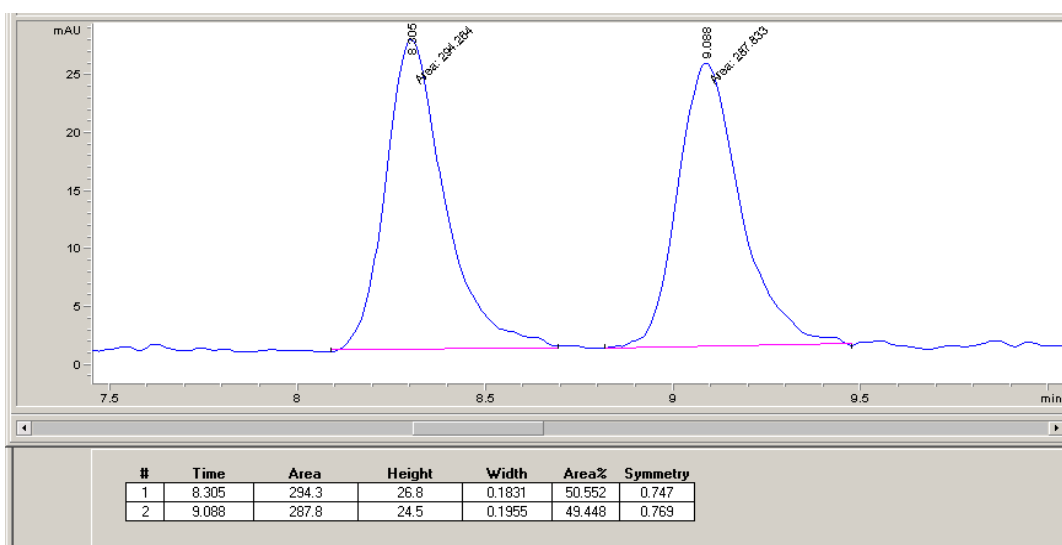
**Enzymatic synthesis of 3f with** P411-SCA-5188**: 99**% e.e.
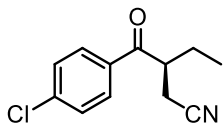


| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 9.137 | 475.7 | 33.4 | 0.2377 | 100.000 | 0.715 |

**(*S*)-3-(4-chlorobenzoyl)pentanenitrile (3g)**



**Chiral-phase HPLC conditions:** Chiralpak IA, 5% *i*-PrOH in hexane, 1.0 mL/min, 25 ºC,

245 nm

**Racemic 3g:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 11.953 | 6483.5 | 387.9 | 0.2786 | 48.941 | 0.64 |
| 2 | 12.588 | 6763.9 | 369.6 | 0.305 | 51.059 | 0.636 |
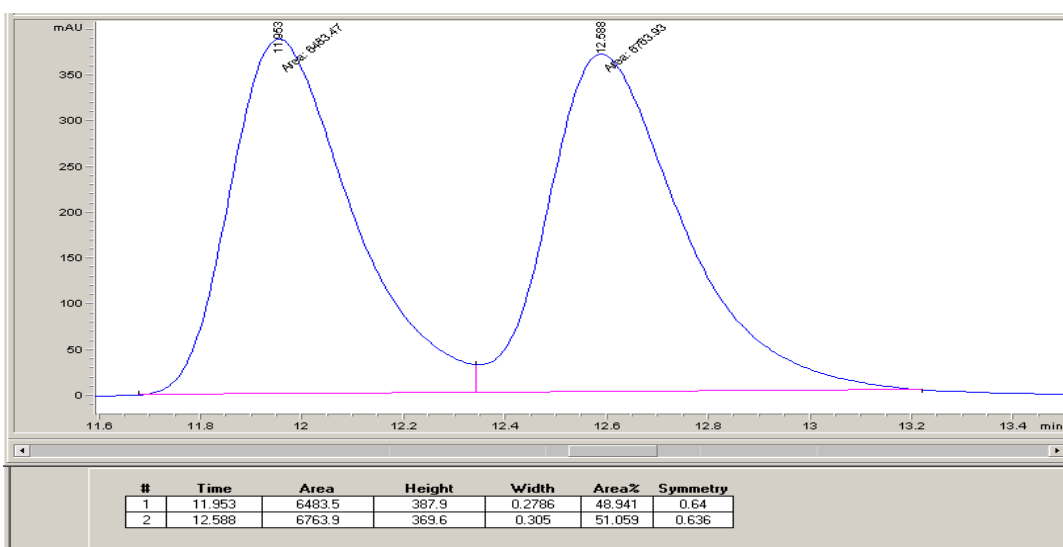
**Enzymatic preparation of 3g with** P411-SCA-5188**:** 90% e.e.
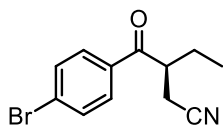


| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 12.013 | 27.4 | 2.3 | 0.1992 | 4.982 | 0.951 |
| 2 | 12.644 | 523.4 | 29.7 | 0.2937 | 95.018 | 0.729 |

**(*S*)-3-(4-bromobenzoyl)pentanenitrile (3h)**



**Chiral-phase HPLC conditions:** Chiralpak IA, 5% *i*-PrOH in hexane, 1.0 mL/min, 25 ℃,

254 nm

**Racemic 3h:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 13.077 | 6160.2 | 309.7 | 0.3315 | 49.863 | 0.614 |
| 2 | 14.412 | 6194 | 237.5 | 0.4347 | 50.137 | 0.535 |
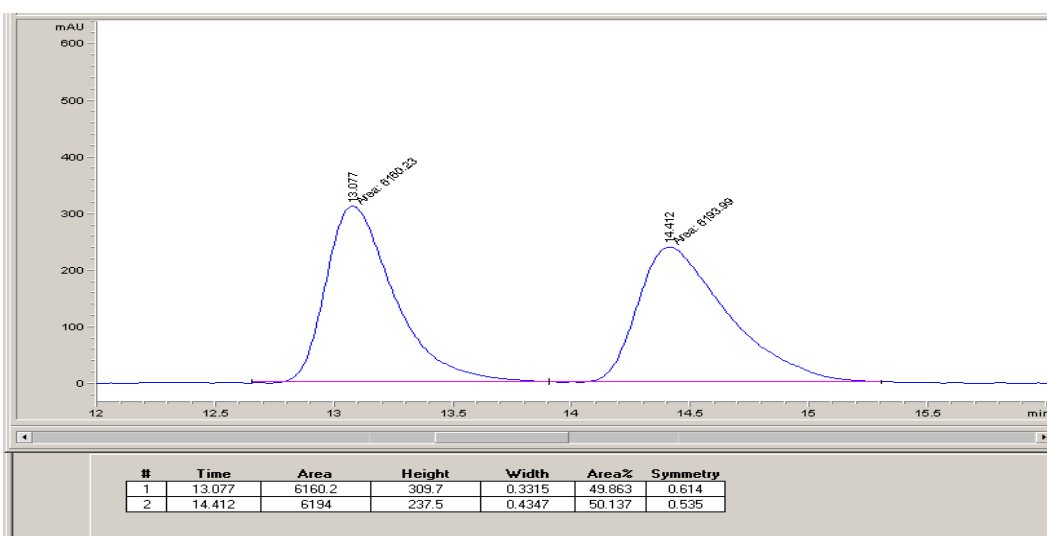
**Enzymatic preparation of 3h with** P411-SCA-5188**:** >99% e.e.
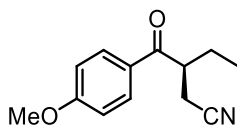


| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 13.204 | 29.9 | 2.5 | 0.2024 | 0.475 | 4.858 |
| 2 | 14.396 | 6260.4 | 280.9 | 0.3714 | 99.525 | 0.595 |

**(*S*)-3-(4-methoxybenzoyl)pentanenitrile (3i)**



**Chiral-phase HPLC conditions:** Chiralpak IA, 5% *i*-PrOH in hexane, 1.0 mL/min, 25 ℃,

254 nm

**Racemic 3i:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 17.24 | 1805.1 | 70.3 | 0.4277 | 50.164 | 0.558 |
| 2 | 19.427 | 1793.3 | 63.2 | 0.4731 | 49.836 | 0.671 |
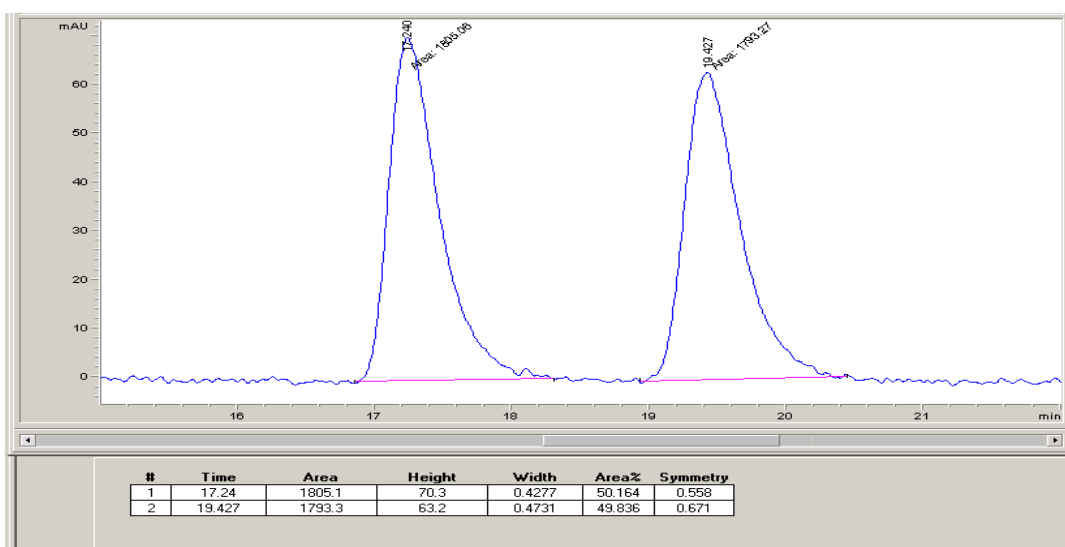
**Enzymatic preparation of 3i with** P411-SCA-5188**:** >99% e.e.
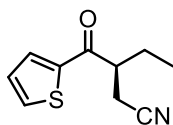


| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 17.399 | 22.6 | 1.5 | 0.2531 | 0.401 | 7.379 |
| 2 | 19.378 | 5607.9 | 192.3 | 0.486 | 99.599 | 0.567 |

**(*S*)-3-(thiophene-2-carbonyl)pentanenitrile (3j)**



**Chiral-phase HPLC conditions:** Chiralpak IB, 15% *i*-PrOH in hexane, 1.0 mL/min, 25

ºC, 254 nm

**Racemic 3j:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 10.729 | 2019.8 | 150.6 | 0.2235 | 49.990 | 0.72 |
| 2 | 11.703 | 2020.6 | 141.7 | 0.2376 | 50.010 | 0.725 |
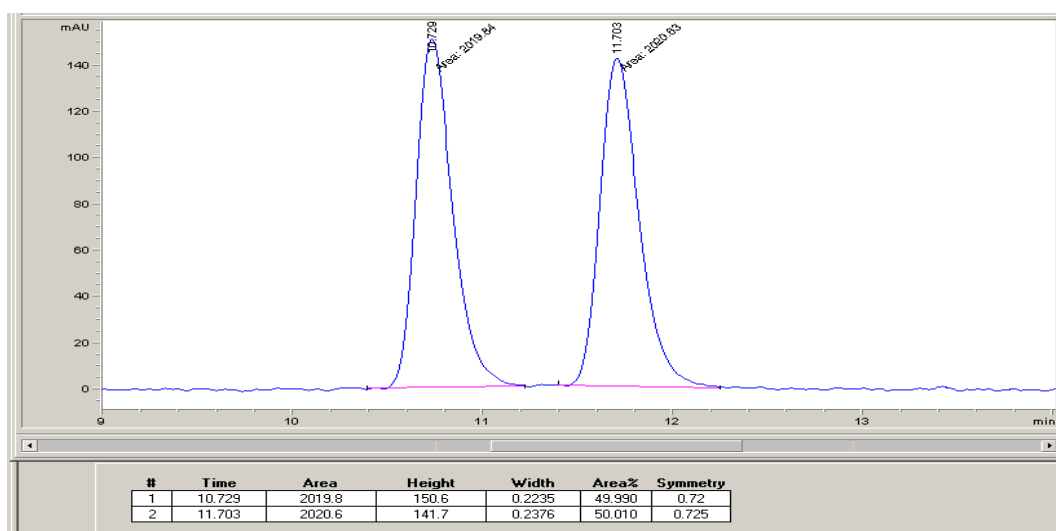
**Enzymatic preparation of 3j with** P411-SCA-5188**:** >99% e.e.
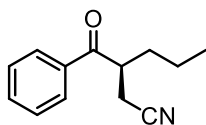


| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 11.766 | 869 | 59.3 | 0.2442 | 100.000 | 0.746 |

**(*S*)-3-benzoylhexanenitrile (3k)**



**Chiral-phase HPLC conditions:** Chiralpak IB, 15% *i*-PrOH in hexane, 1.0 mL/min, 25

ºC, 254 nm

**Racemic 3k:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 7.291 | 1779.9 | 210.8 | 0.1408 | 50.040 | 0.807 |
| 2 | 9.717 | 1777 | 156.1 | 0.1897 | 49.960 | 0.844 |
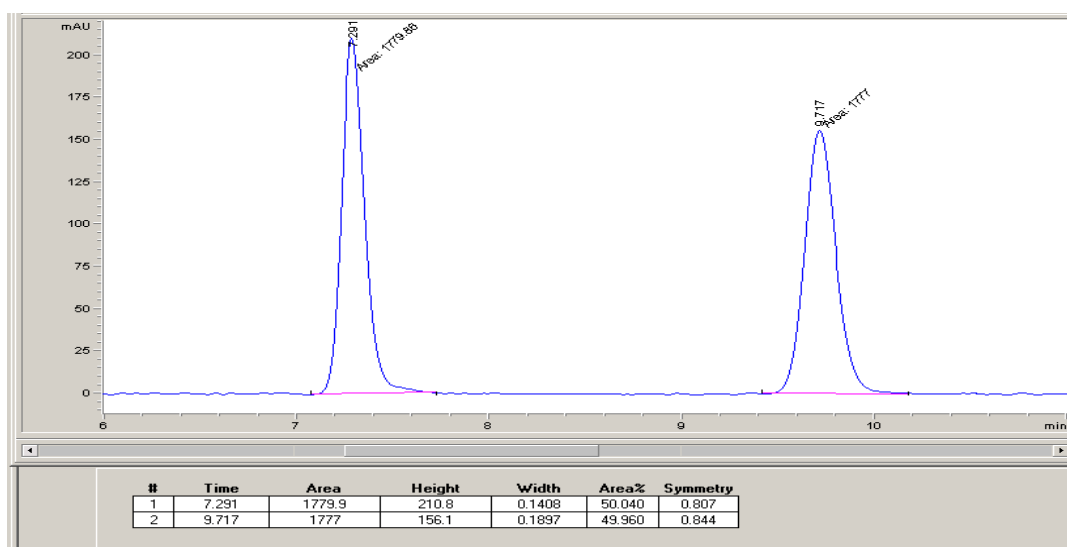
**Enzymatic preparation of 3k with** P411-SCA-5188**:** 91% e.e
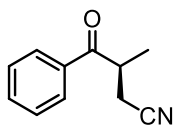


| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 7.314 | 24.5 | 3.2 | 0.126 | 4.338 | 1.142 |
| 2 | 9.74 | 539.9 | 49.6 | 0.1814 | 95.662 | 0.882 |

**(*S*)-3-methyl-4-oxo-4-phenylbutanenitrile (3l)**



**Chiral-phase HPLC conditions:** Chiralpak IB, 15% *i*-PrOH in hexane, 1.0 mL/min, 25 ºC, 254 nm

**Racemic 3l:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 8.6 | 4580.5 | 454.2 | 0.1557 | 49.200 | 0.721 |
| 2 | 9.001 | 4729.5 | 434.7 | 0.1649 | 50.800 | 0.776 |
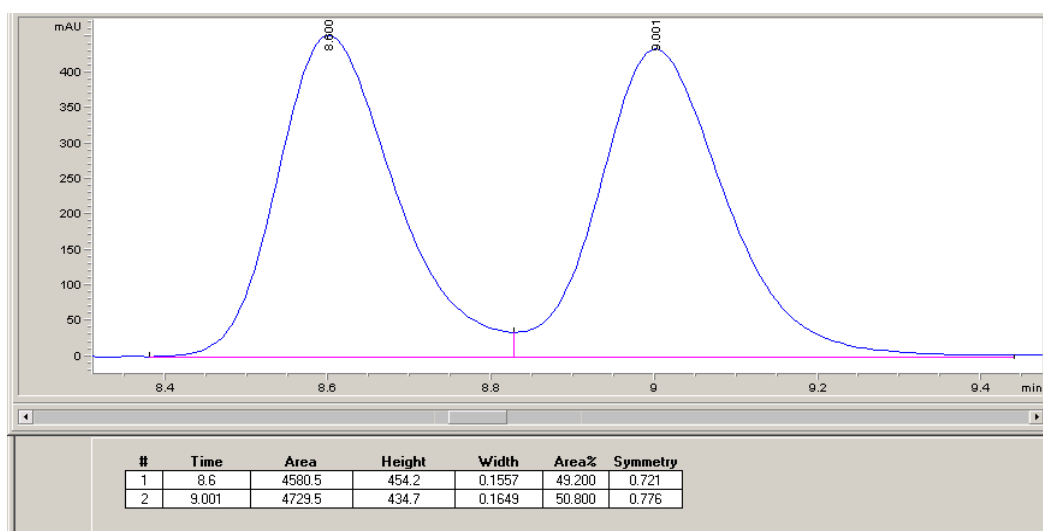
**Enzymatic preparation of 3l with** P411-SCA-5188**:** 99% e.e.
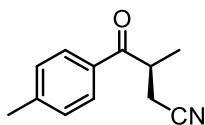


| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 8.97 | 51.5 | 5.5 | 0.1569 | 100.000 | 0.91 |

**(*S*)-3-methyl-4-oxo-4-(p-tolyl)butanenitrile (3m)**



**Chiral-phase HPLC conditions:** Chiralpak IA, 3% *i*-PrOH in hexane, 1.0 mL/min, 25 ℃,
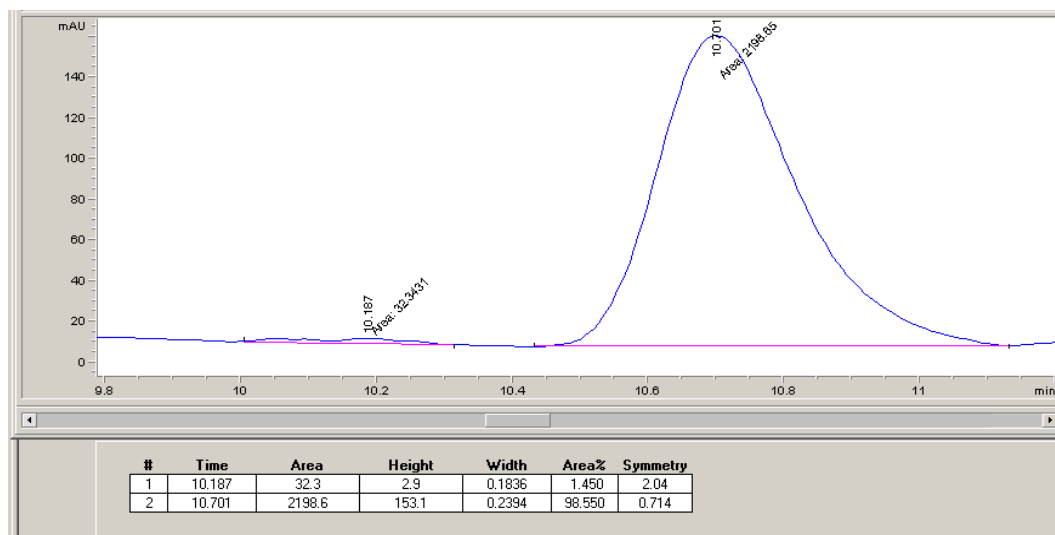
245 nm

**Racemic 3m:**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 15.501 | 3576.8 | 173 | 0.3133 | 46.134 | 0.662 |
| 2 | 16.127 | 4176.3 | 172.3 | 0.3589 | 53.866 | 0.611 |
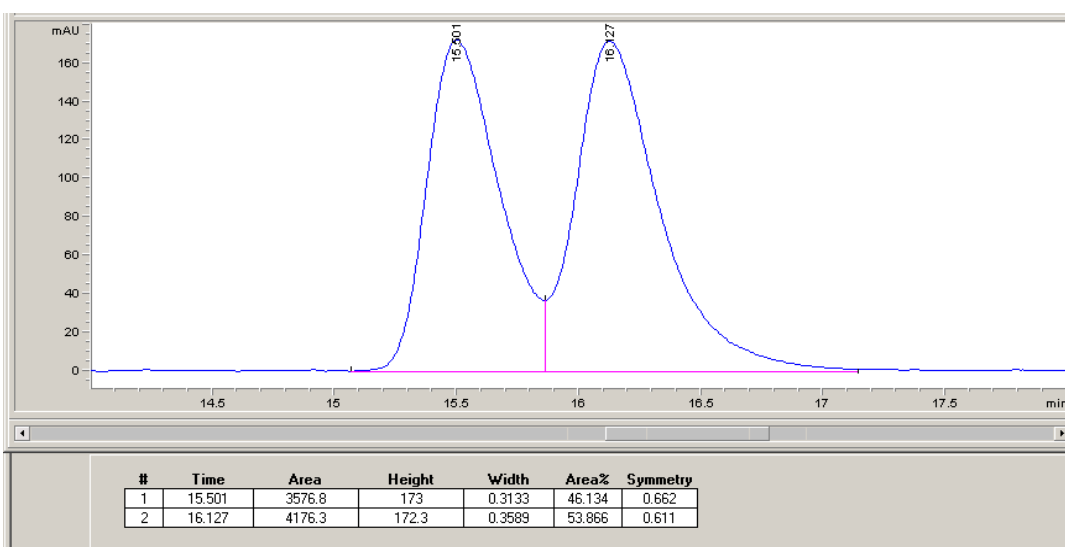
**Enzymatic preparation of 3m with** P411-SCA-5188**:** 90% e.e.



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 15.624 | 108.7 | 5.6 | 0.2311 | 4.896 | 1.238 |
| 2 | 16.17 | 2111.6 | 92.1 | 0.3276 | 95.104 | 0.619 |
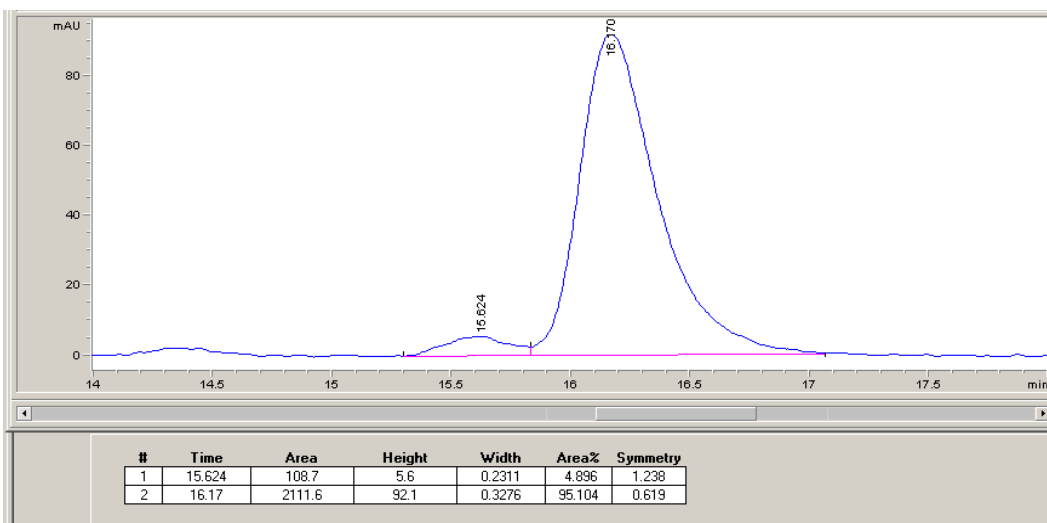
**Molecular Dynamics Simulations**

Molecular dynamics (MD) simulations were prepared with the Amber 16[1] package and the AmberTools 21 package.[2] The structure of the B-chain of the structurally strongly related P411 **E10** variant (PDB ID: 5UCW) was used as the starting point for the preparation of the MD simulations.[3] The beneficial N70S, L177M, W263M and Q437A mutations were manually introduced using Pymol v.1.8.2.[4] Hydrogen atoms were added using the reduce tool from the AmberTools 21 package. For the deprotonated serine residue coordinated to iron (named SEO), the residue was capped with an acyl group (N-terminus) and a methylamidate group (C-terminus) and restrained electrostatic potential (RESP) atomic charges[5] were calculated at the B3LYP/6-31G(d)[6,7] level of theory in Gaussian 09, Rev. D.01,[8] according to the Merz–Singh–Kollman scheme.[9,10] A combination of the ff14SB force field for proteins, the General Amber Force Field (gaff)[11] for the ligand and TIP3P parameters[12] for water molecules was employed.[13] For the ligand, RESP atomic charges were calculated at the B3LYP/6-31G(d) level of theory according to the Merz–Singh–Kollman scheme. Force field parameters and charges for the iron heme-ligand complex were generating using the MCPB.py program[14] with geometry optimizations, force constant and Merz–Singh–Kollman RESP charge calculations at the B3LYP/6-31G(d) level of theory. For the bond between Fe and the side-chain oxygen from SEO, $r_{eq}$ was set to 1.8655 Å and k to 100.0 kcal mol$^{-1}$ Å$^{-2}$ using parmed. For Si–C and Si–O bonds, the $r_{eq}$ and k values derived by Wang et al. were employed. X. Dong, X. Yuan, Z. Song, Q. Wang, *Phys. Chem. Chem. Phys.* **2021**, 23, 12582-12591. The system was solvated in a 10 Å buffer water box and neutralized by addition of sodium ions.

Employing the GPU-accelerated pmemd[15] code implemented in Amber 16, the energy was minimized over 5,000 steps with positional restraints (500 kcal mol$^{-1}$ Å$^{-2}$) applied to all atoms except water molecules and sodium ions, followed by an energy minimization over 5,000 steps with positional restraints (2.0 kcal mol$^{-1}$ Å$^{-2}$) on all peptide backbone atoms (C, Cα, N) as well as all non-hydrogen atoms of the iron heme-ligand complex. Subsequently, the temperature was raised from 0 K to 300 K within 300 ps in an NVT ensemble with positional restraints (30 kcal mol$^{-1}$ Å$^{-2}$) on all peptide backbone atoms (C, Cα, N) as well as all non-hydrogen atoms of the iron heme-ligand complex. Afterwards, the system was equilibrated at 300 K and 1 atm for 2 ns in an NPT ensemble with high positional restraints (30 kcal mol$^{-1}$ Å$^{-2}$) on all peptide backbone atoms (C, Cα, N) as well as all non-hydrogen atoms of the iron heme-ligand complex, followed by an equilibration at 300 K and 1 atm for 2 ns in an NPT ensemble with low positional restraints (0.5 kcal mol$^{-1}$ Å$^{-2}$) on all peptide backbone atoms (C, Cα, N) as well as all non-hydrogen atoms of the iron heme-ligand complex. Production runs were performed in an NPT ensemble for 1,000 ns. For each system, 3 independent replicas were simulated. k-means clustering was performed using cpptraj to analyze each MD trajectory and obtain representative structures for the most populated clusters. Distances were measured between the closest carbon or oxygen atoms of the ligand and carbon or oxygen atoms of the respective amino acid residues. Probability density plots were prepared with Excel 2016 with bin sizes of 0.2 Å for distances. Structures were visualized using Pymol v.1.8.2.[4]

$^1$H, $^{13}$C, and $^{19}$F NMR spectra of the compounds can be found in the Supporting Information of the published paper.

# 3.8 References for Chapter 3

References for Sections 3.1-3.5:

1. Straathof, A. J., Panke, S. & Schmid, A. The Production of Fine Chemicals by Biotransformations. *Curr. Opin. Biotechnol.* **13**, 548–556 (2002).

2. Schoemaker, H. E., Mink, D. & Wubbolts, M. G. Dispelling the Myths— Biocatalysis in Industrial Synthesis. *Science* **299**, 1694–1697 (2003).

3. Nestl, B. M., Nebel, B. A. & Hauer, B. Recent Progress in Industrial Biocatalysis. *Curr. Opin. Chem. Biol.* **15**, 187–193 (2011).

4. Wu, S., Snajdrova, R., Moore, J. C., Baldenius, K. & Bornscheuer, U. T. Biocatalysis: Enzymatic Synthesis for Industrial Applications. *Angew. Chem., Int. Ed.* **60**, 88–119 (2021).

5. Bell, E. L. *et al.* Biocatalysis. *Nat. Rev. Methods Primers* **1**, 46 (2021).

6. Patel, R. N. Biocatalysis for Synthesis of Pharmaceuticals. *Bioorg. Med. Chem.* **26**, 1252–1274 (2018).

7. Wu, S., Zhou, Y. & Li, Z. Biocatalytic Selective Functionalisation of Alkenes *via* Single-Step and One-Pot Multi-Step Reactions. *Chem. Commun.* **55**, 883–896 (2019).

8. McDonald, R. I., Liu, G. & Stahl, S. S. Palladium(II)-Catalyzed Alkene Functionalization *via* Nucleopalladation: Stereochemical Pathways and Enantioselective Catalytic Applications. *Chem. Rev.* **111**, 2981–3019 (2011).

9. Beller, M., Seayad, J., Tillack, A. & Jiao, H. Catalytic Markovnikov and *anti*-Markovnikov Functionalization of Alkenes and Alkynes: Recent Developments and Trends. *Angew. Chem., Int. Ed.* **43**, 3368–3398 (2004).

10. Rodriguez-Ruiz, V. *et al.* Recent Developments in Alkene Hydro-Functionalisation Promoted by Homogeneous Catalysts Based on Earth Abundant Elements: Formation of C–N, C–O and C–P Bond. *Dalton Trans.* **44**, 12029–12059 (2015).

11. Coombs, J. R. & Morken, J. P. Catalytic Enantioselective Functionalization of Unactivated Terminal Alkenes. *Angew. Chem., Int. Ed.* **55**, 2636–2649 (2016).

12. Yanto, Y. *et al.* Asymmetric Bioreduction of Alkenes Using Ene-Reductases YersER and KYE1 and Effects of Organic Solvents. *Org. Lett.* **13**, 2540–2543 (2011).

13. Yang, L.-C., Deng, H. & Renata, H. Recent Progress and Developments in Chemoenzymatic and Biocatalytic Dynamic Kinetic Resolution. *Org. Process Res. Dev.* **26**, 1925–1943 (2022).

14. Demming, R. M. *et al.* Asymmetric Enzymatic Hydration of Unactivated, Aliphatic Alkenes. *Angew. Chem., Int. Ed.* **58**, 173–177 (2019).

15. Stueckler, C. *et al.* Stereocomplementary Bioreduction of *α,β*-Unsaturated Dicarboxylic Acids and Dimethyl Esters using Enoate Reductases: Enzyme- and Substrate-Based Stereocontrol. *Org. Lett.* **9**, 5409–5411 (2007).

16. Wu, S. *et al.* Enantioselective *trans*-Dihydroxylation of Aryl Olefins by Cascade Biocatalysis with Recombinant *Escherichia coli* Coexpressing Monooxygenase and Epoxide Hydrolase. *ACS Catal.* **4**, 409–420 (2014).

17. Williams, J. M. J. *Preparation of Alkenes: A Practical Approach*. (Oxford University Press, Oxford, UK, 1996).

18. Nevesely, T., Wienhold, M., Molloy, J. J. & Gilmour, R. Advances in the $E - Z$ Isomerization of Alkenes Using Small Molecule Photocatalysts. *Chem. Rev.* **122**, 2650–2694 (2022).

19. Litman, Z. C., Wang, Y., Zhao, H. & Hartwig, J. F. Cooperative Asymmetric Reactions Combining Photocatalysis and Enzymatic Catalysis. *Nature* **560**, 355–359 (2018).

20. Mohr, J. T., Moore, J. T. & Stoltz, B. M. Enantioconvergent Catalysis. *Beilstein J. Org. Chem.* **12**, 2038–2045 (2016).

21. Bhat, V., Welin, E. R., Guo, X. & Stoltz, B. M. Advances in Stereoconvergent Catalysis from 2005 to 2015: Transition-Metal-Mediated Stereoablative Reactions, Dynamic Kinetic Resolutions, and Dynamic Kinetic Asymmetric Transformations. *Chem. Rev.* **117**, 4528–4561 (2017).

22. Kroutil, W., Mischitz, M. & Faber, K. Deracemization of (±)-2,3-Disubstituted Oxiranes *via* Biocatalytic Hydrolysis Using Bacterial Epoxide Hydrolases: Kinetics

of an Enantioconvergent Process. *J. Chem. Soc., Perkin Trans.* **1**, 3629–3636 (1997).

23. Chiappe, C. *et al.* Biocatalysis in Ionic Liquids: The Stereoconvergent Hydrolysis of *trans-β*-Methylstyrene Oxide Catalyzed by Soluble Epoxide Hydrolase. *J. Mol. Catal. B Enzym.* **27**, 243–248 (2004).

24. Yang, Y., Cho, I., Qi, X., Liu, P. & Arnold, F. H. An Enzymatic Platform for the Asymmetric Amination of Primary, Secondary and Tertiary C(*sp³*)–H Bonds. *Nat. Chem.* **11**, 987–993 (2019).

25. Mao, R. et al. Enantio- and Diastereoenriched Enzymatic Synthesis of 1,2,3-Polysubstituted Cyclopropanes from (Z/E)-Trisubstituted Enol Acetates. J. Am. Chem. Soc. 145, 16176–16185 (2023).

26. Carreira, E. M. & Kvaerno, L. *α-Functionalization of Enolates*. 3, 69–102 (WileyVCH, 2007).

27. Dugger, R. W., Ragan, J. A. & Ripin, D. H. B. Survey of GMP Bulk Reactions Run in a Research Facility between 1985 and 2002. *Org. Process Res. Dev.* **9**, 253–258 (2005).

28. Wright, T. B. & Evans, P. A. Catalytic Enantioselective Alkylation of Prochiral Enolates. *Chem. Rev.* **121**, 9196–9242 (2021).

29. Evans, D. A., Ennis, M. D. & Mathre, D. J. Asymmetric Alkylation Reactions of Chiral Imide Enolates. A Practical Approach to the Enantioselective Synthesis of *alpha*-Substituted Carboxylic Acid Derivatives. *J. Am. Chem. Soc.* **104**, 1737–1739 (1982).

30. Myers, A. G., Yang, B. H., Chen, H. & Gleason, J. L. Use of Pseudoephedrine as a Practical Chiral Auxiliary for Asymmetric Synthesis. *J. Am. Chem. Soc.* **116**, 9361–9362 (1994).

31. Prier, C. K., Zhang, R. K., Buller, A. R., Brinkmann-Chen, S. & Arnold, F. H. Enantioselective, Intermolecular Benzylic C–H Amination Catalysed by an Engineered Iron-Haem Enzyme. *Nat. Chem.* **9**, 629-634 (2017).

32. Sruthi, P. R. & Anas, S. An Overview of Synthetic Modification of Nitrile Group in Polymers and Applications. *J. Polym. Sci.* **58**, 1039–1061 (2020).

33. Wang, X. *et al.* Nitrile-Containing Pharmaceuticals: Target, Mechanism of Action, and Their SAR Studies. *RSC Med. Chem.* **12**, 1650–1671 (2021).

34. Yang, Y. & Arnold, F. H. Navigating the Unnatural Reaction Space: Directed Evolution of Heme Proteins for Selective Carbene and Nitrene Transfer. *Acc. Chem. Res.* **54**, 1209–1225 (2021).

References for Section 3.6:

1. Mao, R. et al. Enantio- and Diastereoenriched Enzymatic Synthesis of 1,2,3-Polysubstituted Cyclopropanes from (Z/E)-Trisubstituted Enol Acetates. *J. Am. Chem. Soc.* **145**, 16176–16185 (2023).

2. Gibson, D. G. et al. Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases. *Nat. Methods* **6**, 343-345 (2009).

3. Kille, S. et al. Reducing Codon Redundancy and Screening Effort of Combinatorial Protein Libraries Created by Saturation Mutagenesis. *ACS Synth. Biol.* **2**, 83–92 (2013).

4. Barr, I. & Guo, F. Pyridine Hemochromagen Assay for Determining the Concentration of Heme in Purified Protein Solutions. *Bio. Protoc*. **5**, e1594 (2015).

5. Spinnato, D., Schweitzer-Chaput, B., Goti, G., Oseka, M. & Melchiorre, P. A Photochemical Organocatalytic Strategy for the alpha-Alkylation of Ketones by using Radicals. Angew. *Chem., Int. Ed*. **59**, 9485-9490 (2020).

6. Singh, K., Staig, S. J. & Weaver, J. D. Facile Synthesis of Z-Alkenes via Uphill Catalysis. *J. Am. Chem. Soc*. **136**, 5275-5278 (2014).

7. Hock, K. J. et al. Tryptamine Synthesis by Iron Porphyrin Catalyzed C–H Functionalization of Indoles with Diazoacetonitrile. *Angew. Chem., Int. Ed*. **58**, 3630-3634 (2019).

8. Gomez, J. E., Guo, W., Gaspa, S. & Kleij, A. W. Copper-Catalyzed Synthesis of gamma-Amino Acids Featuring Quaternary Stereocenters. *Angew. Chem., Int. Ed*. **56**, 15035-15038 (2017).

*C h a p t e r   4*

# ENZYMATIC ASSEMBLY OF DIVERSE LACTONE STRUCTURES: AN

# INTRAMOLECULAR C–H FUNCTIONALIZATION STRATEGY

Content in this chapter is adapted from published work:

Wackelin, D. J.[†]; Mao, R.[†]; Sicinski, K. M.; Zhao, Y.; Das, A.; Chen, K.; Arnold, F. H. Enzymatic Assembly of Diverse Lactone Structures: An Intramolecular C–H Functionalization Strategy. *Manuscript in preparation*. **2023**.

D.J.W. and K.C. conceived and designed the overall project under F.H.A.'s guidance. D.J.W. and K.C. conducted the initial screening of heme proteins and executed the directed evolution experiments with R.M. and Y.Z.'s assistance. D.J.W., R.M., and K.C. explored and studied the substrate scope and applications, with assistance from K.M.S for product purification. A.D. and D.J.W. performed the crystallization of products. D.J.W., R.M., K.C., and F.H.A. wrote the manuscript, incorporating input from all contributors. D.J.W. and R.M. contributed equally to this work.

**Abstract**

Lactones are cyclic esters with extensive applications in materials science, medicinal chemistry, and the food and perfume industries. Nature's strategy for synthesis of the many lactones found in natural products always relies on a single type of retrosynthetic strategy, a C–O bond disconnection. Here we describe a set of laboratory-engineered enzymes that use a new-to-nature C–C bond-forming strategy to assemble diverse lactone structures. These engineered 'carbene transferases' catalyze intramolecular carbene insertions into benzylic or allylic C–H bonds, which allows for the synthesis of lactones with different ring sizes and ring scaffolds from simple starting materials. Starting a from serine-ligated cytochrome P450 variant previously engineered for other carbene transfer activities, directed evolution generated variant P411-LAS-5247, which exhibits high activity for constructing 5-membered $\varepsilon$-lactone, lactam, and cyclic ketone products (up to 5600 total turnovers (TTN) and >99% enantiomeric excess (e.e.)). Further engineering led to variants P411-LAS-5249 and P411-LAS-5264, which deliver 6-membered $\delta$-lactones and 7-membered $\varepsilon$-lactones, respectively, overcoming the thermodynamically unfavorable ring strain associated with these products compared to the $\gamma$-lactones. This new carbene-transfer activity was further extended to the synthesis of complex lactone scaffolds based on fused, bridged, and spiro rings. The enzymatic platform developed here complements natural biosynthetic strategies for lactone assembly and expands the structural diversity of lactones accessible through C–H functionalization.

## 4.1 Introduction

Lactones are ubiquitous in natural products[1] and active pharmaceutical ingredients[2] and serve as essential building blocks for fine chemicals and polyesters.[1,3,4] Intramolecular carbene C–H insertion, catalyzed by transition metals such as rhodium, iridium, and copper,[5-9] has received significant attention for the synthesis of lactones. Based on a C–C bond disconnection strategy, this approach is simple, efficient, and exhibits high atom economy.[10] Due to the thermodynamic stability of the intermediates and products, however, intramolecular C–H insertion has been mainly limited to the synthesis of five-membered $\gamma$-lactones (Figure 1A).[5-9,11,12] While a few catalytic systems have demonstrated the synthesis of four-membered $\beta$-lactones through favorable five-membered transition states, examples of lactones with different ring sizes are limited, and reactions with these molecules often proceed with poor regio- or enantio-selectivity and thus result in mixtures of lactone products.[12,13] A general method for synthesis of chiral lactones with diverse ring sizes and structures using an intramolecular carbene C–H insertion strategy has yet to be demonstrated.

Nature does not employ C–C bond disconnection strategies to biosynthesize lactones; instead, C–O bond disconnection strategies are universally used, including Baeyer–Villiger oxidation,[14] oxidative lactonization of diols,[15] C–H hydroxylation of fatty acids followed by lactonization,[16] and reductive cyclization of keto-esters (Figure 1B).[17] Though these routes produce lactones with diverse structures and ring sizes, the C–O bond disconnection is a limitation which restricts retrosynthetic versatility and broader synthetic applications. This limitation inspired us to add the chemists' C–C bond disconnection approach to nature's biocatalytic repertoire by engineering enzymes to perform intramolecular lactone cyclization

via C–H insertion and using the macromolecular enzyme scaffold to impart desired regio- and enantioselectivities unavailable to small-molecule catalysts.

Over the past decade, various hemeproteins have been repurposed to perform abiological carbene- and nitrene-transfer reactions.[18-22] Enzymes containing a heme prosthetic group have been engineered to activate a carbene or nitrene precursor and subsequently control the highly reactive carbene/nitrene intermediate to target specific bonds, including C–H, heteroatom–H, and unsaturated C–C bonds. Motivated by these precedents, we aimed to develop a biocatalytic intramolecular carbene C–H insertion strategy for lactone synthesis (Figure 1C). Several challenges must be overcome to realize this goal. First, the enzymes must bind the carbene intermediates and orient them for carbene insertion into desired C–H bonds. Second, the enzymes need to stabilize intermediates with varied ring strains to give lactone products of different sizes. Third, because the stereoconfiguration of the reaction intermediates is influenced by the size of the ring, the enzymes must precisely control the stereoconfiguration of different ring intermediates to achieve stereoselectivity. We proposed that such an enzymatic platform could be generated by directed evolution, and with further evolution, the enzymes could accept a greater diversity of substrates to synthesize other cyclic structures, such as lactams and cyclic ketones important in a broad range of pharmaceutical intermediates.
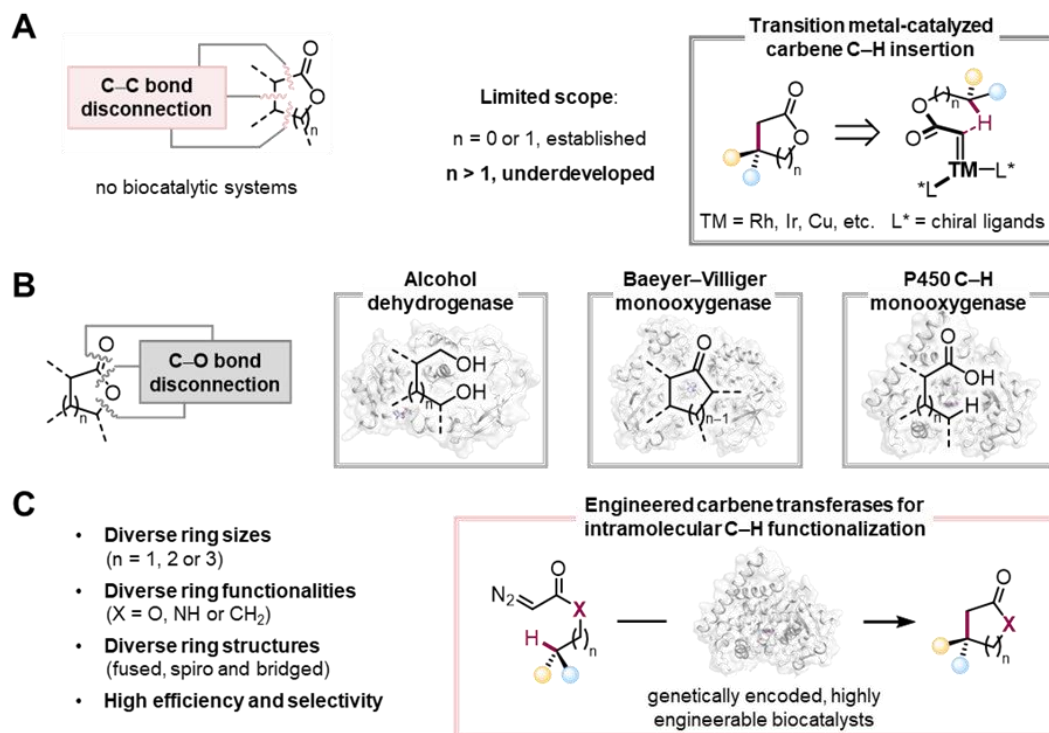
**Figure 4-1. Chemocatalytic and biocatalytic lactone synthesis.** (A) Small-molecule transition-metal-catalyzed intramolecular C–H insertion, which typically produces five-membered lactones. (B) Nature employs C–O disconnections for the construction of lactones. (C) This work: a biocatalytic platform enables access to diverse lactones via intramolecular carbene transfer reactions. R, organic groups; TM, transition metal.

## 4.2 Enzymatic γ-lactone synthesis: Initial screening, directed evolution, and substrate scope

We initiated this investigation by focusing on the intramolecular carbene C–H insertion reaction of diazo compound **1a** to form γ-lactone **2a**, shown in Figure 2. γ-Lactones are the preferred products of intramolecular carbene C–H insertion with small-molecule transition metal catalysts.[23,24] We proposed to start with a biocatalytic system for γ-lactones and then move on to more challenging targets. To this end, we surveyed an in-house collection of over

one hundred cytochrome P450 and P411 (P450 with an axial serine ligand) variants, previously engineered for carbene and nitrene transfer, in the form of whole *Escherichia coli* cells. Most of the hemeproteins exhibited no measurable activity for the desired carbene C–H insertion transformation; the diazo substrate **1a** was nonetheless fully consumed to give a mixture of unwanted products from carbene O–H insertion with water, carbene dimerization, and cycloaddition between carbene dimers and the diazo substrate. However, P411 variant **C10**, known for its promiscuous activities in other carbene-transfer reactions such as internal alkyne cyclopropenation[25] and lactone-carbene C–H insertion,[26] was effective in producing the γ-lactone product **2a** (7% yield, Figure 2A). We thus used **C10** as the starting template for engineering a set of lactone synthase enzymes (LAS).

To enhance catalytic efficiency, we performed directed evolution by targeting active-site residues for site-saturation mutagenesis (SSM). Sequential rounds of mutagenesis and screening identified variant P411-LAS-5243 with four beneficial mutations (Q437I, V328I, L78M, and L436R) that collectively increase TTN fivefold, providing **2a** in 41% yield (Figures 2A and 2B). Site-saturation mutagenesis of the cysteine loop uncovered mutation L401V in variant P411-LAS-5244, which improved the yield to 52% (Figures 2A and 2B). When exhaustive examination of other active-site residues did not result in further activity enhancement, we selected several non-active-site residues, known to contribute to P450's dynamics with native substrates bound, for saturation mutagenesis.[27,28] Library screening led to three additional beneficial mutations (L162I, R190L, and E70S) in final variant P411-LAS-5247, with 3850 total turnovers and 66% yield for product **2a** (Figures 2A and 2B). The evolved enzyme displayed near-perfect stereo-control with >99% enantiopurity for product **2a**.

We investigated the activity of P411-LAS-5247 against an array of diazo compounds for the synthesis of $\gamma$-lactones via intramolecular carbene insertion into benzylic C–H bonds (Figure 1C). A variety of $\gamma$-lactones were synthesized with excellent TTNs and enantioselectivities (up to 3800 TTN and >99% e.e., Figure 2C). Substrates bearing diverse substituents on the phenyl ring, electron-neutral (**1a** and **1b**, Figure 2), -rich (**1c**, Figure 2), or -deficient (**1d** and **1e**, Figure 2), are all compatible with this biotransformation. Notably, **2e** can serve as a synthetic intermediate for the antispasmodic drug baclofen.[10] In addition to diazoacetate substrates, we also tested diazoketone **1f** and diazoamide **1g**, which upon cyclization form non-lactone products. These substrates proceeded smoothly with P411-LAS-5247 to furnish cyclic ketone **2f** and $\gamma$-lactam **2g** with excellent TTNs (up to 5500 TTN, Figure 2D). It is worth noting that the lactam product was obtained without any additional protection-deprotection steps on the nitrogen, underscoring the ability of the enzyme to work with substrates having active functional groups. The absolute stereochemistry for enzymatic product **2c** was assigned as *S* by comparing the elution order of the two enantiomers with a literature report[29] (see Supplementary Information for details). The other $\gamma$-lactones **2** were assigned by analogy.

**Figure 4-2. Directed evolution and substrate-scope study for γ-lactone synthesis.** Reaction conditions: 2.5 mM **1**, *E. coli* whole cells harboring P411 variants ($OD_{600}$ = 2.5) in M9-N aqueous buffer (pH 7.4), 1.25% v/v acetonitrile (co-solvent), room temperature, anaerobic conditions, 24 h. (A) Directed evolution culminating in P411-LAS-5247. Reactions were performed in triplicate (n = 3). Yields reported are based on calibrated HPLC traces and are the means of three independent experiments. (B) Evolutionary trajectory of P411-LAS-5247 from P411-**C10** and locations of beneficial mutations in protein structure based on a close variant **E10** (PDB ID: 5UCW). (C) Scope of γ-lactones using P411-LAS-5247. (D) Synthesis of a 5-membered cyclic ketone and a γ-lactam using P411-LAS-5247.

## 4.3 Expansion of the enzymatic platform for δ- and ε-lactone synthesis

With the biocatalytic C–H insertion strategy established for $\gamma$-lactone synthesis, we sought to extend it to $\delta$-lactones. We chose to test diazo **3b**, which has one additional carbon compared to substrate **1a**, toward the synthesis of **4b** (Figure 3C). Conventional dirhodium catalysts catalyze the conversion of **3** into $\gamma$-lactones as the sole products through carbene insertion into homobenzylic C–H bonds.[30,31] We were thus curious whether the enzymes would exhibit different regioselectivity for C–H insertion. Screening the P411-LAS-5247 lineage against substrate **3b**, we found that P411-LAS-5244 exhibited high efficiency for the formation of $\delta$-lactone **4b** as the exclusive C–H insertion product (48% yield, 3600 TTN, and 90% e.e., Figure 3C). These results show that this biocatalytic platform features a different site preference from the rhodium system - the hemeprotein carbene transferases discriminate C–H bonds based on bond strength, while the geometric distance of the C–H bonds to the carbene center plays a more important role in the regio-selectivity of rhodium catalysts.

Having identified P411-LAS-5244 as a potent enzyme for $\delta$-lactone formation, we further investigated the substrate scope. When challenged with other substrates bearing substituents on the phenyl ring, P411-LAS-5244 demonstrated high activity, but its enantioselectivity was greatly diminished (e.g., 46% e.e. for product **4a**, Figure 3C). We elected to continue directed evolution of this highly active variant, focusing on improving enantioselectivity. Two active-site mutations, I328V and L82K, were identified that boosted enantioselectivity to 83% e.e. for **4a** while maintaining activity. The resulting variant, P411-LAS-5249, was revealed to deliver $\delta$-lactone products bearing diverse functional groups with high efficiency

and enantioselectivities (up to 2800 TTN, and 86% e.e., Figure 3C). The absolute

stereochemistry for enzymatic product **4b** was assigned as *R* by comparing the elution order

of the two enantiomers with a literature report[32] (see Supplementary Information for details).

The other $\delta$-lactones **4** were assigned by analogy.

By recognizing the site preference in this enzymatic C–H insertion system, we then asked

whether 7-membered $\varepsilon$-lactones could be accessed with an appropriate diazo substrate (e.g.,

the transformation of **5a** into **6a**, Figure 3D). 7-Membered lactones are generally much more

difficult to assemble compared to 5- or 6-membered lactones due to the higher enthalpy and

entropy costs of cyclization.[23,24] The P411-LAS-5247 and P411-LAS-5249 lineages,

unfortunately, only displayed minimal or no activity toward the formation of $\varepsilon$-lactone **6a**.

We therefore screened a large collection of enzyme variants, all engineered from the initial

**C10** parent, for $\varepsilon$-lactone formation from **5a**. A variant previously evolved for internal alkyne

cyclopropenation, **C10**-WIRF_GA (Figure 3A, Supplementary Table 5),[25] exhibited

promising activity (16% yield, 940 TTN, Figure 3D).

Directed evolution of **C10**-WIRF_GA introduced nine new mutations and yielded an

efficient variant, P411-LAS-5259, for the synthesis of $\varepsilon$-lactone **6a** (58% yield and 2400

TTN, Figures 3A and 3D), but with unsatisfactory enantioselectivity (51% e.e.). A single

amino acid mutation at residue 87 dramatically boosted the selectivity to 96% e.e. (Figures

3A and 3D), which highlights the crucial role of this residue in controlling the orientation of

carbene intermediates in the protein active site.[22,33,34] This mutation caused the activity to

drop by nearly fivefold, however (Figure 3D). Further evolution was performed to recover

the activity (Figures 3A and 3D), leading to a new variant P411-LAS-5264 with high activity

and enantioselectivity toward the synthesis of **6a** (68% yield, 3900 TTN, and 95% e.e.,

Figures 3A and 3D). Beyond model substrate **5a**, P411-LAS-5264 also accepts

structurally diverse and challenging substrates, such as **5b** and **5c**, for carbene insertion at an

allylic site or a benzylic site of an unprotected indole, affording the corresponding $\varepsilon$-lactone

products **6b** and **6c** with up to 3300 TTN and 90% e.e. (Figure 3D). The absolute

stereochemistry for enzymatic product **6a** was assigned as *S* through X-ray crystallography

(see Supplementary Figure 3). The other $\varepsilon$-lactone **6** were assigned by analogy.
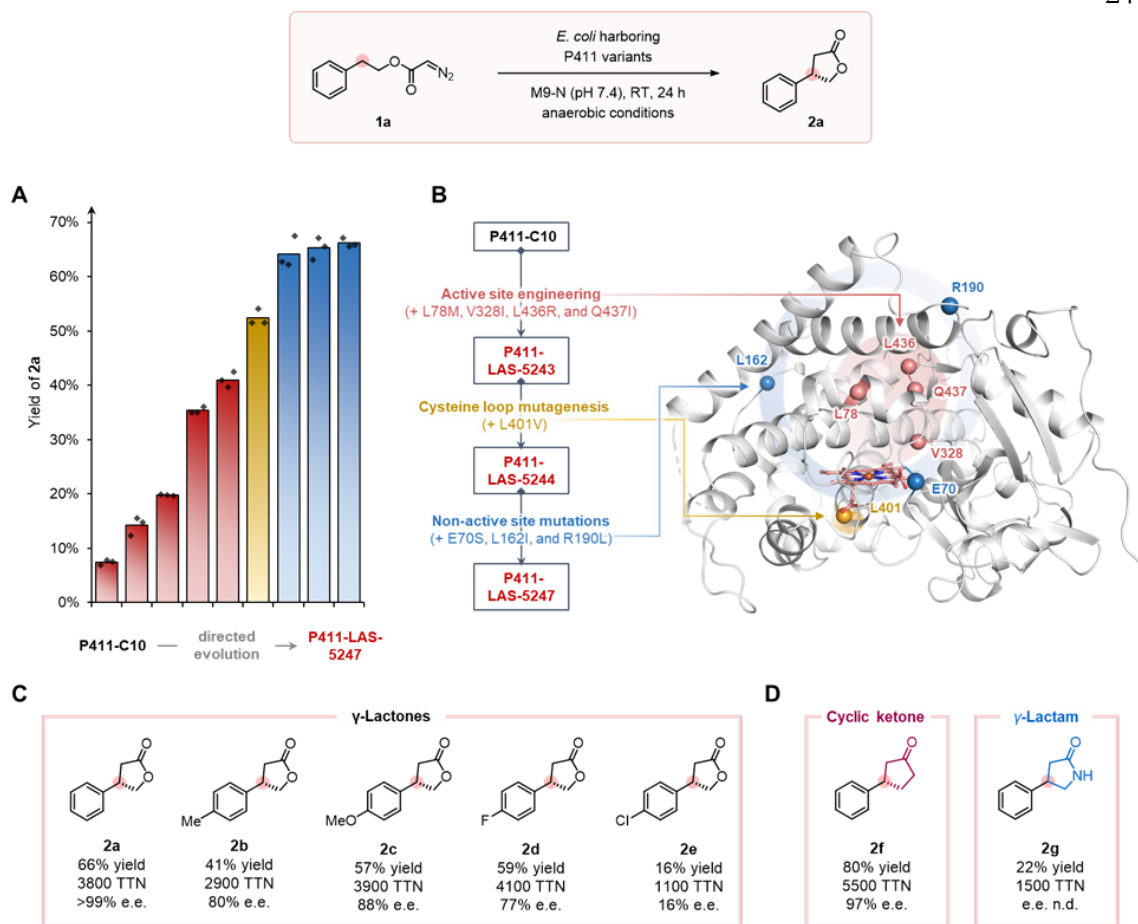
**Figure 4-3. Directed evolution and substrate-scope study for δ- and ε-lactone synthesis.** Reaction conditions: 2.5 mM **3** or **5**, *E. coli* whole cells harboring P411 variants (OD$_{600}$ = 2.5) in M9-N aqueous buffer (pH 7.4), 1.25% v/v acetonitrile (co-solvent), room temperature, anaerobic conditions, 24 h. Reactions were performed in triplicate (n = 3). Yields reported are based on calibrated HPLC traces and are the means of three independent experiments. (A) Evolutionary trajectories of lactone synthases P411-LAS-5247, P411-LAS-5249, and P411-LAS-5264 from P411-**C10**. (B) Locations of beneficial mutations in the protein structure based on a close variant **E10** (PDB ID: 5UCW); (C)

Scope of $\delta$-lactones using P411-LAS-5249. (D) Scope of $\varepsilon$-lactones using P411-LAS-5264.

[a]P411-LAS-5244 was used. [b]P411-LAS-5249 was used.

## 4.4 Complex lactone synthesis through enzymatic carbene C–H insertion

We were also interested in the synthesis of more complex ring scaffolds. The large collection of P411 enzyme variants derived from **C10** provides a biocatalyst pool with which to quickly search for variants capable of catalyzing desired C–H insertions. Interestingly, when paired with different substrates, different enzymes displayed unique stereo- and regio-selectivities to afford lactone products with high structural complexity. For instance, variant P411-LAS-5265 (see Supplementary Table 5) efficiently converted substrates **7a** and **7b**, derived from indane, into $\gamma$- and $\delta$-lactones **8a** and **8b** in a fused ring scaffold (Figures 4A and 4B), which represents a C–H functionalization onto an existing ring with high activity and stereoselectivity. The reaction of **7a** was scaled up and the product isolated to afford 69 mg of **7b** with 76% yield, demonstrating that the synthesis of these complex rings is feasible at small scale. Racemic substrates **7c** and **7d** could be utilized by variants P411-LAS-5259 and P411-LAS-5257 (see Supplementary Table 5), respectively, to target tertiary benzylic C–H bonds for carbene insertion and give the 6-membered spiro-lactone products, **8c** and **8d**, with up to 2500 TTN (Figures 4C and 4D).

The biocatalysts could also discriminate the enantiomers of racemic substrates to give different product outcomes. Racemic substrate **7e**, bearing two distinct sets of benzylic C–H bonds, was transformed by enzyme P411-LAS-5259 into two lactone products, a fused $\gamma$-lactone (**8e**) and a bridged $\delta$-lactone (**8e'**) with high efficiency and selectivity (Figure 4E). Similarly, spiro $\delta$-lactone (**8f**) and bridged 8-membered $\zeta$-lactones (**8f'**) were observed as the

products of racemic substrate **7f** using enzyme P411-LAS-5247 (Figure 4F). Given that

the enzyme's active site is a chiral environment, we rationalized that the two enantiomers of

the racemic substrates could bind in the active site with different orientations, which enables

carbene insertion into geometrically distinct C–H bonds. This process represents a parallel

kinetic resolution (PKR),[35,36] which requires the catalyst to distinguish the substrate

enantiomers and proceed with each enantiomer through different reaction pathways.[37] To

further verify the PKR process, we used the enantiopure starting materials of **7f** for the

biotransformation: P411-LAS-5247 specifically converted (*R*)-**7f** to spiro-lactone **8f** and (*S*)-

**7f** to bridged-lactone **8f'** with high TTNs and selectivities (see Supplementary Figure 2),

suggesting that P411-LAS-5247 can differentiate between the two enantiomers and perform

a parallel kinetic resolution. We also identified another two variants, P411-LAS-5256 and

P411-LAS-5266 (see Supplementary Table 5), which converted racemic **7f** to spiro-lactone

**8f** and bridged-lactone **8f'**, respectively, exhibiting simple product specificity (Figures 4G

and 4H).
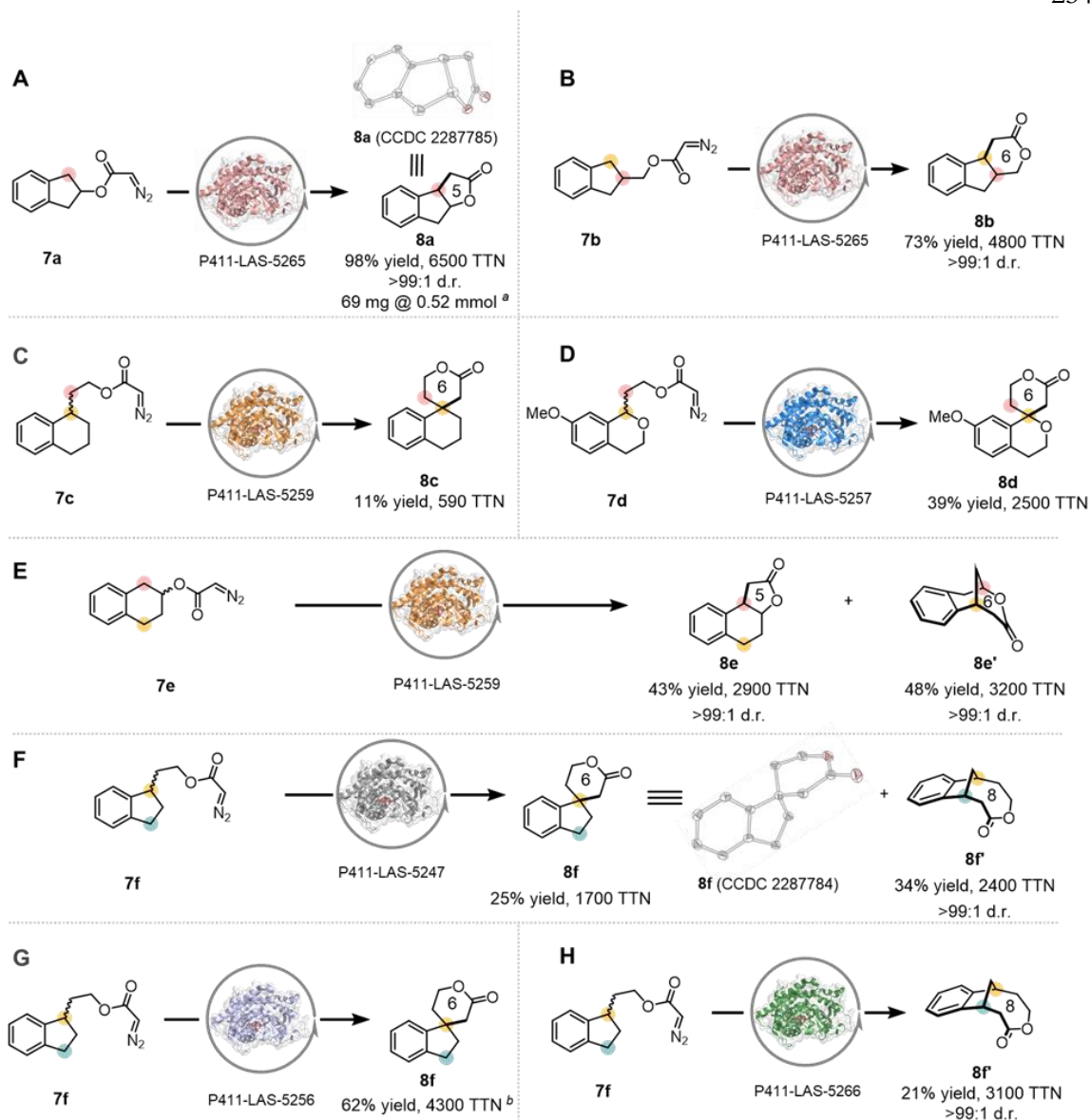
**Figure 4-4. Enzymatic synthesis of complex lactones.** Reaction condition: 2.5 mM **3**, *E. coli* whole cells harboring P411 variants ($OD_{600} = 2.5$) in M9-N aqueous buffer (pH 7.4), 1.25% v/v acetonitrile (co-solvent), room temperature, anaerobic conditions, 48 h. Yields reported are based on calibrated HPLC traces and are the means of three independent experiments. Enzymatic synthesis of (A and B) fused lactones, (C and D) spiro lactones, (E) fused- and bridged-lactones. (F) Parallel kinetic resolution toward **8f** and **8f'** catalyzed by P411-LAS-0003 from racemic diazo **7f.** (G) Specific synthesis of spiro lactone **8f** catalyzed by P411-LAS-0011 from racemic diazo **7f**. (H) Specific

synthesis of bridged-lactone **8f** catalyzed by P411-LAS-0012 from racemic diazo **7f**. [a]76% isolated yield. [b]8% yield, 530 TTN for product **8f'**. [c]**8f'** was not detected.

## 4.5 Summary and conclusion

We have developed an enzymatic platform for asymmetric intramolecular carbene C–H transfer reactions, which enables straightforward construction of diverse lactone products with exceptional efficiency and selectivity. P450-derived carbene transferases were engineered for the selective functionalization of benzylic or allylic C–H bonds, allowing for the efficient assembly of an extensive array of benzylic and allylic lactones, including $\gamma$-, $\delta$-, $\varepsilon$-lactones, as well as fused-, spiro-, and bridged-lactones. The new-to-nature C–C disconnection approach used by these carbene transferases serves as a valuable complement to the existing biotransformations based on C–O disconnections used by natural enzymes for lactone biosynthesis. In addition, many products demonstrated here, such as $\delta$-, $\varepsilon$-, and more complex lactones, as well as unprotected lactams and indoles, have proven to be challenging targets for synthetic transition metal catalysts, but are now readily accessible using this biocatalytic strategy. This enzymatic platform enriches the disconnection strategies for biocatalytic lactone assembly while expanding the structural diversity of lactones accessible to synthetic chemistry. We anticipate the potential of hemeprotein carbene transferases will be further developed to access even broader classes of cyclic compounds, including lactones, lactams, and others, with applications in synthetic chemistry and drug discovery.

## 4.6 Supplementary information for Chapter 4

**General Procedures**

Unless otherwise noted, all chemicals and reagents were obtained from commercial suppliers (Sigma-Aldrich, VWR, TCI America, Fischer Scientific, Alfa Aesar, Acros, and Combi Blocks) and used without further purification. Silica gel chromatography was carried out using AMD Silica Gel 60, 230–400 mesh. $^1$H and $^{13}$C NMR spectra were recorded on a Bruker Prodigy 400 MHz instrument (400 MHz for $^1$H and 101 MHz for $^{13}$C NMR) or a Varian 300 MHz Spectrometer (300 MHz for $^1$H NMR). Chemical shifts ($\delta$) are reported in ppm downfield from tetramethylsilane, using the solvent resonance as the internal standard ($^1$H NMR: $\delta$ = 7.26, $^{13}$C NMR: $\delta$ = 77.16 for CDCl$_3$). Data for $^1$H NMR are reported as follows: chemical shift ($\delta$ ppm), multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, p = pentet, sext = sextet, m = multiplet, dd = doublet of doublets, dt = doublet of triplets, ddd = doublet of doublet of doublets), coupling constant (Hz), integration. Sonication was performed using a Qsonica Q500 sonicator. High-resolution mass spectra were obtained at the California Institute of Technology Mass Spectral Facility. Samples were analyzed by field ionization (FI) using a JEOL AccuTOF GC-Alpha (JMS-T2000GC) mass spectrometer interfaced with an Agilent 8890 GC system. Ions detected by FI are radical cations.

*Escherichia coli* cells were grown using Luria-Bertani medium or HyperBroth (AthenaES) with 100 μg/mL ampicillin (LB$_{amp}$ or HB$_{amp}$). Primer sequences are available upon request. T5 exonuclease, Phusion polymerase, and *Taq* ligase were purchased from New England Biolabs (NEB, Ipswich, MA). M9-N minimal media (abbreviated as M9-N buffer, pH = 7.4)

were used as buffering systems for whole cells, unless otherwise specified. M9-N buffer

was used without a carbon source, it contains 47.7 mM $Na_2HPO_4$, 22.0 mM $KH_2PO_4$, 8.6

mM NaCl, 2.0 mM $MgSO_4$, and 0.1 mM $CaCl_2$., and 0.1 mM $CaCl_2$.

**Chromatography**

Chemical reactions were monitored using thin layer chromatography (Merck 60 silica gel

plates) and a UV lamp for visualization. Analytical reverse-phase high-performance liquid

chromatography (HPLC) was carried out using an Agilent 1200 series instrument and a

Kromasil 100 C18 column (4.6 mm $\times$ 50 mm, 3 µm) or an Agilent C18 column (InfinityLab

Poroshell 120 EC-C18, 4.6 x 50 mm, 2.7 µm; Part Number: 699975-902T) with water and

acetonitrile, both containing 0.1% acetic acid, as the mobile phase. Analytical chiral normal-

phase HPLC was conducted using either an Agilent 1200 series instrument with *n*-hexane

and isopropanol as the mobile phase or JASCO SF-2000 integrated analytical supercritical

fluid chromatography (SFC) system with supercritical $CO_2$ and isopropanol as the mobile

phase. Enantiomers were separated using one of the following chiral columns: Chiralpak IA,

Chiralpak IB, Chiralpak IC, Chiralpak AD-H, and Daicel CHIRALCEL OJ-H (all 4.6 mm $\times$

250 mm, 5 µm).

**Cloning and Site-Saturation Mutagenesis**

The genes encoding all enzymes described in this study were cloned using Gibson assembly[1]

into vector pET22b(+) (Novagen) between restriction sites *Nde*I and *Xho*I in frame with a *C*-

terminal 6×His-tag. Site-saturation mutagenesis was performed using the "22c-trick"[2] or

"NNK" as degenerate codons. The PCR products were digested with *Dpn*I, gel purified,

and ligated using Gibson Mix™.[1] Without further purification after the Gibson step, 1 µL of

the Gibson product was used to transform 50 μL of electrocompetent *E. coli* BL21 E.

cloni® (Lucigen) cells.

**Expression of P411 Variants in 96-Well Plates**

Single colonies from LB$_{amp}$ agar plates were picked using sterile toothpicks and cultured in

deep-well 96-well plates containing LB$_{amp}$ (300 μL/well) at 37 °C, 80% humidity, and 220

rpm shaking overnight. Subsequently, HB$_{amp}$ (1000 μL/well) in a deep-well plate was

inoculated with an aliquot (50 μL/well) of these overnight cultures and allowed to shake for

2.5 hours at 37 °C, 80% humidity, and 220 rpm. The plates were then cooled on ice for 30

minutes, and the cultures were induced with 0.5 mM isopropyl *β*-d-1-thiogalactopyranoside

(IPTG) (final concentration). The expression was supplemented with 1.0 mM 5-

aminolevulinic acid (ALA) (final concentration). Expression was then conducted at 20 °C,

220 rpm for 20–22 hours.

**Plate Reaction Screening in Whole-Cell Format**

*E. coli* cells harboring P411 variants in deep-well 96-well plates were pelleted (3,500 g, 5

min, 4 °C) and resuspended in M9-N buffer (360 μL, pH 7.4) by gentle vortexing. The 96-

well plates were then transferred to an anaerobic chamber. To deep-well plates of cell

suspensions were added the substrate (20 μL/well, 200 mM in acetonitrile). The plates were

sealed with aluminum sealing foil immediately after the addition and shaken in the anaerobic

chamber at room temperature and 600 rpm. After 24 hours, the seals were removed and

acetonitrile (600 μL/well) was added. The plates were tightly sealed with silicone mats,

vigorously vortexed, and centrifuged (4,500 g, 5 min) to precipitate proteins and cell debris.

The supernatant (200 µL/well) was filtered through an AcroPrep 96-well filter plate (0.2 µm) into a shallow 96-well plate for reverse-phase HPLC analysis to determine the yield.

**Flask Expression of P411 Variants**

*E. coli* (E. cloni BL21(DE3)) cells carrying plasmid encoding the appropriate P411 variant were grown overnight in 5 mL Luria-Bertani medium supplemented with 0.1 mg/mL ampicillin (LB$_{amp}$). The preculture (1 mL) was used to inoculate 50 mL of Hyperbroth medium supplemented with 0.1 mg/mL ampicillin (HB$_{amp}$) in a 125-mL Erlenmeyer flask. This culture was incubated at 37 °C and 230 rpm for 2.5 hours. The culture was then cooled on ice for 45 min and induced with 0.5 mM IPTG and 1.0 mM ALA (final concentrations). Expression was conducted at 24 °C, 140 rpm for 16–18 hours. Subsequently, the *E. coli* cells were pelleted by centrifugation (5,000 g, 5 min, and 4 °C). Media were removed, and the pellets were resuspended to an optical density at 600 nm (OD$_{600}$) of 30 in M9-N minimal medium. Aliquots of the cell suspension (3–4 mL) were used to determine protein concentration after lysis by sonication.

**Hemochrome Assay for the Determination of Hemoprotein Concentration**

Protein concentration in the cell was determined using the hemochrome assay on cell lysate.[3] Lysate was obtained by sonication using using a Qsonica Q500 sonicator (6 minutes, 1 second on, 2 seconds off, 35% amplitude, on wet ice). The cell debris was removed by centrifugation (14,000 g, 10 minutes, 4 °C). To a cuvette, 500 µL of the lysate and 500 µL of solution I [0.2 M NaOH, 40% (v/v) pyridine, 0.5 mM K$_3$Fe(CN)$_6$] were added. The UV-Vis spectrum (380–620 nm) of the oxidized state Fe(III) was recorded immediately. Sodium

dithionite (10 μL of 0.5 M solution in water) was added, and the UV-Vis spectrum of the reduced state Fe(II) was recorded immediately. The protein concentration was calculated using the extinction coefficient and dilution factor (2× dilution in volume): $\varepsilon\_[557_{reduced} - 540_{oxidized}] = 23.98$ mM$^{-1}$cm$^{-1}$. The hemochrome assay detects total heme level, which is a good approximation of over-expressed heme enzyme.

**Biotransformation Using Whole *E. coli* Cells**

Suspension of *E. coli* (*E. cloni* BL21(DE3)) cells expressing the appropriate hemoprotein variant in M9-N buffer (typically $OD_{600} = 2.5$) were transferred to a reaction vial. Enzymatic reactions were then set up in an anaerobic chamber. To an Agilent's 2-mL screw top vial were added degassed suspension of *E. coli* expressing P411 variant (typically $OD_{600} = 2.5$, 395 μL) and substrate (typically 5 μL of 200 mM stock solution in acetonitrile). The final volume of the biotransformation was set to be 400 μL, with 1.25% (v/v) acetonitrile. The reaction vials were then capped and shaken in the anaerobic chamber at room temperature and 600 rpm for 24–48 hours. After the completion of the reaction, 600 μL acetonitrile containing 0.67 mM 1,3,5-trimethoxybenzene internal standard were added to the vial. The resulting mixture was transferred to a 1.7-mL microcentrifuge tube, vigorously vortexed, and centrifuged (14,000 g, 5 min, 4 °C). A sample of the supernatant (0.6 mL) was transferred to a vial for HPLC analysis.

To determine the enantiomeric excess (e.e.), three parallel analytical reactions were conducted. Each reaction mixture was extracted with an equal volume of a hexane/ethyl acetate (3:1) solution. The organic extracts were combined and transferred to a 2-dram vial,

and the solvent was removed by blowing air. Residue was resuspended in 100 μL

isopropanol and subjected to normal-phase HPLC to determine the e.e.

**Enzymatic Preparative Synthesis**

The *E. coli* cell suspension harboring a P411 variant was prepared as described in **Sections**

**1.5** and **1.6**. The cell suspension (*E. coli* whole cells harboring P411 variants, $OD_{600} = 2.5$,

suspended in M9-N aqueous buffer) was placed in a flask covered loosely with aluminum

foil and the bubbled with argon for at least 15 minutes. The reaction flask was then

transferred to an anaerobic chamber. To the reaction flask with cell suspension ($OD_{600} = 2.5$),

substrate (200 mM acetonitrile stock solution, final concentration = 2.5 mM, 1.0 equiv.) was

added to make 1.25% v/v acetonitrile (co-solvent). The reaction vial was immediately capped

and sealed with parafilm and stirred in the anaerobic chamber at room temperature for 48

hours. The reaction solution was filtered through celite then extracted with an equal volume

of ethyl acetate (× 3 times). The combined organic layer was then washed with brine and

dried over anhydrous $MgSO_4$. The solvent was removed *in vacuo*, and the product was

purified by flash chromatography.

**Discovery of Initial Activity and Directed Evolution**

**Table 4-S1.** Initial activity screening with engineered P411s for $\gamma$-lactone synthesis.[a]
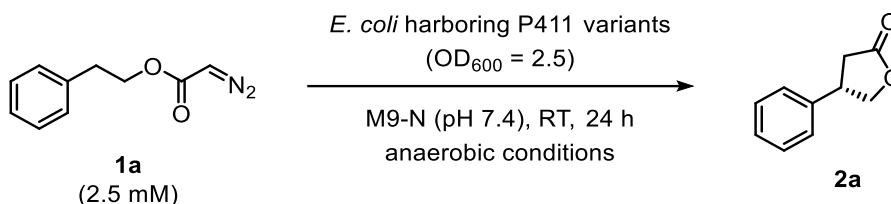


A panel of cytochrome P450 and P411 (P450 with an axial serine ligand) variants previously engineered for carbene nitrene transformations were screened in whole *E. coli* cells against **1a** under anaerobic conditions.[4-6] Among these, P411-LAS-5239 (**C10**), known for its promiscuous activities in different carbene-transfer reactions such as internal alkyne cyclopropenation[5] and lactone-carbene C–H insertion,[4] was effective in producing the $\gamma$-lactone product **2a** (8% yield).

| Entry | Variant / Catalyst | Yield of 3a[a] | Standard deviation of yield | TTN | Standard deviation of TTN |
|---|---|---|---|---|---|
| **1** | P411-LAS-5239 | 7.5% | 0.5% | 530 | 35 |
| **2** | [b] hemin (20 μM) | N.D. | - | N.D. | - |
| **3** | [c] hemin (20 μM) + Na₂S₂O₄ | N.D. | - | N.D. | - |
| **4** | [c] hemin (20 μM) + BSA (20 μM) | N.D. | - | N.D. | - |
| **5** | [c] hemin (20 μM) + Na₂S₂O₄ + BSA (20 μM) | N.D. | - | N.D. | - |
| **6** | [d] cellular background | N.D. | - | N.D. | - |

[a] Experiments were performed using whole *E. coli* cells according to the protocol described in **Sections 1.3** and **1.4**. The yields of **2a** were calculated based on comparing the signal integration ratio of the products and an added internal standard compound (see **Section 5** for more details). All reactions were performed in triplicate, with the reported yields representing the average of three experiments. The signal calibration curves correlating the products and the internal standard compound can be found in **Section 5**. To enhance systematic management of different enzyme variants within the Arnold lab, we recently implemented a new nomenclature system. Variants are named as follows: family name-chemistry abbreviation-entry code. All enzyme variants in this study follow this nomenclature.

[b] Negative control experiments (with free hemin) were performed without the addition of $Na_2S_2O_4$. Under this condition, the resting oxidation state of hemin in aqueous buffer should be Fe(III).

[c] Negative control experiments using free hemin under reduced conditions (Fe(II)) were performed using an excess amount of $Na_2S_2O_4$ (20 mM).

[d] Cellular background control experiments were performed in whole-cell format, using *E. coli* (*E. cloni* BL21(DE3)) cells harboring an engineered tryptophan synthase $\beta$-subunit (Tm9D8*). The gene of this enzyme was also cloned into the pET22b(+) vector (Novagen) between restriction sites *Nde*I and *Xho*I. The protein expression protocol for this experiment followed the standard P411 expression conditions as described in **Sections 1.5**, **1.6**, and **1.7**.

N.D. – no product was detected.

**Table 4-S2.** Directed evolution of P411-LAS-5247 for $\gamma$-lactone synthesis.[a]



**1a**
(2.5 mM)

*E. coli* harboring P411 variants
($OD_{600}$ = 2.5)

M9-N (pH 7.4), RT, 24 h
anaerobic conditions

**2a**
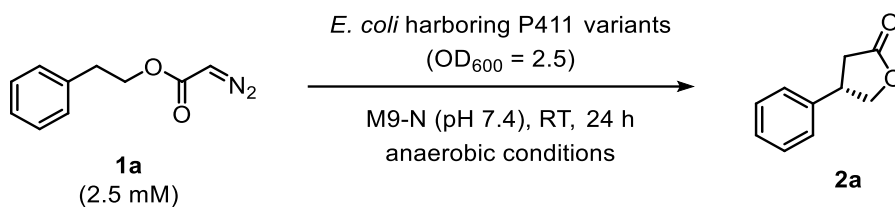
| Entry | Round # | Variant | Yield of 2a[a] | Standard deviation of yield | TTN | Standard deviation of TTN |
|-------|---------|---------|----------------|-----------------------------|-----|---------------------------|
| 1 | 0 | P411-LAS-5239 | 7.5% | 0.5% | 530 | 35 |
| 2 | 1 | P411-LAS-5240 | 14% | 1.7% | 2200 | 260 |
| 3 | 2 | P411-LAS-5241 | 20% | 0.1% | 1100 | 8 |
| 4 | 3 | P411-LAS-5242 | 35% | 0.4% | 2700 | 28 |
| 5 | 4 | P411-LAS-5243 | 41% | 1.5% | 2900 | 200 |
| 6 | 5 | P411-LAS-5244 | 52% | 1.5% | 3900 | 120 |
| 7 | 6 | P411-LAS-5245 | 64% | 2.9% | 4100 | 190 |
| 8 | 7 | P411-LAS-5246 | 65% | 2.0% | 4900 | 150 |
| 9 | 8 | P411-LAS-5247 | 66% | 0.8% | 3800 | 49 |

[a] Experiments were performed using a suspension of *E. coli* cells harboring enzyme variants prepared according to the protocol described in **Sections 1.5**, **1.6**, and **1.7**. Yields were calculated from HPLC calibration curves and the average of triplicate experiments (n = 3).

**Table 4-S3.** Directed evolution of P411-LAS-5246 for $\delta$-lactone **4a** synthesis.[a]
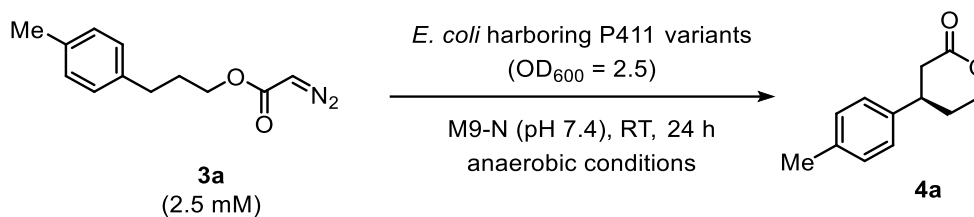


| Entry | Round # | Variant | Yield of 4a[a] | Standard deviation of yield | TTN | Standard deviation of TTN |
|-------|---------|---------|----------------|------------------------------|-----|----------------------------|
| **1** | 0 | P411-LAS-5244 [b] | 52% | 12.4% | 3900 | 930 |
| **2** | 1 | P411-LAS-5248 | 50% | 6.2% | 2900 | 360 |
| **3** | 2 | P411-LAS-5249 [c] | 50% | 1.1% | 2800 | 60 |

[a] Experiments were performed using a suspension of *E. coli* cells harboring enzyme variants prepared according to the protocol described in **Sections 1.5**, **1.6**, and **1.7**. Yields were calculated from HPLC calibration curves and the average of triplicate experiments (n = 3).

[b] 46% e.e. for product **4a**.

[c] 83% e.e. for product **4a**.

**Table 4-S4.** Directed evolution of P411-LAS-5265 for $\varepsilon$-lactone **6a** synthesis.[a]
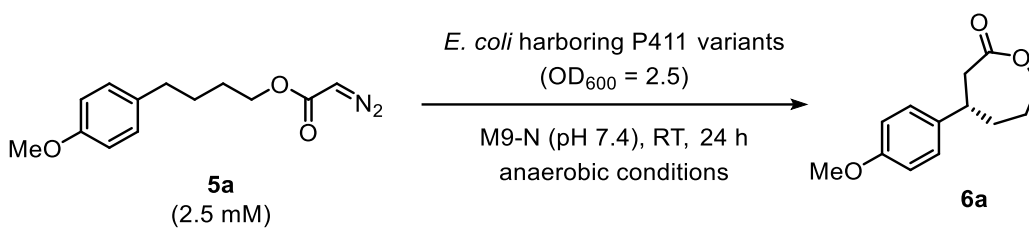


| Entry | Round # | Variant | Yield of 6a[a] | Standard deviation of yield | TTN | Standard deviation of TTN |
|---|---|---|---|---|---|---|
| **1** | 0 | P411-LAS-5250 | 16% | 1.9% | 940 | 110 |
| **2** | 1 | P411-LAS-5251 | 26% | 1.2% | 1100 | 52 |
| **3** | 2 | P411-LAS-5252 | 21% | 0.5% | 1100 | 28 |
| **4** | 3 | P411-LAS-5253 | 31% | 1.9% | 1500 | 94 |
| **5** | 4 | P411-LAS-5254 | 31% | 4.0% | 1300 | 170 |
| **6** | 5 | P411-LAS-5255 | 36% | 3.8% | 1500 | 160 |
| **7** | 6 | P411-LAS-5256 | 44% | 2.3% | 1700 | 90 |
| **8** | 7 | P411-LAS-5257 | 52% | 2.7% | 2200 | 110 |
| **9** | 8 | P411-LAS-5258 | 45% | 2.2% | 2300 | 120 |
| **10** | 9 | P411-LAS-5259 [b] | 58% | 0.3% | 2400 | 14 |
| **11** | 10 | P411-LAS-5260 [c] | 12% | 0.3% | 700 | 19 |
| **12** | 11 | P411-LAS-5261 | 18% | 0.8% | 1200 | 51 |
| **13** | 12 | P411-LAS-5262 | 33% | 1.1% | 2700 | 93 |
| **14** | 13 | P411-LAS-5263 | 39% | 2.2% | 2900 | 160 |
| **15** | 14 | P411-LAS-5264 [d] | 54% | 0.5% | 3100 | 27 |

ᵃ Experiments were performed using a suspension of *E. coli* cells harboring enzyme variants prepared according to the protocol described in **Sections 1.5**, **1.6**, and **1.7**. Yields were calculated from HPLC calibration curves and the average of triplicate experiments (n = 3).

ᵇ 51% e.e. for product **6a**.

ᶜ 96% e.e. for product **6a**.

ᵈ 95% e.e. for product **6a**.

**Nucleotide and Amino Acid Sequences**

The genes encoding the heme proteins shown below were cloned using Gibson assembly[1] into vector pET-22b(+) (Novagen) between restriction sites *Nde*I and *Xho*I in frame with a *C*-terminal 6×His-tag.

**Table 4-S5.** Mutations of the P411 variants in this study.

| Name | Mutations relative to the wild-type P450$_{BM3}$ | Mutations relative to P411-C10 |
|---|---|---|
| P411-LAS-5239 | N70E, A74G, V78L, A82L, F87A, M118S, P142S, F162L, T175I, M177L, A184V, S226R, H236Q, E252G, I263Y, H266V, T268G, A290V, A328V, A330Y, L353V, I366V, C400S, I401L, T436L, L437Q, E442K, ΔFAD | - |
| P411-LAS-5240 | N70E, A74G, V78L, A82L, F87A, M118S, P142S, F162L, T175I, M177L, A184V, S226R, H236Q, E252G, I263Y, H266V, T268G, A290V, A328V, A330Y, L353V, I366V, C400S, I401L, T436L, **L437I**, E442K, ΔFAD | **Q437I** |

| | | |
|---|---|---|
| P411-LAS-5241 | N70E, A74G, V78L, A82L, F87A, M118S, P142S, F162L, T175I, M177L, A184V, S226R, H236Q, E252G, I263Y, H266V, T268G, A290V, **A328I**, A330Y, L353V, I366V, C400S, I401L, T436L, **L437I**, E442K, ΔFAD | **V328I, Q437I** |
| P411-LAS-5242 | N70E, A74G, **V78M**, A82L, F87A, M118S, P142S, F162L, T175I, M177L, A184V, S226R, H236Q, E252G, I263Y, H266V, T268G, A290V, **A328I**, A330Y, L353V, I366V, C400S, I401L, T436L, **L437I**, E442K, ΔFAD | **L78M, V328I, Q437I** |
| P411-LAS-5243 | N70E, A74G, **V78M**, A82L, F87A, M118S, P142S, F162L, T175I, M177L, A184V, S226R, H236Q, E252G, I263Y, H266V, T268G, A290V, **A328I**, A330Y, L353V, I366V, C400S, I401L, **T436R**, **L437I**, E442K, ΔFAD | **L78M, V328I, L436R, Q437I** |
| P411-LAS-5244 | N70E, A74G, **V78M**, A82L, F87A, M118S, P142S, F162L, T175I, M177L, A184V, S226R, H236Q, E252G, I263Y, H266V, T268G, A290V, **A328I**, A330Y, L353V, I366V, C400S, **I401V**, **T436R**, **L437I**, E442K, ΔFAD | **L78M, V328I, L401V, L436R, Q437I** |
| P411-LAS-5245 | N70E, A74G, **V78M**, A82L, F87A, M118S, P142S, **F162I**, T175I, M177L, A184V, S226R, H236Q, E252G, I263Y, H266V, T268G, A290V, **A328I**, A330Y, L353V, I366V, C400S, **I401V**, **T436R**, **L437I**, E442K, ΔFAD | **L78M, L162I, V328I, L401V, L436R, Q437I** |
| P411-LAS-5246 | N70E, A74G, **V78M**, A82L, F87A, M118S, P142S, **F162I**, T175I, M177L, A184V, **R190L**, S226R, H236Q, E252G, I263Y, H266V, T268G, A290V, **A328I**, A330Y, L353V, I366V, C400S, **I401V**, **T436R**, **L437I**, E442K, ΔFAD | **L78M, L162I, R190L, V328I, L401V, L436R, Q437I** |
| P411-LAS-5247 | **N70S,** A74G, **V78M**, A82L, F87A, M118S, P142S, **F162I**, T175I, M177L, A184V, **R190L**, S226R, H236Q, E252G, I263Y, H266V, T268G, A290V, **A328I**, A330Y, L353V, I366V, C400S, **I401V**, **T436R**, **L437I**, E442K, ΔFAD | **E70S, L78M, L162I, R190L, V328I, L401V, L436R, Q437I** |

| P411-LAS-5248 | N70E, A74G, **V78M**, A82L, F87A, M118S, P142S, F162L, T175I, M177L, A184V, S226R, H236Q, E252G, I263Y, H266V, T268G, A290V, **A328V**, A330Y, L353V, I366V, C400S, **I401V**, **T436R**, **L437I**, E442K, ΔFAD | **L78M, L401V, L436R, Q437I** |
|---|---|---|
| P411-LAS-5249 | N70E, A74G, **V78M**, A82L, F87A, **T88S,** M118S, P142S, F162L, T175I, M177L, A184V, S226R, H236Q, E252G, I263Y, H266V, T268G, A290V, **A328V**, A330Y, L353V, I366V, C400S, **I401V**, **T436R**, **L437I**, E442K, ΔFAD | **L78M, T88S, L401V, L436R, Q437I** |
| P411-LAS-5250 | N70E, S72F, V78L, A82L, F87A, M118S, P142S, F162L, T175I, M177L, A184V, S226R, H236Q, E252G, I263W, H266V, T268G, A290V, A328V, A330Y, S332G, L353V, I366V, C400S, I401L, T436R, L437I, E442K, ΔFAD | **S72F, Y263W, S332G, L436R, Q437I** |
| P411-LAS-5251 | N70E, S72F, V78L, A82L, F87A, M118S, P142S, F162L, T175I, M177L, A184V, S226R, H236Q, E252G, I263W, H266V, T268G, A290V, **T327P**, A328V, A330Y, S332G, L353V, I366V, C400S, I401L, T436R, L437I, E442K, ΔFAD | **S72F, Y263W, T327P, S332G, L436R, Q437I** |
| P411-LAS-5252 | N70E, **S72V**, V78L, A82L, F87A, M118S, P142S, F162L, T175I, M177L, A184V, S226R, H236Q, E252G, I263W, H266V, T268G, A290V, **T327P**, A328V, A330Y, S332G, L353V, I366V, C400S, I401L, T436R, L437I, E442K, ΔFAD | **S72V, Y263W, T327P, S332G, L436R, Q437I** |
| P411-LAS-5253 | N70E, **S72V**, V78L, A82L, F87A, M118S, P142S, F162L, T175I, M177L, A184V, S226R, H236Q, E252G, I263W, H266V, T268G, A290V, **T327P**, A328V, **P329P (silent mutation)**, A330Y, S332G, L353V, I366V, C400S, I401L, T436R, L437I, E442K, ΔFAD | **S72V, Y263W, T327P, S332G, L436R, Q437I** |

| P411-LAS-5254 | N70E, **S72V**, V78L, A82L, F87A, M118S, P142S, F162L, T175I, M177L, A184V, H236Q, E252G, I263W, H266V, T268G, A290V, **T327P**, A328V, A330Y, S332G, L353V, I366V, C400S, I401L, T436R, L437I, E442K, ΔFAD | **S72V, R226S, Y263W, T327P, S332G, L436R, Q437I** |
|---|---|---|
| P411-LAS-5255 | N70E, **S72V**, V78L, A82L, F87A, **H92F**, M118S, P142S, F162L, T175I, M177L, A184V, H236Q, E252G, I263W, H266V, T268G, A290V, **T327P**, A328V, A330Y, S332G, L353V, I366V, C400S, I401L, T436R, L437I, E442K, ΔFAD | **S72V, H92F, R226S, Y263W, T327P, S332G, L436R, Q437I** |
| P411-LAS-5256 | N70E, **S72V**, V78L, A82L, F87A, **H92F**, M118S, P142S, T175I, M177L, A184V, H236Q, E252G, I263W, H266V, T268G, A290V, **T327P**, A328V, A330Y, S332G, L353V, I366V, C400S, I401L, T436R, L437I, E442K, ΔFAD | **S72V, H92F, L162F, R226S, Y263W, T327P, S332G, L436R, Q437I** |
| P411-LAS-5257 | N70E, **S72V**, V78L, A82L, F87A, **H92F**, M118S, P142S, T175I, M177L, A184V, H236Q, **E252R**, I263W, H266V, T268G, A290V, **T327P**, A328V, A330Y, S332G, L353V, I366V, C400S, I401L, T436R, L437I, E442K, ΔFAD | **S72V, H92F, L162F, R226S, G252R, Y263W, T327P, S332G, L436R, Q437I** |
| P411-LAS-5258 | N70E, **S72V**, V78L, A82L, F87A, **H92F**, M118S, P142S, T175I, M177L, A184V, H236Q, **E252R**, I263W, H266V, T268G, A290V, **T327P**, A328V, A330Y, S332G, L353V, C400S, I401L, T436R, L437I, E442K, ΔFAD | **S72V, H92F, L162F, R226S, G252R, Y263W, T327P, S332G, V366I, L436R, Q437I** |
| P411-LAS-5259 | N70E, **S72V**, V78L, A82L, F87A, **H92F**, M118S, **P142G**, T175I, M177L, A184V, H236Q, **E252R**, I263W, H266V, T268G, A290V, **T327P**, A328V, A330Y, S332G, L353V, C400S, I401L, T436R, L437I, E442K, ΔFAD | **S72V, H92F, S142G, L162F, R226S, G252R, Y263W, T327P,** |

| | | S332G, V366I, L436R, Q437I |
|---|---|---|
| P411-LAS-5260 | N70E, **S72V**, V78L, A82L, **F87V**, **H92F**, M118S, **P142G**, T175I, M177L, A184V, H236Q, **E252R**, I263W, H266V, T268G, A290V, **T327P**, A328V, A330Y, S332G, L353V, C400S, I401L, T436R, L437I, E442K, ΔFAD | **S72V, A87V, H92F, S142G, L162F, R226S, G252R, Y263W, T327P, S332G, V366I, L436R, Q437I** |
| P411-LAS-5261 | N70E, **S72V**, V78L, A82L, **F87V**, **H92F**, M118S, **P142G**, T175I, M177L, A184V, H236Q, **E252R**, I263W, H266V, T268G, A290V, **T327P**, **A328I**, A330Y, S332G, L353V, C400S, I401L, T436R, L437I, E442K, ΔFAD | **S72V, A87V, H92F, S142G, L162F, R226S, G252R, Y263W, T327P, V328I, S332G, V366I, L436R, Q437I** |
| P411-LAS-5262 | N70E, **S72V**, V78L, **D80E**, A82L, **F87V**, **H92F**, M118S, **P142G**, T175I, M177L, A184V, **I219I (silent mutation)**, H236Q, **G240R**, **E252R**, I263W, H266V, T268G, A290V, **T327P**, **A328I**, A330Y, S332G, L353V, C400S, I401L, T436R, L437I, E442K, ΔFAD | **S72V, D80E, A87V, H92F, S142G, L162F, R226S, G240R, G252R, Y263W, T327P, V328I, S332G, V366I, L436R, Q437I** |
| P411-LAS-5263 | **I2T**, N70E, **S72V**, V78L, **D80E**, A82L, **F87V**, **H92F**, M118S, **P142G**, T175I, M177L, A184V, H236Q, **G240R**, **E252R**, I263W, H266V, T268G, **F279L**, A290V, **T327P**, **A328I**, A330Y, S332G, L353V, C400S, I401L, T436R, L437I, E442K, ΔFAD | **I2T, S72V, D80E, A87V, H92F, S142G, L162F, R226S, G240R, G252R,** |

| | | |
|---|---|---|
| | | **Y263W, F279L, T327P, V328I, S332G, V366I, L436R, Q437I** |
| P411-LAS-5264 | **I2T**, N70E, **S72V**, V78L, **D80E**, A82L, **F87V**, **H92F**, M118S, **P142G**, T175I, M177L, A184V, **A191A (silent mutation)**, H236Q, **G240R**, **E252R**, I263W, H266V, T268G, **F279L**, A290V, **T327P**, **A328I**, A330Y, S332G, L353V, C400S, I401L, T436R, L437I, E442K, ΔFAD | **I2T, S72V, D80E, A87V, H92F, S142G, L162F, R226S, G240R, G252R, Y263W, F279L, T327P, V328I, S332G, V366I, L436R, Q437I** |
| P411-LAS-5265 | N70E, A74G, V78L, A82L, F87A, M118S, P142S, F162L, T175I, M177L, A184V, S226R, H236Q, E252G, **I263W**, H266V, T268G, A290V, A328V, A330Y, L353V, I366V, C400S, I401L, T436L, **L437I**, E442K, ΔFAD | **Y263W, Q437I** |
| P411-LAS-5266 | N70E, A74G, V78L, A82L, **F87P**, M118S, P142S, F162L, T175I, M177L, A184V, S226R, H236Q, E252G, I263Y, **A264S**, H266V, **E267D**, T268G, A290V, **T327P**, A328V, A330Y, **S332A**, L353V, I366V, C400S, I401L, T436L, E442K ΔFAD | **A87P, A264S, E267D, T327P, S332A, Q437L** |
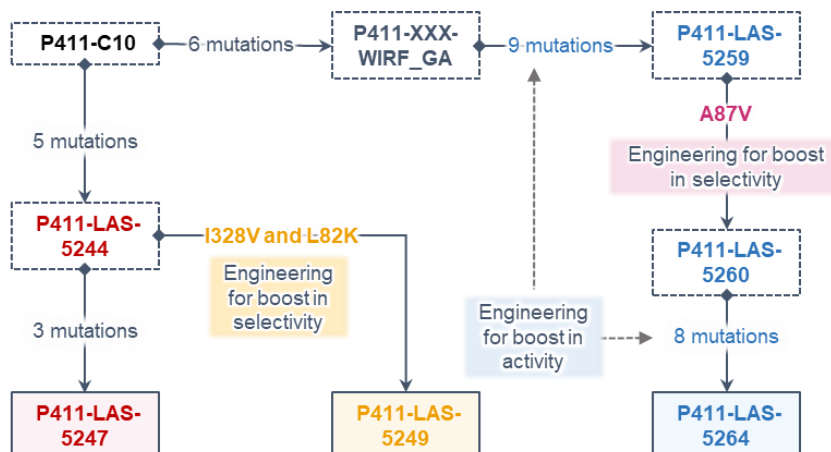
**Figure 4-S1. Evolutionary trajectories of lactone synthases P411-LAS-5247, P411-LAS-5249, and P411-LAS-5264 from P411-C10.**
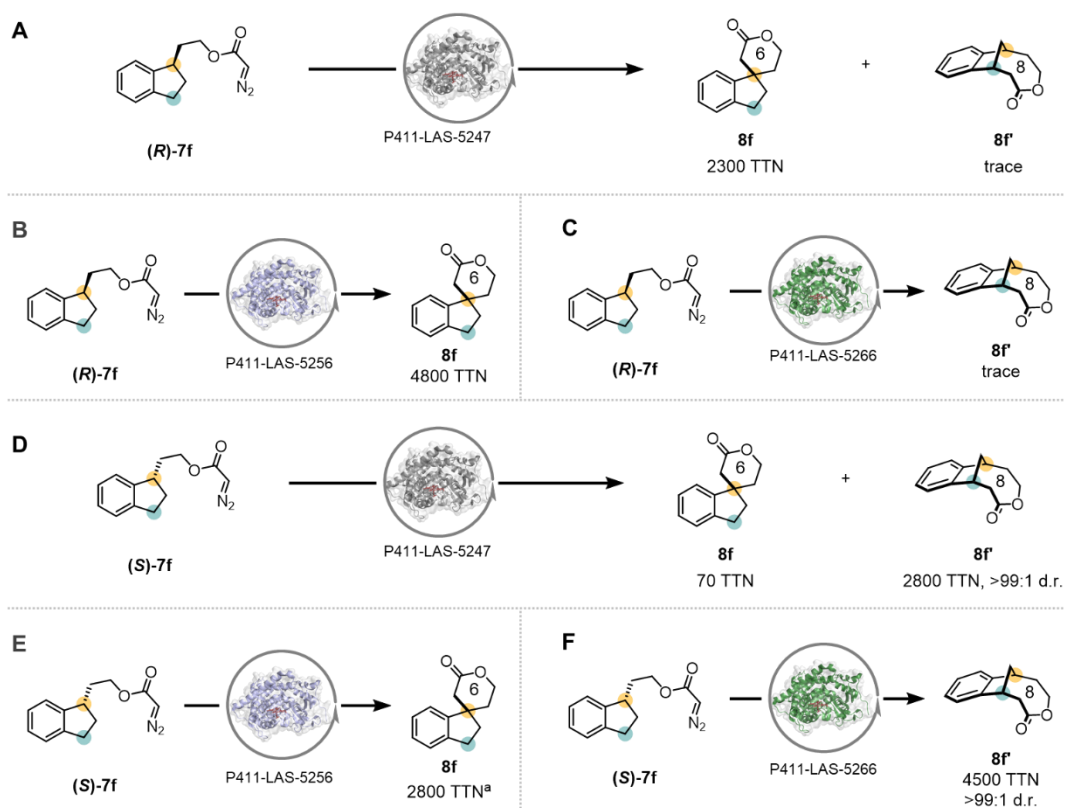


**Figure 4-S2. Reaction of stereopure 7f with P411-LAS-5247, P411-LAS-5256, and P411-LAS-5266.**

[a]550 TTN for product **8f'**

**DNA and amino acid sequences of P411-LAS-5239 (P411-C10)[5]:**

ATGACAATTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATTT
ACCGTTATTAAACACAGATAAACCGGTTCAAGCTTTGATGAAAATTGCGGATG
AATTAGGAGAAATCTTTAAATTCGAGGCGCCTGGTCGTGTAACGCGCTACTTA
TCAAGTCAGCGTCTAATTAAAGAAGCATGCGATGAATCACGCTTTGATAAAGA
GTTAAGTCAAGGTCTGAAATTTCTGCGTGATTTTCTTGGAGACGGGTTAGCCA
CAAGCTGGACGCATGAAAAAATTGGAAAAAAGCGCATAATATCTTACTTCC
AAGCTTTAGTCAGCAGGCAATGAAAGGCTATCATGCGAGTATGGTCGATATCG
CCGTGCAGCTTGTTCAAAAGTGGGAGCGTCTAAATGCAGATGAGCATATTGAA
GTATCGGAAGACATGACACGTTTAACGCTTGATACAATTGGTCTTTGCGGCTTT
AACTATCGCCTTAACAGCTTTTACCGAGATCAGCCTCATCCATTTATTATAAGT
CTGGTCCGTGCACTGGATGAAGTAATGAACAAGCTGCAGCGAGCAAATCCAG
ACGACCCAGCTTATGATGAAAACAAGCGCCAGTTTCAAGAAGATATCAAGGT
GATGAACGACCTAGTAGATAAAATTATTGCAGATCGCAAAGCAAGGGGTGAA
CAAAGCGATGATTTATTAACGCAGATGCTAAACGGAAAAGATCCAGAAACGG
GTGAGCCGCTTGATGACGGGAACATTCGCTATCAAATTATTACATTCTTATATG
CGGGAGTTGAAGGTACAAGTGGTCTTTTATCATTTGCGCTGTATTTCTTAGTGA
AAAATCCACATGTATTACAAAAAGTAGCAGAAGAAGCAGCACGAGTTCTAGT
AGATCCTGTTCCAAGCTACAAACAAGTCAAACAGCTTAAATATGTCGGCATGG
TCTTAAACGAAGCGCTGCGCTTATGGCCAACGGTTCCTTATTTTTCCCTATATG
CAAAAGAAGATACGGTGCTTGGAGGAGAATATCCTTTAGAAAAAGGCGACGA
AGTAATGGTTCTGATTCCTCAGCTTCACCGTGATAAAACAGTTTGGGGAGACG
ATGTGGAGGAGTTCCGTCCAGAGCGTTTTGAAAATCCAAGTGCGATTCCGCAG
CATGCGTTTAAACCGTTTGGAAACGGTCAGCGTGCGTCTCTGGGTCAGCAGTT
CGCTCTTCATGAAGCAACGCTGGTACTTGGTATGATGCTAAAACACTTTGACTT
TGAAGATCATACAAACTACGAGCTCGATATTAAAGAACTGCAGACGTTAAAA
CCTAAAGGCTTTGTGGTAAAAGCAAATCGAAAAAAATTCCGCTTGGCGGTAT
TCCTTCACCTAGCACTGAACAGTCTGCTAAAAAGTACGCAAAAGGCAGAA
AACGCTCATAATACGCCGCTGCTTGTGCTATACGGTTCAAATATGGGTACCGC

TGAAGGAACGGCGCGTGATTTAGCAGATATTGCAATGAGCAAAGGATTTGCA
CCGCAGGTCGCAACGCTTGATTCACACGCCGGAAATCTTCCGCGCGAAGGAGC
TGTATTAATTGTAACGGCGTCTTATAACGGTCATCCGCCTGATAACGCAAAGC
AATTTGTCGACTGGTTAGACCAAGCGTCTGCTGATGAAGTAAAAGGCGTTCGC
TACTCCGTATTTGGATGCGGCGATAAAAACTGGGCTACTACGTATCAAAAAGT
GCCTGCTTTTATCGATGAAACGCTTGCCGCTAAAGGGGCAGAAAACATCGCTG
ACCGCGGTGAAGCAGATGCAAGCGACGACTTTGAAGGCACATATGAAGAATG
GCGTGAACATATGTGGAGTGACGTAGCAGCCTACTTTAACCTCGACATTGAAA
ACAGTGAAGATAATAAATCTACTCTTTCACTTCAATTTGTCGACAGCGCCGCG
GATATGCCGCTTGCGAAAATGCACGGTGCGTTTTCAACGCTCGAGCACCACCA
CCACCACCACTGA

MTIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQ
RLIKEACDESRFDKELSQGLKFLRDFLGDGLATSWTHEKNWKKAHNILLPSFSQQ
AMKGYHASMVDIAVQLVQKWERLNADEHIEVSEDMTRLTLDTIGLCGFNYRLNS
FYRDQPHPFIISLVRALDEVMNKLQRANPDDPAYDENKRQFQEDIKVMNDLVDKI
IADRKARGEQSDDLLTQMLNGKDPETGEPLDDGNIRYQIITFLYAGVEGTSGLLSF
ALYFLVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRLWPTV
PYFSLYAKEDTVLGGEYPLEKGDEVMVLIPQLHRDKTVWGDDVEEFRPERFENPS
AIPQHAFKPFGNGQRASLGQQFALHEATLVLGMMLKHFDFEDHTNYELDIKELQT
LKPKGFVVKAKSKKIPLGGIPSPSTEQSAKKVRKKAENAHNTPLLVLYGSNMGTA
EGTARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNGHPPDNAKQF
VDWLDQASADEVKGVRYSVFGCGDKNWATTYQKVPAFIDETLAAKGAENIADR
GEADASDDFEGTYEEWREHMWSDVAAYFNLDIENSEDNKSTLSLQFVDSAADMP
LAKMHGAFSTLEHHHHHH*

**DNA and amino acid sequences of P411-LAS-5244:**

ATGACAATTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATTT
ACCGTTATTAAACACAGATAAACCGGTTCAAGCTTTGATGAAAATTGCGGATG
AATTAGGAGAAATCTTTAAATTCGAGGCGCCTGGTCGTGTAACGCGCTACTTA
TCAAGTCAGCGTCTAATTAAAGAAGCATGCGATGAATCACGCTTTGATAAAGA
GTTAAGTCAAGGTCTGAAATTTATGCGTGATTTTCTTGGAGACGGGTTAGCCA
CAAGCTGGACGCATGAAAAAATTGGAAAAAAGCGCATAATATCTTACTTCC
AAGCTTTAGTCAGCAGGCAATGAAAGGCTATCATGCGAGTATGGTCGATATCG
CCGTGCAGCTTGTTCAAAAGTGGGAGCGTCTAAATGCAGATGAGCATATTGAA
GTATCGGAAGACATGACACGTTTAACGCTTGATACAATTGGTCTTTGCGGCTTT
AACTATCGCCTTAACAGCTTTTACCGAGATCAGCCTCATCCATTTATTATAAGT
CTGGTCCGTGCACTGGATGAAGTAATGAACAAGCTGCAGCGAGCAAATCCAG
ACGACCCAGCTTATGATGAAAACAAGCGCCAGTTTCAAGAAGATATCAAGGT
GATGAACGACCTAGTAGATAAAATTATTGCAGATCGCAAAGCAAGGGGTGAA
CAAAGCGATGATTTATTAACGCAGATGCTAAACGGAAAAGATCCAGAAACGG
GTGAGCCGCTTGATGACGGGAACATTCGCTATCAAATTATTACATTCTTATATG
CGGGAGTTGAAGGTACAAGTGGTCTTTTATCATTTGCGCTGTATTTCTTAGTGA
AAAATCCACATGTATTACAAAAAGTAGCAGAAGAAGCAGCACGAGTTCTAGT
AGATCCTGTTCCAAGCTACAAACAAGTCAAACAGCTTAAATATGTCGGCATGG
TCTTAAACGAAGCGCTGCGCTTATGGCCAACGATTCCTTATTTTTCCCTATATG
CAAAAGAAGATACGGTGCTTGGAGGAGAATATCCTTTAGAAAAAGGCGACGA
AGTAATGGTTCTGATTCCTCAGCTTCACCGTGATAAAACAGTTTGGGGAGACG
ATGTGGAGGAGTTCCGTCCAGAGCGTTTTGAAAATCCAAGTGCGATTCCGCAG
CATGCGTTTAAACCGTTTGGAAACGGTCAGCGTGCGTCTGTGGGTCAGCAGTT
CGCTCTTCATGAAGCAACGCTGGTACTTGGTATGATGCTAAAACACTTTGACTT
TGAAGATCATACAAACTACGAGCTCGATATTAAAGAACGTATTACGTTAAAAC
CTAAAGGCTTTGTGGTAAAAGCAAAATCGAAAAAAATTCCGCTTGGCGGTATT
CCTTCACCTAGCACTGAACAGTCTGCTAAAAAAGTACGCAAAAAGGCAGAAA
ACGCTCATAATACGCCGCTGCTTGTGCTATACGGTTCAAATATGGGTACCGCT

GAAGGAACGGCGCGTGATTTAGCAGATATTGCAATGAGCAAAGGATTTGCAC
CGCAGGTCGCAACGCTTGATTCACACGCCGGAAATCTTCCGCGCGAAGGAGCT
GTATTAATTGTAACGGCGTCTTATAACGGTCATCCGCCTGATAACGCAAAGCA
ATTTGTCGACTGGTTAGACCAAGCGTCTGCTGATGAAGTAAAAGGCGTTCGCT
ACTCCGTATTTGGATGCGGCGATAAAAACTGGGCTACTACGTATCAAAAAGTG
CCTGCTTTTATCGATGAAACGCTTGCCGCTAAAGGGGCAGAAAACATCGCTGA
CCGCGGTGAAGCAGATGCAAGCGACGACTTTGAAGGCACATATGAAGAATGG
CGTGAACATATGTGGAGTGACGTAGCAGCCTACTTTAACCTCGACATTGAAAA
CAGTGAAGATAATAAATCTACTCTTTCACTTCAATTTGTCGACAGCGCCGCGG
ATATGCCGCTTGCGAAAATGCACGGTGCGTTTTCAACGCTCGAGCACCACCAC
CACCACCACTGA

MTIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQ
RLIKEACDESRFDKELSQGLKFMRDFLGDGLATSWTHEKNWKKAHNILLPSFSQQ
AMKGYHASMVDIAVQLVQKWERLNADEHIEVSEDMTRLTLDTIGLCGFNYRLNS
FYRDQPHPFIISLVRALDEVMNKLQRANPDDPAYDENKRQFQEDIKVMNDLVDKI
IADRKARGEQSDDLLTQMLNGKDPETGEPLDDGNIRYQIITFLYAGVEGTSGLLSF
ALYFLVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRLWPTIP
YFSLYAKEDTVLGGEYPLEKGDEVMVLIPQLHRDKTVWGDDVEEFRPERFENPSA
IPQHAFKPFGNGQRASVGQQFALHEATLVLGMMLKHFDFEDHTNYELDIKERITL
KPKGFVVKAKSKKIPLGGIPSPSTEQSAKKVRKKAENAHNTPLLVLYGSNMGTAE
GTARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNGHPPDNAKQFV
DWLDQASADEVKGVRYSVFGCGDKNWATTYQKVPAFIDETLAAKGAENIADRG
EADASDDFEGTYEEWREHMWSDVAAYFNLDIENSEDNKSTLSLQFVDSAADMPL
AKMHGAFSTLEHHHHHH*

**DNA and amino acid sequences of P411-LAS-5247:**

ATGACAATTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATTT
ACCGTTATTAAACACAGATAAACCGGTTCAAGCTTTGATGAAAATTGCGGATG
AATTAGGAGAAATCTTTAAATTCGAGGCGCCTGGTCGTGTAACGCGCTACTTA
TCAAGTCAGCGTCTAATTAAAGAAGCATGCGATGAATCACGCTTTGATAAAAG
TTTAAGTCAAGGTCTGAAATTTATGCGTGATTTTCTTGGAGACGGGTTAGCCAC
AAGCTGGACGCATGAAAAAATTGGAAAAAAGCGCATAATATCTTACTTCCA
AGCTTTAGTCAGCAGGCAATGAAAGGCTATCATGCGAGTATGGTCGATATCGC
CGTGCAGCTTGTTCAAAAGTGGGAGCGTCTAAATGCAGATGAGCATATTGAAG
TATCGGAAGACATGACACGTTTAACGCTTGATACAATTGGTCTTTGCGGCTTTA
ACTATCGCATTAACAGCTTTTACCGAGATCAGCCTCATCCATTTATTATAAGTC
TGGTCCGTGCACTGGATGAAGTAATGAACAAGCTGCAGCTTGCAAATCCAGAC
GACCCAGCTTATGATGAAAACAAGCGCCAGTTTCAAGAAGATATCAAGGTGA
TGAACGACCTAGTAGATAAAATTATTGCAGATCGCAAAGCAAGGGGTGAACA
AAGCGATGATTTATTAACGCAGATGCTAAACGGAAAAGATCCAGAAACGGGT
GAGCCGCTTGATGACGGGAACATTCGCTATCAAATTATTACATTCTTATATGC
GGGAGTTGAAGGTACAAGTGGTCTTTTATCATTTGCGCTGTATTTCTTAGTGAA
AAATCCACATGTATTACAAAAAGTAGCAGAAGAAGCAGCACGAGTTCTAGTA
GATCCTGTTCCAAGCTACAAACAAGTCAAACAGCTTAAATATGTCGGCATGGT
CTTAAACGAAGCGCTGCGCTTATGGCCAACGATTCCTTATTTTTCCCTATATGC
AAAAGAAGATACGGTGCTTGGAGGAGAATATCCTTTAGAAAAAGGCGACGAA
GTAATGGTTCTGATTCCTCAGCTTCACCGTGATAAAACAGTTTGGGGAGACGA
TGTGGAGGAGTTCCGTCCAGAGCGTTTTGAAAATCCAAGTGCGATTCCGCAGC
ATGCGTTTAAACCGTTTGGAAACGGTCAGCGTGCGTCTGTTGGTCAGCAGTTC
GCTCTTCATGAAGCAACGCTGGTACTTGGTATGATGCTAAAACACTTTGACTTT
GAAGATCATACAAACTACGAGCTCGATATTAAAGAACGTATTACGTTAAAACC
TAAAGGCTTTGTGGTAAAAGCAAAATCGAAAAAAATTCCGCTTGGCGGTATTC
CTTCACCTAGCACTGAACAGTCTGCTAAAAAAGTACGCAAAAAGGCAGAAAA
CGCTCATAATACGCCGCTGCTTGTGCTATACGGTTCAAATATGGGTACCGCTG

AAGGAACGGCGCGTGATTTAGCAGATATTGCAATGAGCAAAGGATTTGCACC
GCAGGTCGCAACGCTTGATTCACACGCCGGAAATCTTCCGCGCGAAGGAGCTG
TATTAATTGTAACGGCGTCTTATAACGGTCATCCGCCTGATAACGCAAAGCAA
TTTGTCGACTGGTTAGACCAAGCGTCTGCTGATGAAGTAAAAGGCGTTCGCTA
CTCCGTATTTGGATGCGGCGATAAAAACTGGGCTACTACGTATCAAAAAGTGC
CTGCTTTTATCGATGAAACGCTTGCCGCTAAAGGGGCAGAAAACATCGCTGAC
CGCGGTGAAGCAGATGCAAGCGACGACTTTGAAGGCACATATGAAGAATGGC
GTGAACATATGTGGAGTGACGTAGCAGCCTACTTTAACCTCGACATTGAAAAC
AGTGAAGATAATAAATCTACTCTTTCACTTCAATTTGTCGACAGCGCCGCGGA
TATGCCGCTTGCGAAAATGCACGGTGCGTTTTCAACGCTCGAGCACCACCACC
ACCACCACTGA

MTIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQ
RLIKEACDESRFDKSLSQGLKFMRDFLGDGLATSWTHEKNWKKAHNILLPSFSQQ
AMKGYHASMVDIAVQLVQKWERLNADEHIEVSEDMTRLTLDTIGLCGFNYRINS
FYRDQPHPFIISLVRALDEVMNKLQLANPDDPAYDENKRQFQEDIKVMNDLVDKII
ADRKARGEQSDDLLTQMLNGKDPETGEPLDDGNIRYQIITFLYAGVEGTSGLLSFA
LYFLVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRLWPTIPY
FSLYAKEDTVLGGEYPLEKGDEVMVLIPQLHRDKTVWGDDVEEFRPERFENPSAI
PQHAFKPFGNGQRASVGQQFALHEATLVLGMMLKHFDFEDHTNYELDIKERITLK
PKGFVVKAKSKKIPLGGIPSPSTEQSAKKVRKKAENAHNTPLLVLYGSNMGTAEG
TARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNGHPPDNAKQFVD
WLDQASADEVKGVRYSVFGCGDKNWATTYQKVPAFIDETLAAKGAENIADRGE
ADASDDFEGTYEEWREHMWSDVAAYFNLDIENSEDNKSTLSLQFVDSAADMPLA
KMHGAFSTLEHHHHHH*

**DNA and amino acid sequences of P411-LAS-5249:**

ATGACAATTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATTT
ACCGTTATTAAACACAGATAAACCGGTTCAAGCTTTGATGAAAATTGCGGATG
AATTAGGAGAAATCTTTAAATTCGAGGCGCCTGGTCGTGTAACGCGCTACTTA
TCAAGTCAGCGTCTAATTAAAGAAGCATGCGATGAATCACGCTTTGATAAAGA
GTTAAGTCAAGGTCTGAAATTTATGCGTGATTTTCTTGGAGACGGGTTAGCCA
GTAGCTGGACGCATGAAAAAATTGGAAAAAAGCGCATAATATCTTACTTCC
AAGCTTTAGTCAGCAGGCAATGAAAGGCTATCATGCGAGTATGGTCGATATCG
CCGTGCAGCTTGTTCAAAAGTGGGAGCGTCTAAATGCAGATGAGCATATTGAA
GTATCGGAAGACATGACACGTTTAACGCTTGATACAATTGGTCTTTGCGGCTTT
AACTATCGCCTTAACAGCTTTTACCGAGATCAGCCTCATCCATTTATTATAAGT
CTGGTCCGTGCACTGGATGAAGTAATGAACAAGCTGCAGCGAGCAAATCCAG
ACGACCCAGCTTATGATGAAAACAAGCGCCAGTTTCAAGAAGATATCAAGGT
GATGAACGACCTAGTAGATAAAATTATTGCAGATCGCAAAGCAAGGGGGTGAA
CAAAGCGATGATTTATTAACGCAGATGCTAAACGGAAAAGATCCAGAAACGG
GTGAGCCGCTTGATGACGGGAACATTCGCTATCAAATTATTACATTCTTATATG
CGGGAGTTGAAGGTACAAGTGGTCTTTTATCATTTGCGCTGTATTTCTTAGTGA
AAAATCCACATGTATTACAAAAAGTAGCAGAAGAAGCAGCACGAGTTCTAGT
AGATCCTGTTCCAAGCTACAAACAAGTCAAACAGCTTAAATATGTCGGCATGG
TCTTAAACGAAGCGCTGCGCTTATGGCCAACGGTTCCTTATTTTTCCCTATATG
CAAAAGAAGATACGGTGCTTGGAGGAGAATATCCTTTAGAAAAAGGCGACGA
AGTAATGGTTCTGATTCCTCAGCTTCACCGTGATAAAACAGTTTGGGGAGACG
ATGTGGAGGAGTTCCGTCCAGAGCGTTTTGAAAATCCAAGTGCGATTCCGCAG
CATGCGTTTAAACCGTTTGGAAACGGTCAGCGTGCGTCTGTGGGTCAGCAGTT
CGCTCTTCATGAAGCAACGCTGGTACTTGGTATGATGCTAAAACACTTTGACTT
TGAAGATCATACAAACTACGAGCTCGATATTAAAGAACGTATTACGTTAAAAC
CTAAAGGCTTTGTGGTAAAAGCAAAATCGAAAAAAATTCCGCTTGGCGGTATT
CCTTCACCTAGCACTGAACAGTCTGCTAAAAAAGTACGCAAAAAGGCAGAAA
ACGCTCATAATACGCCGCTGCTTGTGCTATACGGTTCAAATATGGGTACCGCT

GAAGGAACGGCGCGTGATTTAGCAGATATTGCAATGAGCAAAGGATTTGCAC
CGCAGGTCGCAACGCTTGATTCACACGCCGGAAATCTTCCGCGCGAAGGAGCT
GTATTAATTGTAACGGCGTCTTATAACGGTCATCCGCCTGATAACGCAAAGCA
ATTTGTCGACTGGTTAGACCAAGCGTCTGCTGATGAAGTAAAAGGCGTTCGCT
ACTCCGTATTTGGATGCGGCGATAAAAACTGGGCTACTACGTATCAAAAAGTG
CCTGCTTTTATCGATGAAACGCTTGCCGCTAAAGGGGCAGAAACATCGCTGA
CCGCGGTGAAGCAGATGCAAGCGACGACTTTGAAGGCACATATGAAGAATGG
CGTGAACATATGTGGAGTGACGTAGCAGCCTACTTTAACCTCGACATTGAAAA
CAGTGAAGATAATAAATCTACTCTTTCACTTCAATTTGTCGACAGCGCCGCGG
ATATGCCGCTTGCGAAAATGCACGGTGCGTTTTCAACGCTCGAGCACCACCAC
CACCACCACTGA

MTIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQ
RLIKEACDESRFDKELSQGLKFMRDFLGDGLASSWTHEKNWKKAHNILLPSFSQQ
AMKGYHASMVDIAVQLVQKWERLNADEHIEVSEDMTRLTLDTIGLCGFNYRLNS
FYRDQPHPFIISLVRALDEVMNKLQRANPDDPAYDENKRQFQEDIKVMNDLVDKI
IADRKARGEQSDDLLTQMLNGKDPETGEPLDDGNIRYQIITFLYAGVEGTSGLLSF
ALYFLVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRLWPTV
PYFSLYAKEDTVLGGEYPLEKGDEVMVLIPQLHRDKTVWGDDVEEFRPERFENPS
AIPQHAFKPFGNGQRASVGQQFALHEATLVLGMMLKHFDFEDHTNYELDIKERIT
LKPKGFVVKAKSKKIPLGGIPSPSTEQSAKKVRKKAENAHNTPLLVLYGSNMGTA
EGTARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNGHPPDNAKQF
VDWLDQASADEVKGVRYSVFGCGDKNWATTYQKVPAFIDETLAAKGAENIADR
GEADASDDFEGTYEEWREHMWSDVAAYFNLDIENSEDNKSTLSLQFVDSAADMP
LAKMHGAFSTLEHHHHHH*

**DNA and amino acid sequences of P411-LAS-5250 (P411-C10-WIRF_GA)[5]:**

ATGACAATTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATTT
ACCGTTATTAAACACAGATAAACCGGTTCAAGCTTTGATGAAAATTGCGGATG
AATTAGGAGAAATCTTTAAATTCGAGGCGCCTGGTCGTGTAACGCGCTACTTA
TCAAGTCAGCGTCTAATTAAAGAAGCATGCGATGAATCACGCTTTGATAAAGA
GTTATTTCAAGCGCTGAAATTTCTGCGTGATTTTCTTGGAGACGGGTTAGCCAC
AAGCTGGACGCATGAAAAAATTGGAAAAAGCGCATAATATCTTACTTCCA
AGCTTTAGTCAGCAGGCAATGAAAGGCTATCATGCGAGTATGGTCGATATCGC
CGTGCAGCTTGTTCAAAAGTGGGAGCGTCTAAATGCAGATGAGCATATTGAAG
TATCGGAAGACATGACACGTTTAACGCTTGATACAATTGGTCTTTGCGGCTTTA
ACTATCGCCTTAACAGCTTTTACCGAGATCAGCCTCATCCATTTATTATAAGTC
TGGTCCGTGCACTGGATGAAGTAATGAACAAGCTGCAGCGAGCAAATCCAGA
CGACCCAGCTTATGATGAAAACAAGCGCCAGTTTCAAGAAGATATCAAGGTG
ATGAACGACCTAGTAGATAAAATTATTGCAGATCGCAAAGCAAGGGGTGAAC
AAAGCGATGATTTATTAACGCAGATGCTAAACGGAAAAGATCCAGAAACGGG
TGAGCCGCTTGATGACGGGAACATTCGCTATCAAATTATTACATTCTTATGGG
CGGGAGTTGAAGGTACAAGTGGTCTTTTATCATTTGCGCTGTATTTCTTAGTGA
AAAATCCACATGTATTACAAAAAGTAGCAGAAGAAGCAGCACGAGTTCTAGT
AGATCCTGTTCCAAGCTACAAACAAGTCAAACAGCTTAAATATGTCGGCATGG
TCTTAAACGAAGCGCTGCGCTTATGGCCAACGGTTCCTTATTTTGGTCTATATG
CAAAAGAAGATACGGTGCTTGGAGGAGAATATCCTTTAGAAAAAGGCGACGA
AGTAATGGTTCTGATTCCTCAGCTTCACCGTGATAAACAGTTTGGGGAGACG
ATGTGGAGGAGTTCCGTCCAGAGCGTTTTGAAAATCCAAGTGCGATTCCGCAG
CATGCGTTTAAACCGTTTGGAAACGGTCAGCGTGCGTCTCTGGGTCAGCAGTT
CGCTCTTCATGAAGCAACGCTGGTACTTGGTATGATGCTAAAACACTTTGACTT
TGAAGATCATACAAACTACGAGCTCGATATTAAGAACGTATTACGTTAAAAC
CTAAAGGCTTTGTGGTAAAAGCAAAATCGAAAAAAATTCCGCTTGGCGGTATT
CCTTCACCTAGCACTGAACAGTCTGCTAAAAAGTACGCAAAAAGGCAGAAA
ACGCTCATAATACGCCGCTGCTTGTGCTATACGGTTCAAATATGGGTACCGCT

GAAGGAACGGCGCGTGATTTAGCAGATATTGCAATGAGCAAAGGATTTGCAC
CGCAGGTCGCAACGCTTGATTCACACGCCGGAAATCTTCCGCGCGAAGGAGCT
GTATTAATTGTAACGGCGTCTTATAACGGTCATCCGCCTGATAACGCAAAGCA
ATTTGTCGACTGGTTAGACCAAGCGTCTGCTGATGAAGTAAAAGGCGTTCGCT
ACTCCGTATTTGGATGCGGCGATAAAAACTGGGCTACTACGTATCAAAAAGTG
CCTGCTTTTATCGATGAAACGCTTGCCGCTAAAGGGGCAGAAAACATCGCTGA
CCGCGGTGAAGCAGATGCAAGCGACGACTTTGAAGGCACATATGAAGAATGG
CGTGAACATATGTGGAGTGACGTAGCAGCCTACTTTAACCTCGACATTGAAAA
CAGTGAAGATAATAAATCTACTCTTTCACTTCAATTTGTCGACAGCGCCGCGG
ATATGCCGCTTGCGAAAATGCACGGTGCGTTTTCAACGCTCGAGCACCACCAC
CACCACCACTGA

MTIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQ
RLIKEACDESRFDKELFQALKFLRDFLGDGLATSWTHEKNWKKAHNILLPSFSQQ
AMKGYHASMVDIAVQLVQKWERLNADEHIEVSEDMTRLTLDTIGLCGFNYRLNS
FYRDQPHPFIISLVRALDEVMNKLQRANPDDPAYDENKRQFQEDIKVMNDLVDKI
IADRKARGEQSDDLLTQMLNGKDPETGEPLDDGNIRYQIITFLWAGVEGTSGLLSF
ALYFLVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRLWPTV
PYFGLYAKEDTVLGGEYPLEKGDEVMVLIPQLHRDKTVWGDDVEEFRPERFENPS
AIPQHAFKPFGNGQRASLGQQFALHEATLVLGMMLKHFDFEDHTNYELDIKERIT
LKPKGFVVKAKSKKIPLGGIPSPSTEQSAKKVRKKAENAHNTPLLVLYGSNMGTA
EGTARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNGHPPDNAKQF
VDWLDQASADEVKGVRYSVFGCGDKNWATTYQKVPAFIDETLAAKGAENIADR
GEADASDDFEGTYEEWREHMWSDVAAYFNLDIENSEDNKSTLSLQFVDSAADMP
LAKMHGAFSTLEHHHHHH*

**DNA and amino acid sequences of P411-LAS-5256:**

ATGACAATTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATTT
ACCGTTATTAAACACAGATAAACCGGTTCAAGCTTTGATGAAAATTGCGGATG
AATTAGGAGAAATCTTTAAATTCGAGGCGCCTGGTCGTGTAACGCGCTACTTA
TCAAGTCAGCGTCTAATTAAAGAAGCATGCGATGAATCACGCTTTGATAAAGA
GTTAGTGCAAGCGCTGAAATTTCTGCGTGATTTTCTTGGAGACGGGTTAGCCA
CAAGCTGGACGTTTGAAAAAATTGGAAAAAGCGCATAATATCTTACTTCCA
AGCTTTAGTCAGCAGGCAATGAAAGGCTATCATGCGAGTATGGTCGATATCGC
CGTGCAGCTTGTTCAAAAGTGGGAGCGTCTAAATGCAGATGAGCATATTGAAG
TATCGGAAGACATGACACGTTTAACGCTTGATACAATTGGTCTTTGCGGCTTTA
ACTATCGCTTTAACAGCTTTTACCGAGATCAGCCTCATCCATTTATTATAAGTC
TGGTCCGTGCACTGGATGAAGTAATGAACAAGCTGCAGCGAGCAAATCCAGA
CGACCCAGCTTATGATGAAAACAAGCGCCAGTTTCAAGAAGATATCAAGGTG
ATGAACGACCTAGTAGATAAAATTATTGCAGATCGCAAAGCAAGTGGTGAAC
AAAGCGATGATTTATTAACGCAGATGCTAAACGGAAAAGATCCAGAAACGGG
TGAGCCGCTTGATGACGGGAACATTCGCTATCAAATTATTACATTCTTATGGG
CGGGAGTTGAAGGTACAAGTGGTCTTTTATCATTTGCGCTGTATTTCTTAGTGA
AAAATCCACATGTATTACAAAAAGTAGCAGAAGAAGCAGCACGAGTTCTAGT
AGATCCTGTTCCAAGCTACAAACAAGTCAAACAGCTTAAATATGTCGGCATGG
TCTTAAACGAAGCGCTGCGCTTATGGCCACCGGTTCCGTATTTTGGTCTATATG
CAAAAGAAGATACGGTGCTTGGAGGAGAATATCCTTTAGAAAAAGGCGACGA
AGTAATGGTTCTGATTCCTCAGCTTCACCGTGATAAAACAGTTTGGGGAGACG
ATGTGGAGGAGTTCCGTCCAGAGCGTTTTGAAAATCCAAGTGCGATTCCGCAG
CATGCGTTTAAACCGTTTGGAAACGGTCAGCGTGCGTCTCTGGGTCAGCAGTT
CGCTCTTCATGAAGCAACGCTGGTACTTGGTATGATGCTAAAACACTTTGACTT
TGAAGATCATACAAACTACGAGCTCGATATTAAAGAACGTATTACGTTAAAAC
CTAAAGGCTTTGTGGTAAAAGCAAAATCGAAAAAAATTCCGCTTGGCGGTATT
CCTTCACCTAGCACTGAACAGTCTGCTAAAAAAGTACGCAAAAAGGCAGAAA
ACGCTCATAATACGCCGCTGCTTGTGCTATACGGTTCAAATATGGGTACCGCT

GAAGGAACGGCGCGTGATTTAGCAGATATTGCAATGAGCAAAGGATTTGCAC
CGCAGGTCGCAACGCTTGATTCACACGCCGGAAATCTTCCGCGCGAAGGAGCT
GTATTAATTGTAACGGCGTCTTATAACGGTCATCCGCCTGATAACGCAAAGCA
ATTTGTCGACTGGTTAGACCAAGCGTCTGCTGATGAAGTAAAAGGCGTTCGCT
ACTCCGTATTTGGATGCGGCGATAAAAACTGGGCTACTACGTATCAAAAAGTG
CCTGCTTTTATCGATGAAACGCTTGCCGCTAAAGGGGCAGAAAACATCGCTGA
CCGCGGTGAAGCAGATGCAAGCGACGACTTTGAAGGCACATATGAAGAATGG
CGTGAACATATGTGGAGTGACGTAGCAGCCTACTTTAACCTCGACATTGAAAA
CAGTGAAGATAATAAATCTACTCTTTCACTTCAATTTGTCGACAGCGCCGCGG
ATATGCCGCTTGCGAAAATGCACGGTGCGTTTTCAACGCTCGAGCACCACCAC
CACCACCACTGA

MTIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQ
RLIKEACDESRFDKELVQALKFLRDFLGDGLATSWTFEKNWKKAHNILLPSFSQQ
AMKGYHASMVDIAVQLVQKWERLNADEHIEVSEDMTRLTLDTIGLCGFNYRFNS
FYRDQPHPFIISLVRALDEVMNKLQRANPDDPAYDENKRQFQEDIKVMNDLVDKI
IADRKASGEQSDDLLTQMLNGKDPETGEPLDDGNIRYQIITFLWAGVEGTSGLLSF
ALYFLVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRLWPPV
PYFGLYAKEDTVLGGEYPLEKGDEVMVLIPQLHRDKTVWGDDVEEFRPERFENPS
AIPQHAFKPFGNGQRASLGQQFALHEATLVLGMMLKHFDFEDHTNYELDIKERIT
LKPKGFVVKAKSKKIPLGGIPSPSTEQSAKKVRKKAENAHNTPLLVLYGSNMGTA
EGTARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNGHPPDNAKQF
VDWLDQASADEVKGVRYSVFGCGDKNWATTYQKVPAFIDETLAAKGAENIADR
GEADASDDFEGTYEEWREHMWSDVAAYFNLDIENSEDNKSTLSLQFVDSAADMP
LAKMHGAFSTLEHHHHHH*

**DNA and amino acid sequences of P411-LAS-5266:**

ATGACAATTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATTT
ACCGTTATTAAACACAGATAAACCGGTTCAAGCTTTGATGAAAATTGCGGATG
AATTAGGAGAAATCTTTAAATTCGAGGCGCCTGGTCGTGTAACGCGCTACTTA
TCAAGTCAGCGTCTAATTAAAGAAGCATGCGATGAATCACGCTTTGATAAAGA
GTTAGTGCAAGCGCTGAAATTTCTGCGTGATTTTCTTGGAGACGGGTTAGCCA
CAAGCTGGACGTTTGAAAAAATTGGAAAAAGCGCATAATATCTTACTTCCA
AGCTTTAGTCAGCAGGCAATGAAAGGCTATCATGCGAGTATGGTCGATATCGC
CGTGCAGCTTGTTCAAAAGTGGGAGCGTCTAAATGCAGATGAGCATATTGAAG
TATCGGAAGACATGACACGTTTAACGCTTGATACAATTGGTCTTTGCGGCTTTA
ACTATCGCTTTAACAGCTTTTACCGAGATCAGCCTCATCCATTTATTATAAGTC
TGGTCCGTGCACTGGATGAAGTAATGAACAAGCTGCAGCGAGCAAATCCAGA
CGACCCAGCTTATGATGAAAACAAGCGCCAGTTTCAAGAAGATATCAAGGTG
ATGAACGACCTAGTAGATAAAATTATTGCAGATCGCAAAGCAAGTGGTGAAC
AAAGCGATGATTTATTAACGCAGATGCTAAACGGAAAAGATCCAGAAACGGG
TGAGCCGCTTGATGACCGTAACATTCGCTATCAAATTATTACATTCTTATGGGC
GGGAGTTGAAGGTACAAGTGGTCTTTTATCATTTGCGCTGTATTTCTTAGTGAA
AAATCCACATGTATTACAAAAGTAGCAGAAGAAGCAGCACGAGTTCTAGTA
GATCCTGTTCCAAGCTACAAACAAGTCAAACAGCTTAAATATGTCGGCATGGT
CTTAAACGAAGCGCTGCGCTTATGGCCACCGGTTCCGTATTTTGGTCTATATGC
AAAAGAAGATACGGTGCTTGGAGGAGAATATCCTTTAGAAAAAGGCGACGAA
GTAATGGTTCTGATTCCTCAGCTTCACCGTGATAAAACAGTTTGGGGAGACGA
TGTGGAGGAGTTCCGTCCAGAGCGTTTTGAAAATCCAAGTGCGATTCCGCAGC
ATGCGTTTAAACCGTTTGGAAACGGTCAGCGTGCGTCTCTGGGTCAGCAGTTC
GCTCTTCATGAAGCAACGCTGGTACTTGGTATGATGCTAAAACACTTTGACTTT
GAAGATCATACAAACTACGAGCTCGATATTAAAGAACGTATTACGTTAAAACC
TAAAGGCTTTGTGGTAAAAGCAAAATCGAAAAAAATTCCGCTTGGCGGTATTC
CTTCACCTAGCACTGAACAGTCTGCTAAAAAAGTACGCAAAAAGGCAGAAAA
CGCTCATAATACGCCGCTGCTTGTGCTATACGGTTCAAATATGGGTACCGCTG

AAGGAACGGCGCGTGATTTAGCAGATATTGCAATGAGCAAAGGATTTGCACC
GCAGGTCGCAACGCTTGATTCACACGCCGGAAATCTTCCGCGCGAAGGAGCTG
TATTAATTGTAACGGCGTCTTATAACGGTCATCCGCCTGATAACGCAAAGCAA
TTTGTCGACTGGTTAGACCAAGCGTCTGCTGATGAAGTAAAAGGCGTTCGCTA
CTCCGTATTTGGATGCGGCGATAAAAACTGGGCTACTACGTATCAAAAAGTGC
CTGCTTTTATCGATGAAACGCTTGCCGCTAAAGGGGCAGAAACATCGCTGAC
CGCGGTGAAGCAGATGCAAGCGACGACTTTGAAGGCACATATGAAGAATGGC
GTGAACATATGTGGAGTGACGTAGCAGCCTACTTTAACCTCGACATTGAAAAC
AGTGAAGATAATAAATCTACTCTTTCACTTCAATTTGTCGACAGCGCCGCGGA
TATGCCGCTTGCGAAAATGCACGGTGCGTTTTCAACGCTCGAGCACCACCACC
ACCACCACTGA

MTIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQ
RLIKEACDESRFDKELVQALKFLRDFLGDGLATSWTFEKNWKKAHNILLPSFSQQ
AMKGYHASMVDIAVQLVQKWERLNADEHIEVSEDMTRLTLDTIGLCGFNYRFNS
FYRDQPHPFIISLVRALDEVMNKLQRANPDDPAYDENKRQFQEDIKVMNDLVDKI
IADRKASGEQSDDLLTQMLNGKDPETGEPLDDRNIRYQIITFLWAGVEGTSGLLSF
ALYFLVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRLWPPV
PYFGLYAKEDTVLGGEYPLEKGDEVMVLIPQLHRDKTVWGDDVEEFRPERFENPS
AIPQHAFKPFGNGQRASLGQQFALHEATLVLGMMLKHFDFEDHTNYELDIKERIT
LKPKGFVVKAKSKKIPLGGIPSPSTEQSAKKVRKKAENAHNTPLLVLYGSNMGTA
EGTARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNGHPPDNAKQF
VDWLDQASADEVKGVRYSVFGCGDKNWATTYQKVPAFIDETLAAKGAENIADR
GEADASDDFEGTYEEWREHMWSDVAAYFNLDIENSEDNKSTLSLQFVDSAADMP
LAKMHGAFSTLEHHHHHH*

**DNA and amino acid sequences of P411-LAS-5259:**

ATGACAATTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATTT
ACCGTTATTAAACACAGATAAACCGGTTCAAGCTTTGATGAAAATTGCGGATG
AATTAGGAGAAATCTTTAAATTCGAGGCGCCTGGTCGTGTAACGCGCTACTTA
TCAAGTCAGCGTCTAATTAAAGAAGCATGCGATGAATCACGCTTTGATAAAGA
GTTAGTGCAAGCGCTGAAATTTCTGCGTGATTTTCTTGGAGACGGGTTAGCCA
CAAGCTGGACGTTTGAAAAAATTGGAAAAAGCGCATAATATCTTACTTCCA
AGCTTTAGTCAGCAGGCAATGAAAGGCTATCATGCGAGTATGGTCGATATCGC
CGTGCAGCTTGTTCAAAAGTGGGAGCGTCTAAATGCAGATGAGCATATTGAAG
TAGGTGAAGACATGACACGTTTAACGCTTGATACAATTGGTCTTTGCGGCTTT
AACTATCGCTTTAACAGCTTTTACCGAGATCAGCCTCATCCATTTATTATAAGT
CTGGTCCGTGCACTGGATGAAGTAATGAACAAGCTGCAGCGAGCAAATCCAG
ACGACCCAGCTTATGATGAAAACAAGCGCCAGTTTCAAGAAGATATCAAGGT
GATGAACGACCTAGTAGATAAAATTATTGCAGATCGCAAAGCAAGTGGTGAA
CAAAGCGATGATTTATTAACGCAGATGCTAAACGGAAAAGATCCAGAAACGG
GTGAGCCGCTTGATGACCGTAACATTCGCTATCAAATTATTACATTCTTATGGG
CGGGAGTTGAAGGTACAAGTGGTCTTTTATCATTTGCGCTGTATTTCTTAGTGA
AAAATCCACATGTATTACAAAAAGTAGCAGAAGAAGCAGCACGAGTTCTAGT
AGATCCTGTTCCAAGCTACAAACAAGTCAAACAGCTTAAATATGTCGGCATGG
TCTTAAACGAAGCGCTGCGCTTATGGCCACCGGTTCCGTATTTTGGTCTATATG
CAAAAGAAGATACGGTGCTTGGAGGAGAATATCCTTTAGAAAAAGGCGACGA
AGTAATGGTTCTGATTCCTCAGCTTCACCGTGATAAAACAATTTGGGGAGACG
ATGTGGAGGAGTTCCGTCCAGAGCGTTTTGAAAATCCAAGTGCGATTCCGCAG
CATGCGTTTAAACCGTTTGGAAACGGTCAGCGTGCGTCTCTGGGTCAGCAGTT
CGCTCTTCATGAAGCAACGCTGGTACTTGGTATGATGCTAAAACACTTTGACTT
TGAAGATCATACAAACTACGAGCTCGATATTAAAGAACGTATTACGTTAAAAC
CTAAAGGCTTTGTGGTAAAAGCAAAATCGAAAAAAATTCCGCTTGGCGGTATT
CCTTCACCTAGCACTGAACAGTCTGCTAAAAAAGTACGCAAAAAGGCAGAAA
ACGCTCATAATACGCCGCTGCTTGTGCTATACGGTTCAAATATGGGTACCGCT

GAAGGAACGGCGCGTGATTTAGCAGATATTGCAATGAGCAAAGGATTTGCAC
CGCAGGTCGCAACGCTTGATTCACACGCCGGAAATCTTCCGCGCGAAGGAGCT
GTATTAATTGTAACGGCGTCTTATAACGGTCATCCGCCTGATAACGCAAAGCA
ATTTGTCGACTGGTTAGACCAAGCGTCTGCTGATGAAGTAAAAGGCGTTCGCT
ACTCCGTATTTGGATGCGGCGATAAAAACTGGGCTACTACGTATCAAAAAGTG
CCTGCTTTTATCGATGAAACGCTTGCCGCTAAAGGGGCAGAAAACATCGCTGA
CCGCGGTGAAGCAGATGCAAGCGACGACTTTGAAGGCACATATGAAGAATGG
CGTGAACATATGTGGAGTGACGTAGCAGCCTACTTTAACCTCGACATTGAAAA
CAGTGAAGATAATAAATCTACTCTTTCACTTCAATTTGTCGACAGCGCCGCGG
ATATGCCGCTTGCGAAAATGCACGGTGCGTTTTCAACGCTCGAGCACCACCAC
CACCACCACTGA

MTIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQ
RLIKEACDESRFDKELVQALKFLRDFLGDGLATSWTFEKNWKKAHNILLPSFSQQ
AMKGYHASMVDIAVQLVQKWERLNADEHIEVGEDMTRLTLDTIGLCGFNYRFNS
FYRDQPHPFIISLVRALDEVMNKLQRANPDDPAYDENKRQFQEDIKVMNDLVDKI
IADRKASGEQSDDLLTQMLNGKDPETGEPLDDRNIRYQIITFLWAGVEGTSGLLSF
ALYFLVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRLWPPV
PYFGLYAKEDTVLGGEYPLEKGDEVMVLIPQLHRDKTIWGDDVEEFRPERFENPS
AIPQHAFKPFGNGQRASLGQQFALHEATLVLGMMLKHFDFEDHTNYELDIKERIT
LKPKGFVVKAKSKKIPLGGIPSPSTEQSAKKVRKKAENAHNTPLLVLYGSNMGTA
EGTARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNGHPPDNAKQF
VDWLDQASADEVKGVRYSVFGCGDKNWATTYQKVPAFIDETLAAKGAENIADR
GEADASDDFEGTYEEWREHMWSDVAAYFNLDIENSEDNKSTLSLQFVDSAADMP
LAKMHGAFSTLEHHHHHH*

**DNA and amino acid sequences of P411-LAS-5264:**

ATGACAACTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATT
TACCGTTATTAAACACAGATAAACCGGTTCAAGCTTTGATGAAAATTGCGGAT
GAATTAGGAGAAATCTTTAAATTCGAGGCGCCTGGTCGTGTAACGCGCTACTT
ATCAAGTCAGCGTCTAATTAAAGAAGCATGCGATGAATCACGCTTTGATAAAG
AGTTAGTGCAAGCGCTGAAATTTCTGCGTGAATTTCTTGGAGACGGGTTAGTG
ACAAGCTGGACGTTTGAAAAAAATTGGAAAAAGCGCATAATATCTTACTTCC
AAGCTTTAGTCAGCAGGCAATGAAAGGCTATCATGCGAGTATGGTCGATATCG
CCGTGCAGCTTGTTCAAAAGTGGGAGCGTCTAAATGCAGATGAGCATATTGAA
GTAGGTGAAGACATGACACGTTTAACGCTTGATACAATTGGTCTTTGCGGCTT
TAACTATCGCTTTAACAGCTTTTACCGAGATCAGCCTCATCCATTTATTATAAG
TCTGGTCCGTGCACTGGATGAAGTAATGAACAAGCTGCAGCGAGCGAATCCA
GACGACCCAGCTTATGATGAAAACAAGCGCCAGTTTCAAGAAGATATCAAGG
TGATGAACGACCTAGTAGATAAAATCATTGCAGATCGCAAAGCAAGTGGTGA
ACAAAGCGATGATTTATTAACGCAGATGCTAAACAGAAAAGATCCAGAAACG
GGTGAGCCGCTTGATGACCGTAACATTCGCTATCAAATTATTACATTCTTATGG
GCGGGAGTTGAAGGTACAAGTGGTCTTTTATCATTTGCGCTGTATCTCTTAGTG
AAAAATCCACATGTATTACAAAAAGTAGCAGAAGAAGCAGCACGAGTTCTAG
TAGATCCTGTTCCAAGCTACAAACAAGTCAAACAGCTTAAATATGTCGGCATG
GTCTTAAACGAAGCGCTGCGCTTATGGCCACCGATTCCGTATTTTGGTCTATAT
GCAAAGAAGATACGGTGCTTGGAGGAGAATATCCTTTAGAAAAAGGCGACG
AAGTAATGGTTCTGATTCCTCAGCTTCACCGTGATAAAACAATTTGGGGAGAC
GATGTGGAGGAGTTCCGTCCAGAGCGTTTTGAAAATCCAAGTGCGATTCCGCA
GCATGCGTTTAAACCGTTTGGAAACGGTCAGCGTGCGTCTCTGGGTCAGCAGT
TCGCTCTTCATGAAGCAACGCTGGTACTTGGTATGATGCTAAAACACTTTGACT
TTGAAGATCATACAAACTACGAGCTCGATATTAAGAACGTATTACGTTAAAA
CCTAAAGGCTTTGTGGTAAAAGCAAAATCGAAAAAATTCCGCTTGGCGGTAT
TCCTTCACCTAGCACTGAACAGTCTGCTAAAAAGTACGCAAAAGGCAGAA
AACGCTCATAATACGCCGCTGCTTGTGCTATACGGTTCAAATATGGGTACCGC

TGAAGGAACGGCGCGTGATTTAGCAGATATTGCAATGAGCAAAGGATTTGCA
CCGCAGGTCGCAACGCTTGATTCACACGCCGGAAATCTTCCGCGCGAAGGAGC
TGTATTAATTGTAACGGCGTCTTATAACGGTCATCCGCCTGATAACGCAAAGC
AATTTGTCGACTGGTTAGACCAAGCGTCTGCTGATGAAGTAAAAGGCGTTCGC
TACTCCGTATTTGGATGCGGCGATAAAAACTGGGCTACTACGTATCAAAAAGT
GCCTGCTTTTATCGATGAAACGCTTGCCGCTAAAGGGGCAGAAAACATCGCTG
ACCGCGGTGAAGCAGATGCAAGCGACGACTTTGAAGGCACATATGAAGAATG
GCGTGAACATATGTGGAGTGACGTAGCAGCCTACTTTAACCTCGACATTGAAA
ACAGTGAAGATAATAAATCTACTCTTTCACTTCAATTTGTCGACAGCGCCGCG
GATATGCCGCTTGCGAAAATGCACGGTGCGTTTTCAACGCTCGAGCACCACCA
CCACCACCACTGA

MTTKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSS
QRLIKEACDESRFDKELVQALKFLREFLGDGLVTSWTFEKNWKKAHNILLPSFSQ
QAMKGYHASMVDIAVQLVQKWERLNADEHIEVGEDMTRLTLDTIGLCGFNYRF
NSFYRDQPHPFIISLVRALDEVMNKLQRANPDDPAYDENKRQFQEDIKVMNDLVD
KIIADRKASGEQSDDLLTQMLNRKDPETGEPLDDRNIRYQIITFLWAGVEGTSGLL
SFALYLLVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRLWP
PIPYFGLYAKEDTVLGGEYPLEKGDEVMVLIPQLHRDKTIWGDDVEEFRPERFENP
SAIPQHAFKPFGNGQRASLGQQFALHEATLVLGMMLKHFDFEDHTNYELDIKERI
TLKPKGFVVKAKSKKIPLGGIPSPSTEQSAKKVRKKAENAHNTPLLVLYGSNMGT
AEGTARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNGHPPDNAKQ
FVDWLDQASADEVKGVRYSVFGCGDKNWATTYQKVPAFIDETLAAKGAENIAD
RGEADASDDFEGTYEEWREHMWSDVAAYFNLDIENSEDNKSTLSLQFVDSAADM
PLAKMHGAFSTLEHHHHHH*

**DNA and amino acid sequences of P411-LAS-5265 (P411-C10-WI)[5]:**

ATGACAATTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATTT
ACCGTTATTAAACACAGATAAACCGGTTCAAGCTTTGATGAAAATTGCGGATG
AATTAGGAGAAATCTTTAAATTCGAGGCGCCTGGTCGTGTAACGCGCTACTTA
TCAAGTCAGCGTCTAATTAAAGAAGCATGCGATGAATCACGCTTTGATAAAGA
GTTAAGTCAAGGTCTGAAATTTCTGCGTGATTTTCTTGGAGACGGGTTAGCCA
CAAGCTGGACGCATGAAAAAATTGGAAAAAGCGCATAATATCTTACTTCC
AAGCTTTAGTCAGCAGGCAATGAAAGGCTATCATGCGAGTATGGTCGATATCG
CCGTGCAGCTTGTTCAAAAGTGGGAGCGTCTAAATGCAGATGAGCATATTGAA
GTATCGGAAGACATGACACGTTTAACGCTTGATACAATTGGTCTTTGCGGCTTT
AACTATCGCCTTAACAGCTTTTACCGAGATCAGCCTCATCCATTTATTATAAGT
CTGGTCCGTGCACTGGATGAAGTAATGAACAAGCTGCAGCGAGCAAATCCAG
ACGACCCAGCTTATGATGAAAACAAGCGCCAGTTTCAAGAAGATATCAAGGT
GATGAACGACCTAGTAGATAAAATTATTGCAGATCGCAAAGCAAGGGGTGAA
CAAAGCGATGATTTATTAACGCAGATGCTAAACGGAAAAGATCCAGAAACGG
GTGAGCCGCTTGATGACGGGAACATTCGCTATCAAATTATTACATTCTTATGG
GCGGGAGTTGAAGGTACAAGTGGTCTTTTATCATTTGCGCTGTATTTCTTAGTG
AAAAATCCACATGTATTACAAAAAGTAGCAGAAGAAGCAGCACGAGTTCTAG
TAGATCCTGTTCCAAGCTACAAACAAGTCAAACAGCTTAAATATGTCGGCATG
GTCTTAAACGAAGCGCTGCGCTTATGGCCAACGGTTCCTTATTTTCCCTATAT
GCAAAAGAAGATACGGTGCTTGGAGGAGAATATCCTTTAGAAAAAGGCGACG
AAGTAATGGTTCTGATTCCTCAGCTTCACCGTGATAAAACAGTTTGGGGAGAC
GATGTGGAGGAGTTCCGTCCAGAGCGTTTTGAAAATCCAAGTGCGATTCCGCA
GCATGCGTTTAAACCGTTTGGAAACGGTCAGCGTGCGTCTCTGGGTCAGCAGT
TCGCTCTTCATGAAGCAACGCTGGTACTTGGTATGATGCTAAAACACTTTGACT
TTGAAGATCATACAAACTACGAGCTCGATATTAAAGAACTGATTACGTTAAAA
CCTAAAGGCTTTGTGGTAAAAGCAAAATCGAAAAAAATTCCGCTTGGCGGTAT
TCCTTCACCTAGCACTGAACAGTCTGCTAAAAAGTACGCAAAAGGCAGAA
AACGCTCATAATACGCCGCTGCTTGTGCTATACGGTTCAAATATGGGTACCGC

TGAAGGAACGGCGCGTGATTTAGCAGATATTGCAATGAGCAAAGGATTTGCA
CCGCAGGTCGCAACGCTTGATTCACACGCCGGAAATCTTCCGCGCGAAGGAGC
TGTATTAATTGTAACGGCGTCTTATAACGGTCATCCGCCTGATAACGCAAAGC
AATTTGTCGACTGGTTAGACCAAGCGTCTGCTGATGAAGTAAAAGGCGTTCGC
TACTCCGTATTTGGATGCGGCGATAAAAACTGGGCTACTACGTATCAAAAAGT
GCCTGCTTTTATCGATGAAACGCTTGCCGCTAAAGGGGCAGAAAACATCGCTG
ACCGCGGTGAAGCAGATGCAAGCGACGACTTTGAAGGCACATATGAAGAATG
GCGTGAACATATGTGGAGTGACGTAGCAGCCTACTTTAACCTCGACATTGAAA
ACAGTGAAGATAATAAATCTACTCTTTCACTTCAATTTGTCGACAGCGCCGCG
GATATGCCGCTTGCGAAAATGCACGGTGCGTTTTCAACGCTCGAGCACCACCA
CCACCACCACTGA

MTIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQ
RLIKEACDESRFDKELSQGLKFLRDFLGDGLATSWTHEKNWKKAHNILLPSFSQQ
AMKGYHASMVDIAVQLVQKWERLNADEHIEVSEDMTRLTLDTIGLCGFNYRLNS
FYRDQPHPFIISLVRALDEVMNKLQRANPDDPAYDENKRQFQEDIKVMNDLVDKI
IADRKARGEQSDDLLTQMLNGKDPETGEPLDDGNIRYQIITFLWAGVEGTSGLLSF
ALYFLVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRLWPTV
PYFSLYAKEDTVLGGEYPLEKGDEVMVLIPQLHRDKTVWGDDVEEFRPERFENPS
AIPQHAFKPFGNGQRASLGQQFALHEATLVLGMMLKHFDFEDHTNYELDIKELIT
LKPKGFVVKAKSKKIPLGGIPSPSTEQSAKKVRKKAENAHNTPLLVLYGSNMGTA
EGTARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNGHPPDNAKQF
VDWLDQASADEVKGVRYSVFGCGDKNWATTYQKVPAFIDETLAAKGAENIADR
GEADASDDFEGTYEEWREHMWSDVAAYFNLDIENSEDNKSTLSLQFVDSAADMP
LAKMHGAFSTLEHHHHHH*

**DNA and amino acid sequences of P411-LAS-5266:**

ATGACAATTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATTT
ACCGTTATTAAACACAGATAAACCGGTTCAAGCTTTGATGAAAATTGCGGATG
AATTAGGAGAAATCTTTAAATTCGAGGCGCCTGGTCGTGTAACGCGCTACTTA
TCAAGTCAGCGTCTAATTAAAGAAGCATGCGATGAATCACGCTTTGATAAAGA
GTTAAGTCAAGGTCTGAAATTTCTGCGTGATTTTCTTGGAGACGGGTTACCGA
CAAGCTGGACGCATGAAAAAATTGGAAAAAAGCGCATAATATCTTACTTCC
AAGCTTTAGTCAGCAGGCAATGAAAGGCTATCATGCGAGTATGGTCGATATCG
CCGTGCAGCTTGTTCAAAAGTGGGAGCGTCTAAATGCAGATGAGCATATTGAA
GTATCGGAAGACATGACACGTTTAACGCTTGATACAATTGGTCTTTGCGGCTTT
AACTATCGCCTTAACAGCTTTTACCGAGATCAGCCTCATCCATTTATTATAAGT
CTGGTCCGTGCACTGGATGAAGTAATGAACAAGCTGCAGCGAGCAAATCCAG
ACGACCCAGCTTATGATGAAAACAAGCGCCAGTTTCAAGAAGATATCAAGGT
GATGAACGACCTAGTAGATAAAATTATTGCAGATCGCAAAGCAAGGGGTGAA
CAAAGCGATGATTTATTAACGCAGATGCTAAACGGAAAAGATCCAGAAACGG
GTGAGCCGCTTGATGACGGGAACATTCGCTATCAAATTATTACATTCTTATATA
GTGGAGTTGATGGTACAAGTGGTCTTTATCATTTGCGCTGTATTTCTTAGTGA
AAAATCCACATGTATTACAAAAAGTAGCAGAAGAAGCAGCACGAGTTCTAGT
AGATCCTGTTCCAAGCTACAAACAAGTCAAACAGCTTAAATATGTCGGCATGG
TCTTAAACGAAGCGCTGCGCTTATGGCCACCGGTTCCTTATTTTGCGCTATATG
CAAAAGAAGATACGGTGCTTGGAGGAGAATATCCTTTAGAAAAAGGCGACGA
AGTAATGGTTCTGATTCCTCAGCTTCACCGTGATAAAACAGTTTGGGGAGACG
ATGTGGAGGAGTTCCGTCCAGAGCGTTTTGAAAATCCAAGTGCGATTCCGCAG
CATGCGTTTAAACCGTTTGGAAACGGTCAGCGTGCGTCTCTGGGTCAGCAGTT
CGCTCTTCATGAAGCAACGCTGGTACTTGGTATGATGCTAAAACACTTTGACTT
TGAAGATCATACAAACTACGAGCTCGATATTAAAGAACTGCTTACGTTAAAAC
CTAAAGGCTTTGTGGTAAAAGCAAAATCGAAAAAAATTCCGCTTGGCGGTATT
CCTTCACCTAGCACTGAACAGTCTGCTAAAAAAGTACGCAAAAAGGCAGAAA
ACGCTCATAATACGCCGCTGCTTGTGCTATACGGTTCAAATATGGGTACCGCT

GAAGGAACGGCGCGTGATTTAGCAGATATTGCAATGAGCAAAGGATTTGCAC
CGCAGGTCGCAACGCTTGATTCACACGCCGGAAATCTTCCGCGCGAAGGAGCT
GTATTAATTGTAACGGCGTCTTATAACGGTCATCCGCCTGATAACGCAAAGCA
ATTTGTCGACTGGTTAGACCAAGCGTCTGCTGATGAAGTAAAAGGCGTTCGCT
ACTCCGTATTTGGATGCGGCGATAAAAACTGGGCTACTACGTATCAAAAAGTG
CCTGCTTTTATCGATGAAACGCTTGCCGCTAAAGGGGCAGAAAACATCGCTGA
CCGCGGTGAAGCAGATGCAAGCGACGACTTTGAAGGCACATATGAAGAATGG
CGTGAACATATGTGGAGTGACGTAGCAGCCTACTTTAACCTCGACATTGAAAA
CAGTGAAGATAATAAATCTACTCTTTCACTTCAATTTGTCGACAGCGCCGCGG
ATATGCCGCTTGCGAAAATGCACGGTGCGTTTTCAACGCTCGAGCACCACCAC
CACCACCACTGA

MTIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQ
RLIKEACDESRFDKELSQGLKFLRDFLGDGLPTSWTHEKNWKKAHNILLPSFSQQ
AMKGYHASMVDIAVQLVQKWERLNADEHIEVSEDMTRLTLDTIGLCGFNYRLNS
FYRDQPHPFIISLVRALDEVMNKLQRANPDDPAYDENKRQFQEDIKVMNDLVDKI
IADRKARGEQSDDLLTQMLNGKDPETGEPLDDGNIRYQIITFLYSGVDGTSGLLSF
ALYFLVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRLWPPV
PYFALYAKEDTVLGGEYPLEKGDEVMVLIPQLHRDKTVWGDDVEEFRPERFENPS
AIPQHAFKPFGNGQRASLGQQFALHEATLVLGMMLKHFDFEDHTNYELDIKELLT
LKPKGFVVKAKSKKIPLGGIPSPSTEQSAKKVRKKAENAHNTPLLVLYGSNMGTA
EGTARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNGHPPDNAKQF
VDWLDQASADEVKGVRYSVFGCGDKNWATTYQKVPAFIDETLAAKGAENIADR
GEADASDDFEGTYEEWREHMWSDVAAYFNLDIENSEDNKSTLSLQFVDSAADMP
LAKMHGAFSTLEHHHHHH*

**Synthesis of Substrates**

2-(2-Tosylhydrazineylidene)acetyl chloride
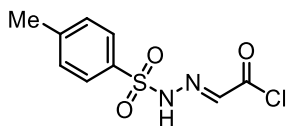


4-Methylbenzenesulfonohydrazide (55.9 g, 0.3 mol, 1.0 equiv.) was dissolved in aqueous hydrochloric acid (2 M, 180 mL) and warmed to 50 °C (solution 1). 2-Oxoacetic acid (44.4 g of 50% in water, 0.3 mol, 1.0 equiv.) was dissolved in water (300 mL) and heated to 50 °C (solution 2). Pre-warmed solution 1 was slowly transferred to solution 2. The reaction mixture was then stirred at 60 °C for 4 h until all the hydrozone product had crashed out. The mixture was cooled to 0 °C and kept for 2 h. The product 2-(2-tosylhydrazineylidene) acetic acid (~70 g, 97% yield) was collected by filtration, washed with hexane: ether (10:1, 20 mL × 3) and dried under vaccum.

2-(2-Tosylhydrazineylidene)acetic acid (70 g, 0.29 mmol, 1.0 equiv.) was dissolved in dry dichloromethane (300 mL). Thionyl chloride (50 mL) and N,N-dimethyl formaldehyde (4 drops, cat.) were added to the solution. The reaction mixture was stirred at room temperature for 1 h and then heated to reflux (~ 50 °C) for 5 h until the starting material was completely dissolved and the reaction turned clear and light yellow. After the reaction was cooled to room temperature, organic solvent and the excess thionyl chloride were removed under reduced pressure. The resulting mixture was treated with ether (20 mL) and sonicated for 5 min. Hexane (150 mL) was then slowly added to the mixture to completely crash out the acyl chloride product. The mixture was cooled to 0 °C and kept for 2 h. The product 2-(2-

tosylhydrazineylidene)acetyl chloride (~74 g, 98% yield, pale yellow) was collected by filtration, washed with hexane (20 mL × 2) and dried under vacuum.

Diazo Synthesis
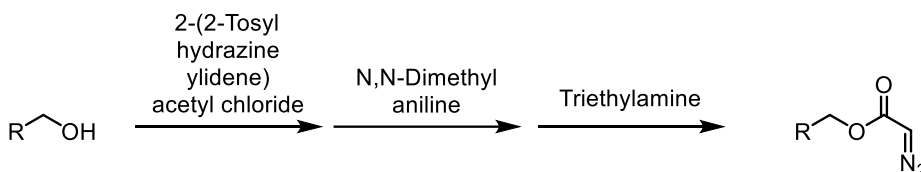


An alcohol substrate (1.0 equiv.) was dissolved in dry dichloromethane (conc. ~ 0.2–0.5 M) and kept at 0 °C. 2-(2-Tosylhydrazineylidene)acetyl chloride (1.05 equiv.) was then added to the solution. The mixture was stirred at 0 °C for 10 min before the addition of N,N-dimethyl aniline (1.3 equiv.). The resulting mixture was then stirred for another 10 min. Triethylamine (2.0 equiv.) was added to the reaction, which was then allowed to slowly warm up to room temperature over 20 min. The reaction was concentrated under reduced pressure and quenched by citric acid (saturated aqueous solution). The resulting mixture was transferred to a separatory funnel. Dichloromethane and water were used in minimum amount to wash the reaction container and transfer everything to the separatory funnel. Hexane/ethyl acetate (13:1) was used for extraction for three times. The combined organic layer was then washed with saturated citric acid solution and brine, dried over $Na_2SO_4$, filtered, and concentrated under reduced pressure. The resulting crude product was purified through a silica column using pentane/ether (1:0 to 10:1) as eluents. The yellow-colored fractions were concentrated to afford the diazo product as a yellow liquid (80–96% yield).

**4-methoxyphenethyl 2-diazoacetate (1c)**
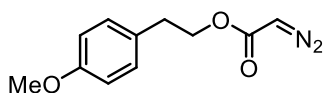


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.13 (d, *J* = 8.6 Hz, 2H), 6.85 (d, *J* = 8.6 Hz, 2H), 4.72 (s, 1H), 4.33 (t, *J* = 7.0 Hz, 2H), 3.79 (s, 3H), 2.89 (t, *J* = 7.0 Hz, 2H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 158.46 , 130.02 , 129.78 , 114.06 , 65.67 , 55.40 , 34.56 .

**HRMS** (FI): calcd for $C_{11}H_{12}N_2O_3$ [M]$^{+•}$ 220.08479; found 220.08468.

**4-fluorophenethyl 2-diazoacetate (1d)**
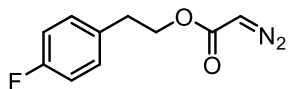


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.21 – 7.12 (m, 2H), 7.04 – 6.94 (m, 2H), 4.72 (s, 1H), 4.34 (t, *J* = 6.9 Hz, 2H), 2.92 (t, *J* = 6.9 Hz, 2H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 161.86 (d, *J* = 244.6 Hz), 133.45 (d, *J* = 3.3 Hz), 130.47 (d, *J* = 8.0 Hz), 115.46 (d, *J* = 21.3 Hz), 65.30 , 34.64 .

**HRMS** (FD): calcd for $C_{10}H_9N_2O_2F$ [M]$^{+•}$ 208.06481; found 208.06525.

**2-diazo-*N*-phenethylacetamide (1g)**
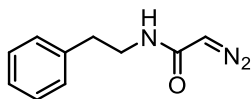


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.36 – 7.21 (m, 5H), 5.11 (s, 1H), 4.68 (s, 1H), 3.65 – 3.53 (m, 2H), 2.85 (t, *J* = 6.9 Hz, 2H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 165.54 , 138.85 , 128.91 , 126.70 , 47.27 , 41.22 , 36.15 .

**HRMS** (FI): calcd for $C_7H_{13}N_2O_4$ [M]$^{+\bullet}$ 189.08753; found 189.09037.

**3-(*p*-tolyl)propyl 2-diazoacetate (3a)**
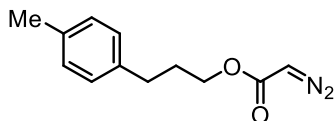


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.14 – 7.03 (m, 4H), 4.74 (s, 1H), 4.18 (t, *J* = 6.5 Hz, 2H), 2.65 (dd, *J* = 8.6, 6.7 Hz, 2H), 2.32 (s, 3H), 2.03 – 1.89 (m, 2H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 138.13 , 135.62 , 129.26 , 128.40 , 64.35 , 31.73 , 30.63 , 21.13 .

**HRMS** (FI): calcd for $C_{12}H_{14}N_2O_2$ [M]$^{+\bullet}$ 218.10553; found 218.10529.

**3-(4-methoxyphenyl)propyl 2-diazoacetate (3c)**
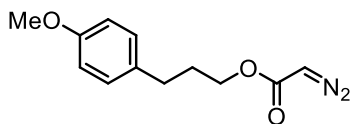


**$^1$H NMR** (400 MHz, Chloroform-*d*) δ 7.17 – 6.98 (m, 2H), 6.88 – 6.79 (m, 2H), 4.74 (s, 1H), 4.17 (t, *J* = 6.5 Hz, 2H), 3.79 (s, 3H), 2.63 (dd, *J* = 8.5, 6.8 Hz, 2H), 2.00 – 1.88 (m, 2H).

**$^{13}$C NMR** (101 MHz, Chloroform-*d*) δ 158.05 , 133.27 , 129.43 , 114.00 , 64.32 , 55.41 , 31.27 , 30.74 .

**HRMS** (FI): calcd for $C_{12}H_{14}N_2O_3$ [M]$^{+\bullet}$ 234.10044; found 234.10034.

**3-(4-fluorophenyl)propyl 2-diazoacetate (3d)**
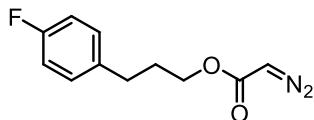


**$^1$H NMR** (400 MHz, Chloroform-*d*) δ 7.17 – 7.09 (m, 2H), 7.02 – 6.92 (m, 2H), 4.74 (s, 1H), 4.17 (t, *J* = 6.5 Hz, 2H), 2.66 (dd, *J* = 8.5, 6.8 Hz, 2H), 2.01 – 1.89 (m, 2H).

**$^{13}$C NMR** (101 MHz, Chloroform-*d*) δ 161.49 (d, *J* = 243.6 Hz), 136.81 (d, *J* = 3.2 Hz), 129.85 (d, *J* = 7.8 Hz), 115.33 (d, *J* = 21.2 Hz), 64.13 , 31.42 , 30.63 .

**HRMS** (FD): calcd for $C_{11}H_{11}N_2O_2F$ [M]$^{+\bullet}$ 222.08046; found 222.07990.

**3-(furan-2-yl)propyl 2-diazoacetate (3e)**
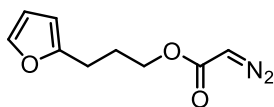


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.38 – 7.29 (m, 1H), 6.28 (dd, *J* = 3.2, 1.9 Hz, 1H), 6.01 (dt, *J* = 2.8, 0.9 Hz, 1H), 4.74 (s, 1H), 4.20 (t, *J* = 6.4 Hz, 2H), 2.76 – 2.69 (m, 2H), 2.00 (dq, *J* = 7.4, 6.4 Hz, 2H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 154.89 , 141.21 , 110.29 , 105.40 , 64.15 , 27.42 , 24.57 .

**HRMS** (ESI): calcd for $C_9H_{10}N_2O_3Na$ [M+Na]$^{+•}$ 217.0613; found 217.0596.

**3-(thiophen-2-yl)propyl 2-diazoacetate (3f)**
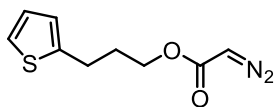


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.13 (dd, *J* = 5.1, 1.2 Hz, 1H), 6.92 (dd, *J* = 5.1, 3.4 Hz, 1H), 6.80 (dq, *J* = 3.3, 1.1 Hz, 1H), 4.75 (s, 1H), 4.22 (t, *J* = 6.4 Hz, 2H), 2.96 – 2.87 (m, 2H), 2.03 (dq, *J* = 7.5, 6.4 Hz, 2H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 143.86 , 126.96 , 124.65 , 123.42 , 63.93 , 30.85 , 26.34 .

**HRMS** (FI): calcd for $C_9H_{10}N_2O_2S$ [M]$^{+•}$ 210.04630; found 210.04534.

**4-(4-methoxyphenyl)butyl 2-diazoacetate (5a)**
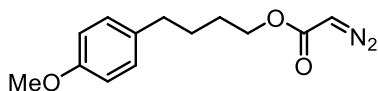


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.12 – 7.05 (m, 2H), 6.87 – 6.79 (m, 2H), 4.72 (s, 1H), 4.21 – 4.13 (m, 2H), 3.79 (s, 3H), 2.63 – 2.54 (m, 2H), 1.66 (d, *J* = 6.9 Hz, 4H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 157.91 , 134.19 , 129.39 , 113.89 , 64.90 , 55.39 , 34.62 , 28.45 , 27.95 .

**HRMS** (FI): calcd for $C_{13}H_{16}N_2O_3$ [M]⁺• 248.11609; found 248.11580.

**(*E*)-6-(4-methoxyphenyl)hex-5-en-1-yl 2-diazoacetate (5b)**
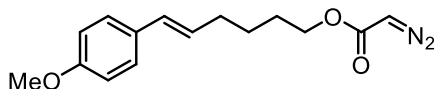


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.27 (d, *J* = 9.0 Hz, 2H), 6.84 (d, *J* = 8.7 Hz, 2H), 6.33 (d, *J* = 15.8 Hz, 1H), 6.11 – 5.96 (m, 1H), 4.73 (s, 1H), 4.18 (t, *J* = 6.6 Hz, 2H), 3.80 (s, 3H), 2.22 (q, *J* = 6.9 Hz, 2H), 1.70 (dt, *J* = 15.3, 6.8 Hz, 2H), 1.58 – 1.49 (m, 2H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 158.85 , 130.64 , 129.82 , 128.13 , 127.15 , 114.05 , 64.94 , 55.43 , 32.64 , 28.45 , 25.83 .

**HRMS** (FI): calcd for $C_{15}H_{18}N_2O_3$ [M]⁺• 274.13174; found 274.13224.

**4-(1*H*-indol-3-yl)butyl 2-diazoacetate (5c)**
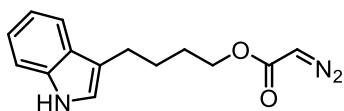


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.87 (s, 1H), 7.52 (dd, *J* = 8.0, 1.1 Hz, 1H), 7.28 (dd, *J* = 8.1, 0.9 Hz, 1H), 7.21 – 7.08 (m, 1H), 7.04 (ddd, *J* = 8.1, 7.0, 1.1 Hz, 1H), 6.93 – 6.87 (m, 1H), 4.64 (s, 1H), 4.13 (t, *J* = 6.2 Hz, 2H), 2.77 – 2.68 (m, 2H), 1.76 – 1.62 (m, 4H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 136.48 , 127.58 , 122.04 , 121.34 , 119.27 , 119.00 , 116.37 , 111.21 , 65.01 , 28.75 , 26.45 , 24.84 .

**HRMS (FI):** calcd for $C_{14}H_{15}N_3O_2$ [M]$^{+\bullet}$ 257.11643; found 257.11652.

**2,3-dihydro-1*H*-inden-2-yl 2-diazoacetate (7a)**
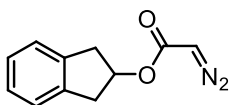


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.31 – 7.22 (m, 4H), 5.65 (dt, *J* = 6.5, 3.3 Hz, 1H), 4.75 (s, 1H), 3.36 (dd, *J* = 17.0, 6.5 Hz, 2H), 3.07 (dd, *J* = 17.0, 3.1 Hz, 2H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 140.31 , 126.80 , 124.65 , 75.91 , 39.70 .

**HRMS (FI):** calcd for $C_{11}H_{10}N_2O_3$ [M]$^{+\bullet}$ 202.07423; found 202.07387.

**(2,3-dihydro-1*H*-inden-2-yl)methyl 2-diazoacetate (7b)**
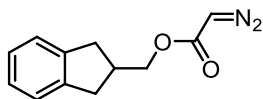


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.31 – 7.21 (m, 2H), 7.17 (dt, *J* = 5.2, 3.6 Hz, 2H), 4.78 (s, 1H), 4.23 (d, *J* = 6.8 Hz, 2H), 3.16 – 3.05 (m, 2H), 2.92 – 2.72 (m, 3H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 142.25 , 126.41 , 124.61 , 67.94 , 38.40 , 35.89 .

**HRMS** (FI): calcd for $C_{12}H_{12}N_2O_2$ [M]$^{+\bullet}$ 216.08988; found 216.08981.

**2-(1,2,3,4-tetrahydronaphthalen-1-yl)ethyl 2-diazoacetate (7c)**
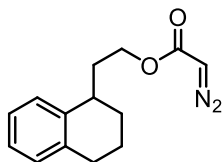


**¹H NMR** δ 7.17 – 7.04 (m, 4H), 4.76 (s, 1H), 4.29 (ddd, *J* = 7.2, 6.3, 3.7 Hz, 2H), 2.92 (dd, *J* = 9.4, 4.9 Hz, 1H), 2.81 – 2.72 (m, 2H), 2.13 – 2.02 (m, 1H), 1.95 – 1.81 (m, 3H), 1.73 (dtd, *J* = 10.0, 5.9, 3.8 Hz, 2H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 140.16 , 137.18 , 129.36 , 128.64 , 125.90 , 125.78 , 63.38 , 35.78 , 34.50 , 29.63 , 27.68 , 19.72 .

**HRMS** (FI): calcd for $C_{14}H_{16}N_2O_2$ [M]$^{+\bullet}$ 244.12118; found 244.12117.

**2-(7-methoxyisochroman-1-yl)ethyl 2-diazoacetate (7d)**
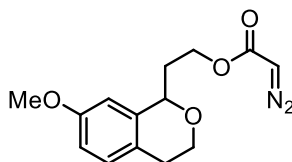


**$^1$H NMR** (400 MHz, Chloroform-*d*) δ 7.03 (d, *J* = 8.4 Hz, 1H), 6.74 (dd, *J* = 8.4, 2.6 Hz, 1H), 6.59 (d, *J* = 2.6 Hz, 1H), 4.81 (dd, *J* = 8.9, 2.9 Hz, 1H), 4.72 (s, 1H), 4.36 (ddd, *J* = 7.4, 5.9, 4.0 Hz, 2H), 3.82 – 3.67 (m, 4H), 2.96 – 2.83 (m, 1H), 2.63 (dt, *J* = 15.9, 3.8 Hz, 1H), 2.26 (ddd, *J* = 15.0, 7.3, 3.2 Hz, 1H), 2.13 (ddd, *J* = 9.0, 6.6, 5.3 Hz, 1H).

**$^{13}$C NMR** (101 MHz, Chloroform-*d*) δ 158.07 , 138.55 , 130.02 , 126.08 , 112.59 , 109.91 , 72.84 , 63.51 , 61.87 , 55.46 , 35.07 , 28.31 .

**HRMS (FI):** calcd for $C_{14}H_{16}N_2O_4$ [M]$^{+\bullet}$ 276.1101; found 276.11068.

**1,2,3,4-tetrahydronaphthalen-2-yl 2-diazoacetate (7e)**
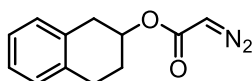


**$^1$H NMR** (400 MHz, Chloroform-*d*) δ 7.17 – 7.03 (m, 4H), 5.30 (dddd, *J* = 8.4, 7.0, 5.1, 3.2 Hz, 1H), 4.74 (s, 1H), 3.14 (dd, *J* = 16.7, 5.2 Hz, 1H), 2.99 – 2.80 (m, 3H), 2.12 – 1.92 (m, 2H).

**$^{13}$C NMR** (101 MHz, Chloroform-*d*) δ 135.92 , 133.97 , 129.77 , 129.04 , 126.51 , 126.40 , 70.83 , 35.15 , 28.44 , 26.81 .

**HRMS (FI):** calcd for $C_{12}H_{12}N_2O_2$ [M]$^{+\bullet}$ 216.08988; found 216.08980.

**2-(2,3-dihydro-1H-inden-1-yl)ethyl 2-diazoacetate (7f)**
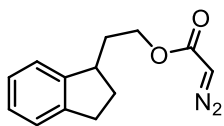


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.20 – 7.05 (m, 4H), 4.67 (s, 1H), 4.23 (t, *J* = 6.8 Hz, 2H), 3.17 – 3.08 (m, 1H), 2.91 – 2.71 (m, 2H), 2.31 – 2.07 (m, 2H), 1.75 – 1.53 (m, 2H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 146.48 , 143.96 , 126.66 , 126.28 , 124.65 , 123.59 , 63.71 , 41.73 , 33.98 , 32.27 , 31.51 .

**HRMS (**FI**): calcd for C₁₃H₁₄N₂O₂ [M]⁺• 230.10553; found 230.10561.**

**Synthesis of Reference Products**

Compounds **2a**–**2g** and **4b** were purchased from commercial suppliers including Sigma-Aldrich, VWR, TCI America, Fischer Scientific, Alfa Aesar, Acros, and Combi Blocks and were used without additional purification. Other reference products were obtained using the method described below or through enzymatic preparative synthesis.
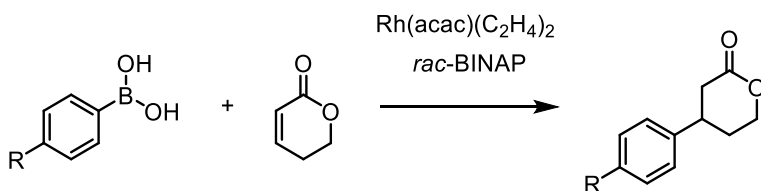


A typical procedure is given for the preparation of 6-membered lactones: 1,4-Dioxane (2.0 mL) was added to a flask charged with Rh(acac)(C$_2$H$_4$)$_2$ (12 μmol), rac-BINAP (14 μmol), and phenylboronic acid (2.00 mmol) and flushed with nitrogen, which was followed by addition of water (0.2 mL) and 5,6-dihydro-2H-pyran-2-one (0.40 mmol). The resulting mixture was then stirred at 100 °C for 3 h. After evaporation of the solvent, the residue was dissolved in ethyl acetate. The solution was washed with saturated aqueous sodium bicarbonate and dried over anhydrous Na$_2$SO$_4$. Chromatography on silica gel (hexane:EtOAc 2:1 to 1:1) gave products as a colorless oil.

**4-(*p*-Tolyl)tetrahydro-2*H*-pyran-2-one (4a)**
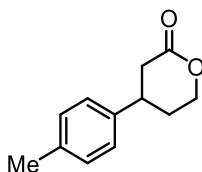
**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.40 – 7.35 (m, 2H), 7.33 – 7.28 (m, 1H), 7.25 – 7.21 (m, 2H), 4.67 (dd, *J* = 9.1, 7.8 Hz, 1H), 4.28 (dd, *J* = 9.1, 7.9 Hz, 1H), 3.79 (p, *J* = 8.4 Hz, 1H), 2.93 (dd, *J* = 17.5, 8.7 Hz, 1H), 2.68 (dd, *J* = 17.5, 9.1 Hz, 1H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 176.51 , 139.54 , 129.30 , 127.88 , 126.84 , 41.27 , 35.86 .
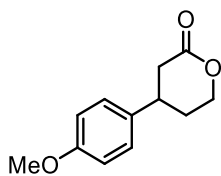
Spectral data were in accordance with literature values.[7]

**4-Phenyltetrahydro-2*H*-pyran-2-one (4b)**
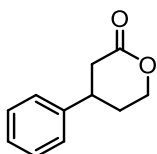
**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.41 – 7.32 (m, 2H), 7.31 – 7.25 (m, 2H), 7.24 – 7.18 (m, 2H), 4.51 (ddd, *J* = 11.4, 4.9, 3.9 Hz, 1H), 4.40 (ddd, *J* = 11.4, 10.4, 3.8 Hz, 1H), 3.24 (tdd, *J* = 10.5, 5.9, 4.6 Hz, 1H), 2.93 (ddd, *J* = 17.7, 6.0, 1.7 Hz, 1H), 2.64 (dd, *J* = 17.6, 10.6 Hz, 1H), 2.19 (dqd, *J* = 14.0, 3.9, 1.6 Hz, 1H), 2.11 – 1.99 (m, 1H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 170.78 , 142.89 , 129.13 , 127.38 , 126.59 , 68.78 , 37.65 , 37.60 , 30.46 .

Spectral data were in accordance with literature values.[8]

**4-(4-Methoxyphenyl)tetrahydro-2*H*-pyran-2-one (4c)**



**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.17 – 7.09 (m, 2H), 6.94 – 6.85 (m, 2H), 4.50 (ddd, *J* = 11.4, 4.9, 3.9 Hz, 1H), 4.38 (ddd, *J* = 11.4, 10.4, 3.8 Hz, 1H), 3.80 (s, 3H), 3.19 (tdd, *J* = 10.5, 5.9, 4.5 Hz, 1H), 2.90 (ddd, *J* = 17.6, 5.9, 1.7 Hz, 1H), 2.59 (dd, *J* = 17.7, 10.6 Hz, 1H), 2.15 (ddtd, *J* = 13.9, 4.5, 3.9, 1.7 Hz, 1H), 2.00 (dtd, *J* = 14.1, 10.5, 4.9 Hz, 1H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 170.90 , 158.79 , 134.99 , 127.56 , 114.45 , 68.80 , 55.46 , 37.91 , 36.81 , 30.67 .

Spectral data were in accordance with literature values.[9]

**4-(4-Fluorophenyl)tetrahydro-2*H*-pyran-2-one (4d)**
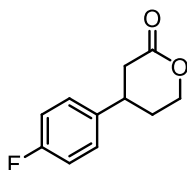


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.22 – 7.13 (m, 2H), 7.10 – 7.00 (m, 2H), 4.51 (ddd, *J* = 11.5, 4.9, 3.8 Hz, 1H), 4.39 (ddd, *J* = 11.5, 10.5, 3.7 Hz, 1H), 3.23 (tdd, *J* = 10.6, 5.9, 4.5 Hz, 1H), 2.91 (ddd, *J* = 17.6, 6.0, 1.7 Hz, 1H), 2.59 (dd, *J* = 17.6, 10.6 Hz, 1H), 2.21 – 2.13 (m, 1H), 2.01 (dtd, *J* = 14.1, 10.6, 4.9 Hz, 1H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 170.51 , 163.21 , 160.76 , 138.60 , 128.13 , 128.05 , 116.09 , 115.88 , 68.68 , 37.82 , 36.98 , 30.56 .

**¹⁹F NMR** (376 MHz, Chloroform-*d*) δ -115.32 .

Spectral data were in accordance with literature values. [9]

**4-(Furan-2-yl)tetrahydro-2*H*-pyran-2-one (4e)**
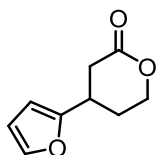


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.36 (dd, *J* = 2.0, 0.8 Hz, 1H), 6.32 (dd, *J* = 3.2, 1.8 Hz, 1H), 6.09 (dt, *J* = 3.2, 0.9 Hz, 1H), 4.47 – 4.33 (m, 2H), 3.37 (dddd, *J* = 11.6, 6.1, 3.6,

1.0 Hz, 1H), 2.93 (ddd, *J* = 17.5, 6.2, 1.2 Hz, 1H), 2.72 (dd, *J* = 17.5, 9.2 Hz, 1H), 2.28 –

2.20 (m, 1H), 2.04 (dtd, *J* = 13.8, 8.8, 4.7 Hz, 1H).

**13C NMR** (101 MHz, Chloroform-*d*) δ 170.13 , 155.53 , 142.13 , 110.38 , 105.20 , 68.08 ,

34.60 , 31.14 , 27.99 .[10]

**4-(Thiophen-2-yl)tetrahydro-2*H*-pyran-2-one (4f)**
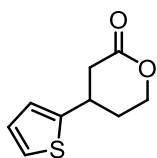


**1H NMR** (400 MHz, Chloroform-*d*) δ 7.22 (dd, *J* = 5.2, 1.2 Hz, 1H), 6.98 (dd, *J* = 5.1, 3.5

Hz, 1H), 6.90 – 6.84 (m, 1H), 4.50 (dt, *J* = 11.6, 4.6 Hz, 1H), 4.39 (ddd, *J* = 11.7, 10.1, 3.8

Hz, 1H), 3.58 – 3.49 (m, 1H), 3.03 (ddd, *J* = 17.7, 5.9, 1.6 Hz, 1H), 2.70 (dd, *J* = 17.6, 10.3

Hz, 1H), 2.29 (ddq, *J* = 11.0, 4.2, 2.6, 2.1 Hz, 1H), 2.06 (dtd, *J* = 14.5, 10.1, 4.8 Hz, 1H).

**13C NMR** (101 MHz, Chloroform-*d*) δ 169.89 , 146.43 , 127.18 , 124.10 , 123.48 , 68.39 ,

38.24 , 33.19 , 31.46 .[11]

**4-(4-Methoxyphenyl)oxepan-2-one (6a)**
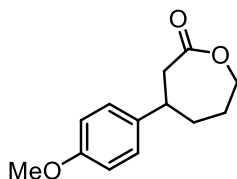


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.13 – 7.07 (m, 2H), 6.85 (d, *J* = 8.7 Hz, 2H), 4.42 – 4.20 (m, 2H), 3.79 (s, 3H), 3.05 – 2.71 (m, 3H), 2.21 – 2.01 (m, 2H), 1.99 – 1.70 (m, 2H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 174.76 , 158.47 , 137.84 , 128.01 , 127.35 , 114.39 , 114.24 , 69.23 , 55.43 , 41.96 , 40.00 , 38.09 , 29.17 .

Spectral data were in accordance with literature values.[12]

**(*E*)-4-(4-Methoxystyryl)oxepan-2-one (6b)**
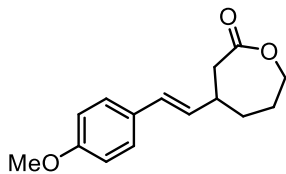


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.27 (d, *J* = 8.5 Hz, 3H), 6.88 – 6.82 (m, 2H), 6.40 (dd, *J* = 15.7, 1.1 Hz, 1H), 5.99 (dd, *J* = 15.9, 7.4 Hz, 1H), 4.38 – 4.27 (m, 1H), 4.23 (ddd, *J* = 12.7, 9.3, 1.1 Hz, 1H), 3.80 (s, 3H), 2.84 – 2.69 (m, 2H), 2.67 – 2.55 (m, 1H), 2.13 – 1.99 (m, 2H), 1.95 – 1.80 (m, 1H), 1.73 – 1.60 (m, 1H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 174.69 , 159.29, 130.35 , 129.76 , 129.34 , 127.48 , 114.14 , 69.24 , 55.45 , 40.59 , 37.52 , 35.81 , 28.16 .

**HRMS** (FI): calcd for $C_{15}H_{18}O_3$ [M]$^{+\bullet}$ 246.12559; found 246.12724.

**4-(1*H*-Indol-3-yl)oxepan-2-one (6c)**
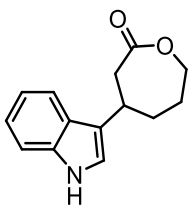


**¹H NMR** (400 MHz, Chloroform-*d*) δ 8.03 (s, 1H), 7.65 (dt, *J* = 7.8, 1.0 Hz, 1H), 7.38 (dt, *J* = 8.1, 0.9 Hz, 1H), 7.22 (ddd, *J* = 8.2, 7.1, 1.2 Hz, 1H), 7.14 (ddd, *J* = 8.0, 7.1, 1.1 Hz, 1H), 7.05 – 7.01 (m, 1H), 4.44 – 4.29 (m, 2H), 3.33 (ddd, *J* = 11.6, 7.8, 4.5 Hz, 1H), 3.09 – 2.99 (m, 2H), 2.40 – 2.27 (m, 1H), 2.19 – 2.06 (m, 1H), 2.05 – 1.95 (m, 2H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 175.16 , 136.48 , 125.98 , 122.56 , 120.25 , 119.97 , 119.79 , 119.04 , 111.47 , 69.41 , 41.51 , 36.13 , 32.24 , 28.93 .

**HRMS** (FI): calcd for $C_{14}H_{15}NO_2$ [M]⁺• 229.11028; found 229.11145.

**3,3a,8,8a-Tetrahydro-2*H*-indeno[2,1-*b*]furan-2-one (6b)**
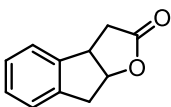


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.30 – 7.20 (m, 5H), 5.30 (ddd, *J* = 5.8, 3.8, 2.7 Hz, 1H), 4.02 (ddt, *J* = 9.3, 5.7, 1.3 Hz, 1H), 3.35 – 3.29 (m, 2H), 3.05 (dd, *J* = 17.8, 9.3 Hz, 1H), 2.75 (dd, *J* = 17.8, 1.5 Hz, 1H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 142.15 , 128.46 , 127.85 , 125.43 , 124.79 , 84.43 , 45.58 , 39.07 , 35.48 .

Spectral data were in accordance with literature values.[13]

**4,4a,9,9a-Tetrahydroindeno[2,1-*c*]pyran-3(1*H*)-one (6c)**
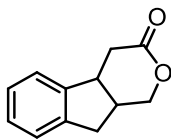


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.26 – 7.16 (m, 4H), 4.39 (dd, *J* = 11.4, 5.0 Hz, 1H), 4.09 (dd, *J* = 11.4, 8.8 Hz, 1H), 3.81 (dt, *J* = 9.8, 7.8 Hz, 1H), 3.30 (dd, *J* = 16.8, 9.7 Hz, 1H), 3.10 – 2.97 (m, 2H), 2.81 (dd, *J* = 16.8, 3.9 Hz, 1H), 2.58 (dd, *J* = 15.2, 7.8 Hz, 1H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 172.79 , 143.96 , 141.31 , 127.91 , 127.56 , 125.06 , 124.27 , 70.26 , 41.04 , 35.81 , 34.93 , 34.90 .

**HRMS** (FI): calcd for $C_{12}H_{12}O_2$ [M]$^{+•}$ 188.08373; found 188.08404.

**3,4,5',6'-Tetrahydro-2H-spiro[naphthalene-1,4'-pyran]-2'(3'*H*)-one (8c)**
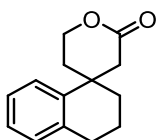


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.24 – 7.06 (m, 4H), 4.50 – 4.39 (m, 2H), 2.87 – 2.76 (m, 3H), 2.65 (dd, *J* = 16.7, 1.1 Hz, 1H), 2.26 (ddd, *J* = 14.2, 8.1, 5.8 Hz, 1H), 1.95 (dddd, *J* = 14.5, 5.3, 4.4, 1.1 Hz, 1H), 1.91 – 1.77 (m, 4H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 172.01 , 142.44 , 136.63 , 126.81 , 126.68 , 77.48 , 66.40 , 37.20 , 36.23 , 18.98 .

**HRMS** (FI): calcd for $C_{12}H_{16}O$ [M]$^{+•}$ 176.12012; found 176.12080.

**5-Methoxy-2,3,5',6'-tetrahydrospiro[indene-1,4'-pyran]-2'(3'*H*)-one (8d)**
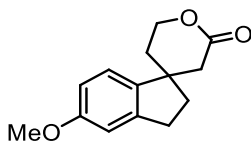


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.16 – 6.99 (m, 1H), 6.77 (ddd, *J* = 8.3, 2.7, 0.7 Hz, 1H), 6.58 (d, *J* = 2.7 Hz, 1H), 5.20 – 5.03 (m, 2H), 4.15 (dd, *J* = 11.2, 1.4 Hz, 1H), 3.92 – 3.85 (m, 2H), 3.82 (d, *J* = 1.5 Hz, 3H), 3.06 (d, *J* = 4.1 Hz, 1H), 2.68 (dd, *J* = 13.1, 5.3 Hz, 1H), 2.55 (ddd, *J* = 13.1, 2.6, 1.5 Hz, 1H), 2.51 – 2.41 (m, 1H), 2.41 – 2.27 (m, 1H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 175.49 , 158.82 , 141.79 , 129.26 , 128.51 , 111.83 , 110.31 , 77.36 , 72.66 , 67.06 , 61.47 , 55.43 , 40.17 , 36.74 , 34.99 .

**HRMS** (FI): calcd for $C_{14}H_{16}O_4$ [M]$^{+\bullet}$ 248.10486; found 248.10591.

**3a,4,5,9b-Tetrahydronaphtho[2,1-*b*]furan-2(1*H*)-one (8e)**
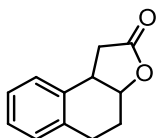


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.24 – 7.09 (m, 4H), 4.98 (td, *J* = 6.8, 3.4 Hz, 1H), 3.81 (dt, *J* = 9.6, 6.3 Hz, 1H), 3.09 (dd, *J* = 17.6, 9.5 Hz, 1H), 2.91 (ddd, *J* = 16.4, 9.5, 4.4 Hz, 1H), 2.75 – 2.65 (m, 1H), 2.58 (dd, *J* = 17.6, 5.6 Hz, 1H), 2.19 (dtd, *J* = 13.2, 6.4, 4.4 Hz, 1H), 1.99 (dddd, *J* = 14.0, 9.7, 4.6, 3.4 Hz, 1H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 176.51 , 136.03 , 135.68 , 129.17 , 128.98 , 127.21 , 126.98 , 78.87, 37.86 , 37.74 , 26.54 , 24.67 .

**HRMS** (FI): calcd for $C_{12}H_{12}O_2$ [M]$^{+\bullet}$ 188.08373; found 188.08441.

**1,2,5,6-Tetrahydro-4*H*-2,6-methanobenzo[*d*]oxocin-4-one (8e')**
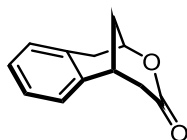


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.24 – 7.16 (m, 2H), 7.15 – 7.07 (m, 2H), 5.13 (dtd, *J* = 5.2, 3.6, 1.8 Hz, 1H), 3.35 – 3.21 (m, 2H), 3.10 (dd, *J* = 17.9, 3.8 Hz, 1H), 2.83 (dd, *J* = 17.9, 5.7 Hz, 1H), 2.65 (dt, *J* = 17.9, 2.2 Hz, 1H), 2.33 (ddt, *J* = 13.6, 4.7, 2.1 Hz, 1H), 2.23 – 2.10 (m, 1H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 175.16 , 136.48 , 125.98 , 122.56 , 120.25 , 119.97 , 119.79 , 119.04 , 111.47 , 69.41 , 41.51 , 36.13 , 32.24 , 28.93 .

**HRMS** (FI): calcd for $C_{12}H_{12}O_2$ [M]$^{+\bullet}$ 188.08373; found 188.08414.

**2,3,5',6'-Tetrahydrospiro[indene-1,4'-pyran]-2'(3'*H*)-one (6f)**
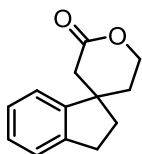


**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.30 – 7.19 (m, 4H), 7.19 – 7.11 (m, 1H), 4.57 – 4.42 (m, 2H), 2.96 (tt, *J* = 6.8, 0.7 Hz, 2H), 2.73 – 2.59 (m, 2H), 2.26 – 2.13 (m, 1H), 2.08 (t, *J* = 7.1 Hz, 2H), 1.91 (dddd, *J* = 14.1, 5.8, 4.6, 1.1 Hz, 1H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 170.86 , 147.65 , 142.73 , 127.63 , 127.04 , 125.03 , 122.29 , 77.23 , 67.00 , 45.58 , 41.67 , 38.82 , 33.76 , 29.62 .

**HRMS** (FI): calcd for $C_{13}H_{14}O_2$ [M]$^{+•}$ 202.09938; found 202.09993.

**1,5,6,7-Tetrahydro-1,7-methanobenzo[*e*]oxonin-3(2*H*)-one (6c)**



**¹H NMR** (400 MHz, Chloroform-*d*) δ 7.30 – 7.13 (m, 5H), 4.14 – 3.97 (m, 2H), 3.68 (dt, *J* = 9.4, 4.9 Hz, 1H), 3.53 – 3.44 (m, 1H), 2.90 (dd, *J* = 12.5, 4.9 Hz, 1H), 2.66 – 2.55 (m, 2H), 2.39 (dddd, *J* = 14.8, 11.4, 5.0, 3.5 Hz, 1H), 2.22 (dt, *J* = 13.7, 1.0 Hz, 1H), 2.00 (ddt, *J* = 14.8, 4.5, 3.1 Hz, 1H).

**¹³C NMR** (101 MHz, Chloroform-*d*) δ 175.54 , 145.68 , 144.08 , 127.80 , 127.57 , 124.89 , 124.59 , 66.44 , 42.68 , 42.55 , 41.15 , 36.87 , 36.32 .

**HRMS** (FI): calcd for $C_{13}H_{14}O_2$ [M]$^{+•}$ 202.09938; found 202.10012.
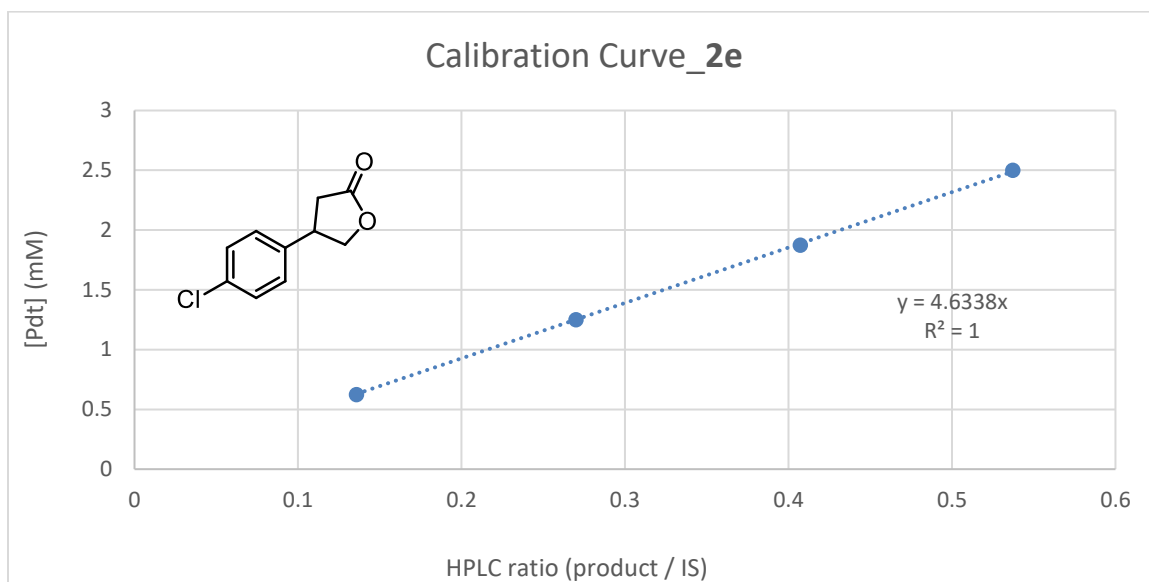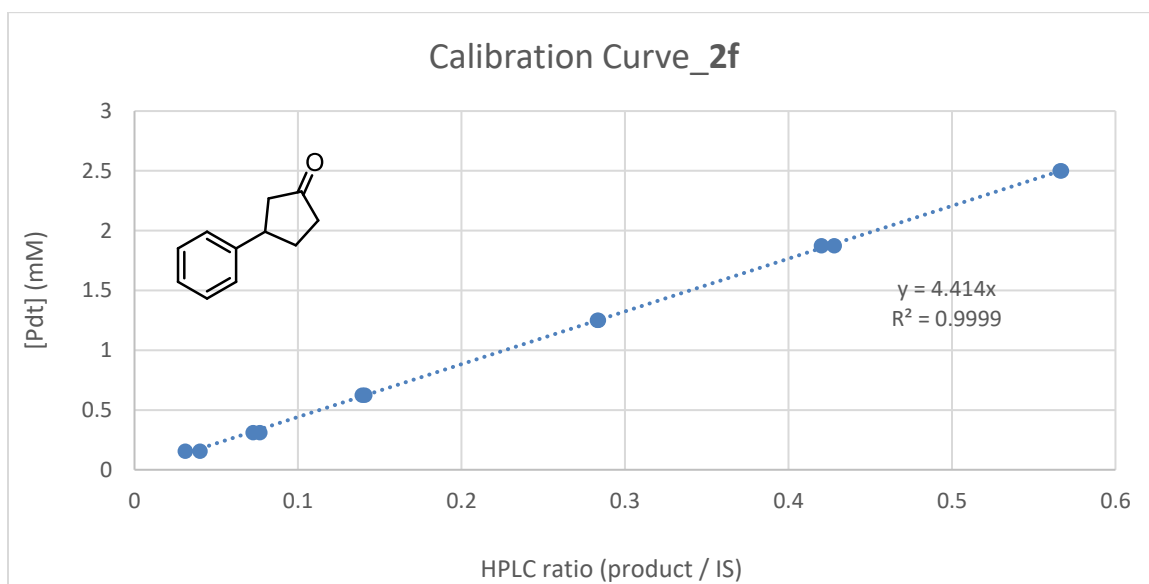
**HPLC Calibration Curves**

Calibration curves of synthesized reference compounds were created for the determination of yield and TTN.

For compounds **2a–6c**: For each substrate, four different concentrations of product (0.625, 1.25, 1.875, and 2.5 mM) with 0.67 mM internal standard in 600 µL acetonitrile solutions were mixed each with 400 µL water. The mixtures were vortexed and then analyzed by HPLC based on UV absorbance at 210 nm. All data points represent the average of duplicate runs. The calibration curves depict the ratio of product area to internal standard area (x-axis) against product concentration in mM (y-axis). Notes: Pdt = product area, IS = internal standard area, [Pdt] = product concentration in reaction, [PC] = protein concentration in reaction, Avg. TTN = average total turnover number, SD TTN = standard deviation of TTN, Avg. Yield = average yield, SD Yield = standard deviation of yield.
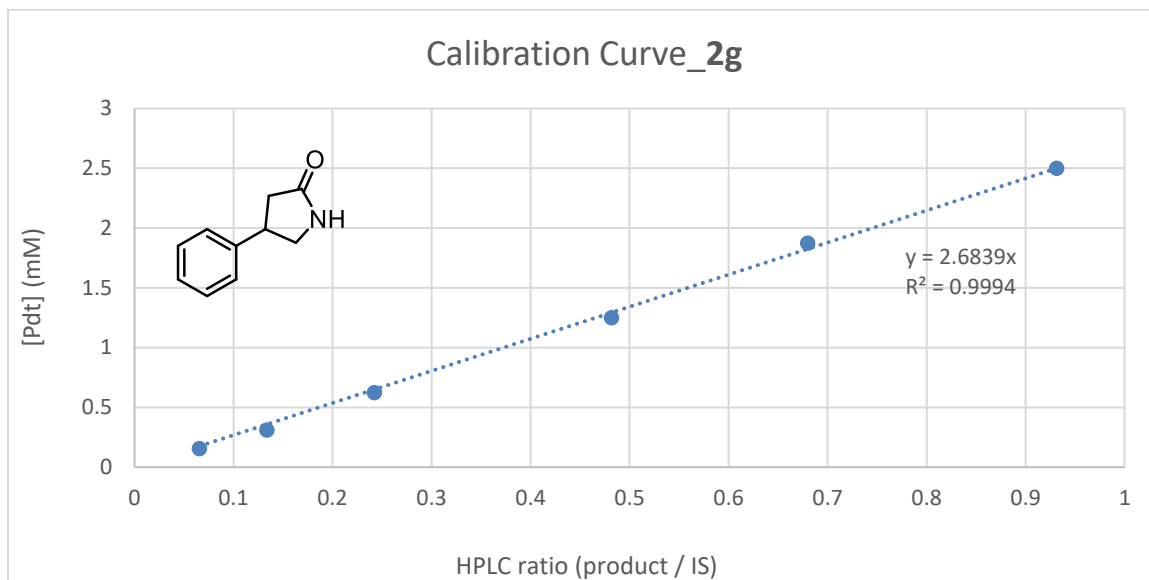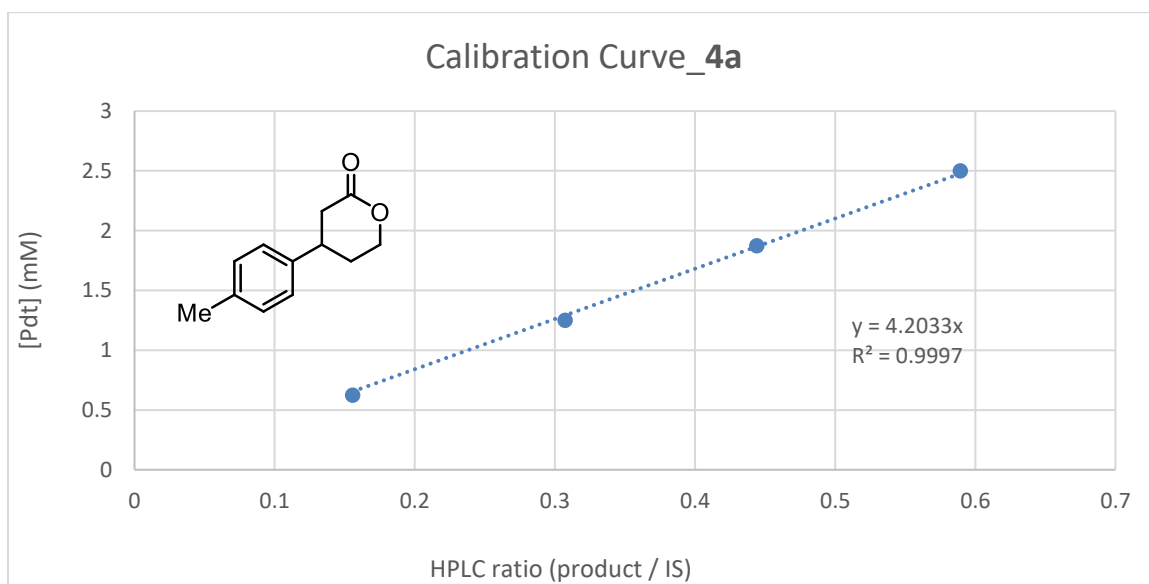
For compounds **8a–8f'**: For each substrate, four different concentrations of product (0.625, 1.25, 1.875, and 2.5 mM) with 0.67 mM internal standard in 600 µL hexane solutions were mixed each with 400 µL water. The mixtures were vortexed and centrifuged. The supernatant was transferred to a new vial and analyzed by HPLC based on UV absorbance at 210 nm. All data points represent the average of duplicate runs. The calibration curves depict the ratio of product area to internal standard area (x-axis) against product concentration in mM (y-axis). Notes: Pdt = product area, IS = internal standard area, [Pdt] = product concentration in reaction, [PC] = protein concentration in reaction, Avg. TTN = average total turnover number, SD TTN = standard deviation of TTN, Avg. Yield = average yield, SD Yield = standard deviation of yield.
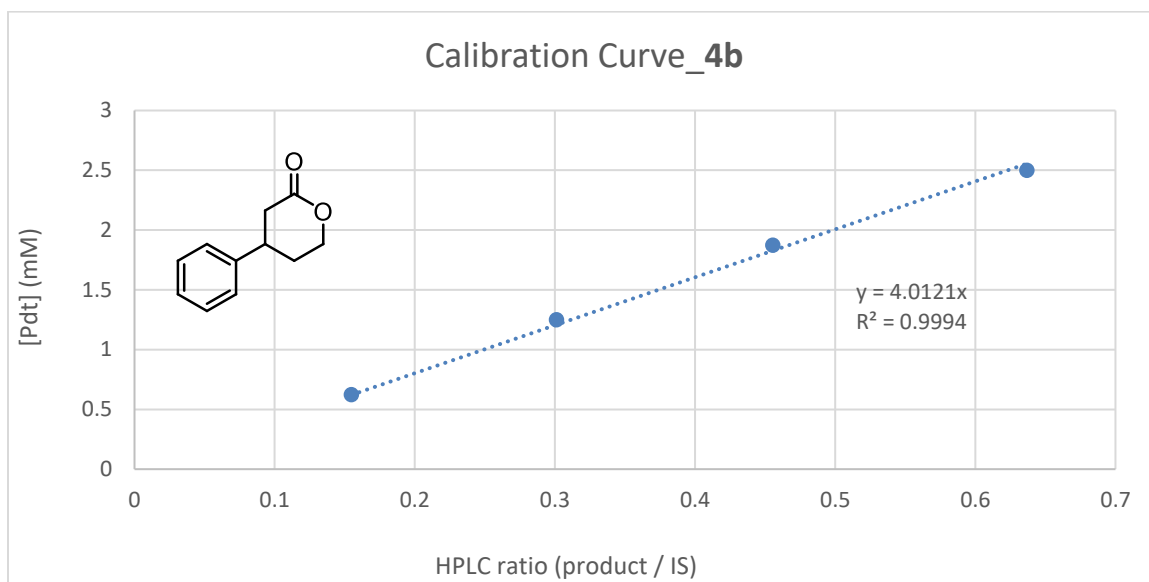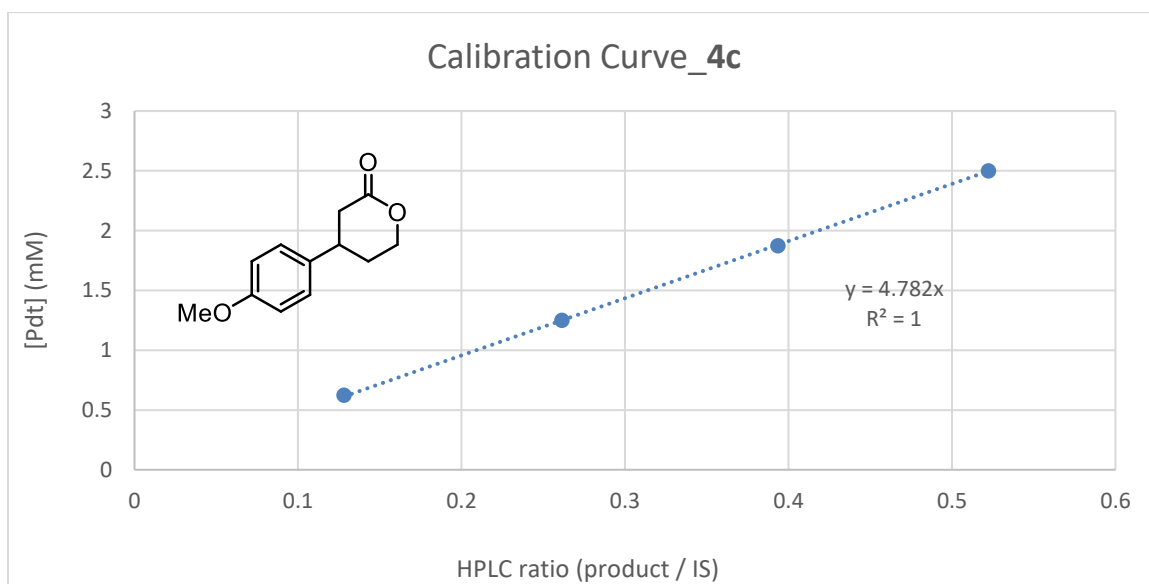
**4-Phenyldihydrofuran-2(3*H*)-one (2a)**
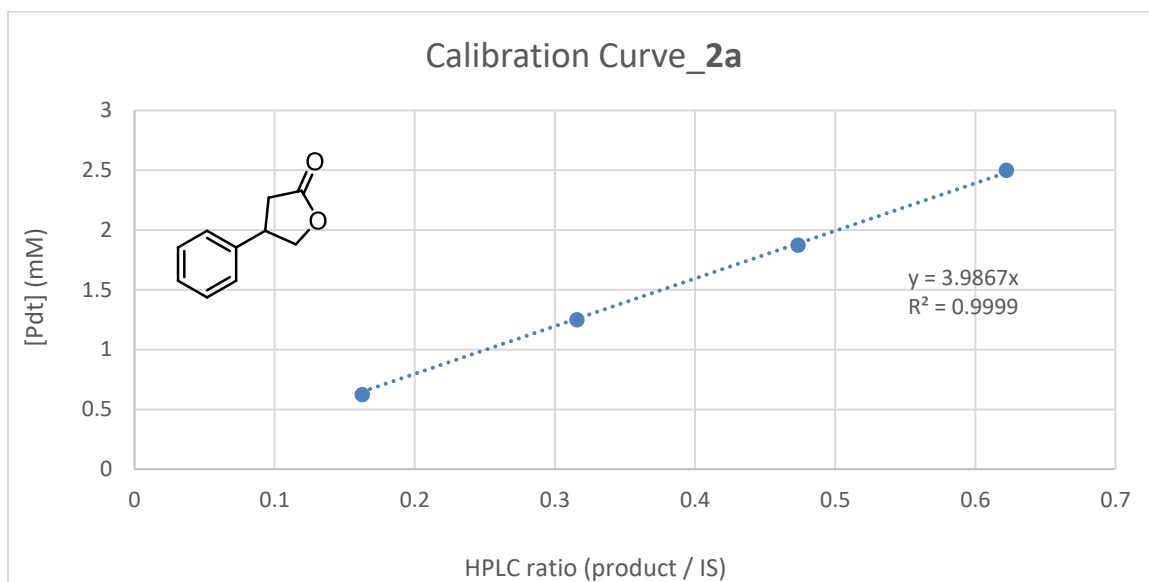


**4-(*p*-Tolyl)dihydrofuran-2(3*H*)-one (2b)**
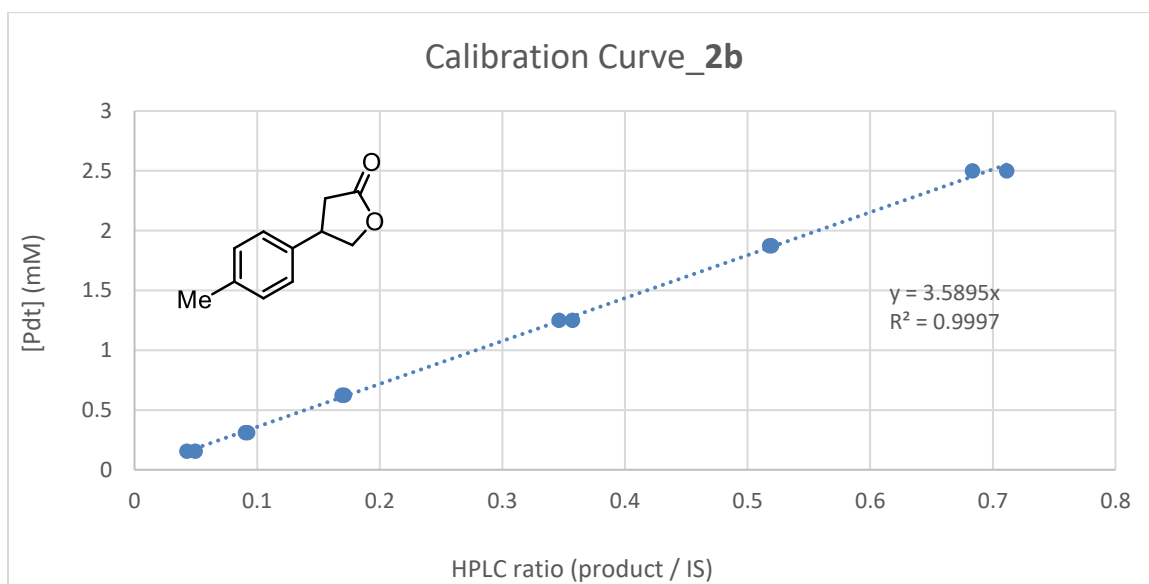
**4-(4-Methoxyphenyl)dihydrofuran-2(3*H*)-one (2c)**
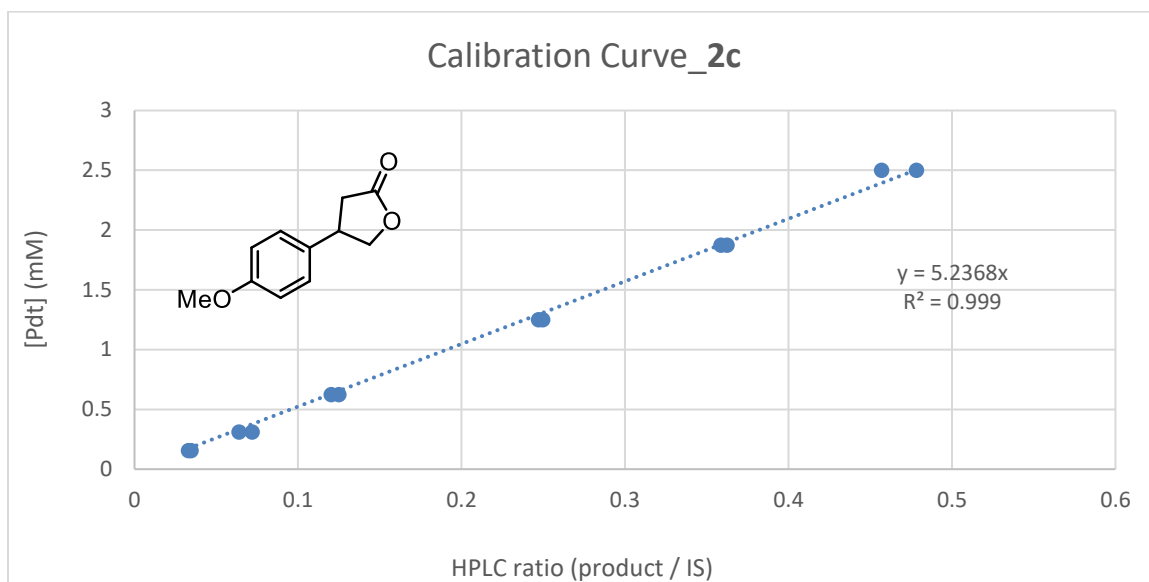


**4-(4-Fluorophenyl)dihydrofuran-2(3*H*)-one (2d)**
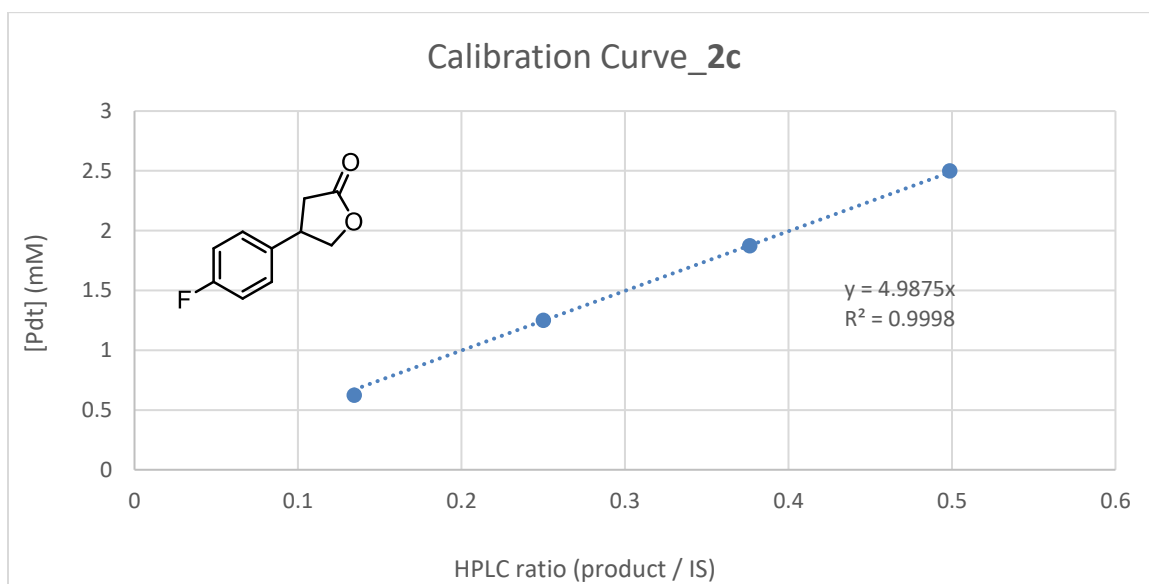
**4-(4-Chlorophenyl)dihydrofuran-2(3*H*)-one (2e)**



**3-Phenylcyclopentan-1-one (2f)**

**4-Phenylpyrrolidin-2-one (2g)**



Calibration Curve_**2g**

y = 2.6839x
R² = 0.9994

[Pdt] (mM)

HPLC ratio (product / IS)

**4-(*p*-Tolyl)tetrahydro-2*H*-pyran-2-one (4a)**



Calibration Curve_**4a**

y = 4.2033x
R² = 0.9997

[Pdt] (mM)

HPLC ratio (product / IS)

**4-Phenyltetrahydro-2_H_-pyran-2-one (4b)**



Calibration Curve_**4b**

y = 4.0121x
R² = 0.9994

[Pdt] (mM)

HPLC ratio (product / IS)

**4-(4-Methoxyphenyl)tetrahydro-2_H_-pyran-2-one (4c)**



Calibration Curve_**4c**

y = 4.782x
R² = 1

[Pdt] (mM)

HPLC ratio (product / IS)

**4-(4-Fluorophenyl)tetrahydro-2*H*-pyran-2-one (4d)**
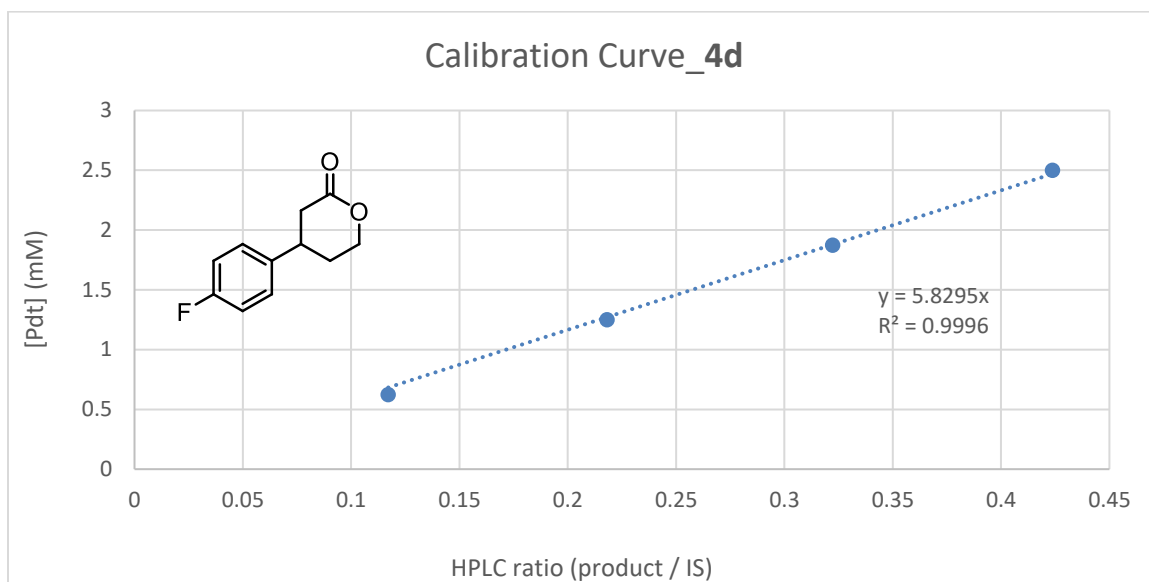


**4-(Furan-2-yl)tetrahydro-2*H*-pyran-2-one (4e)**
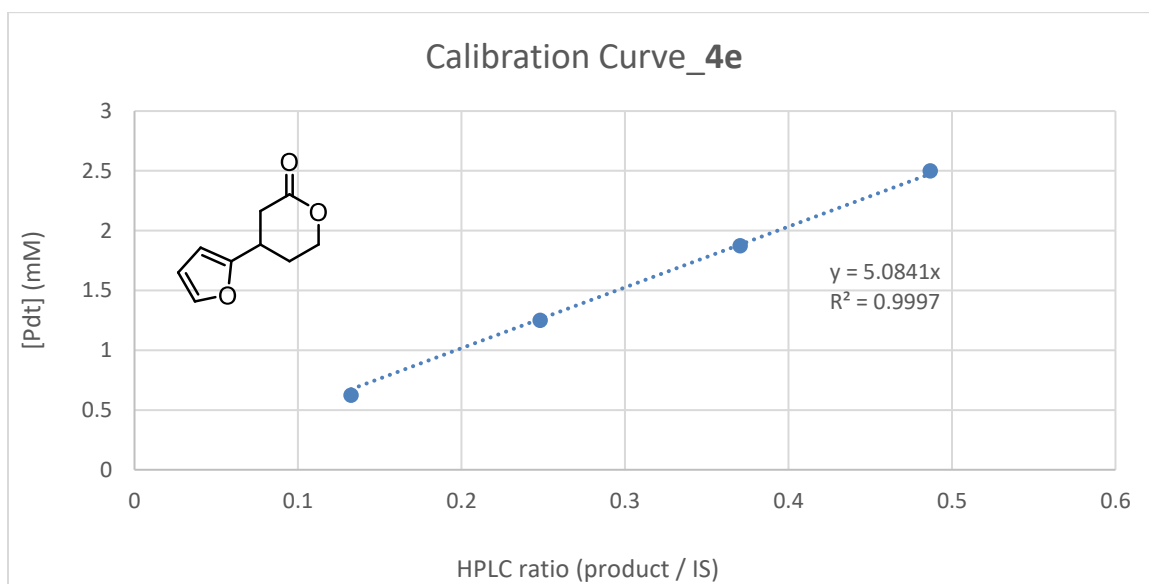
**4-(Thiophen-2-yl)tetrahydro-2*H*-pyran-2-one (4f)**
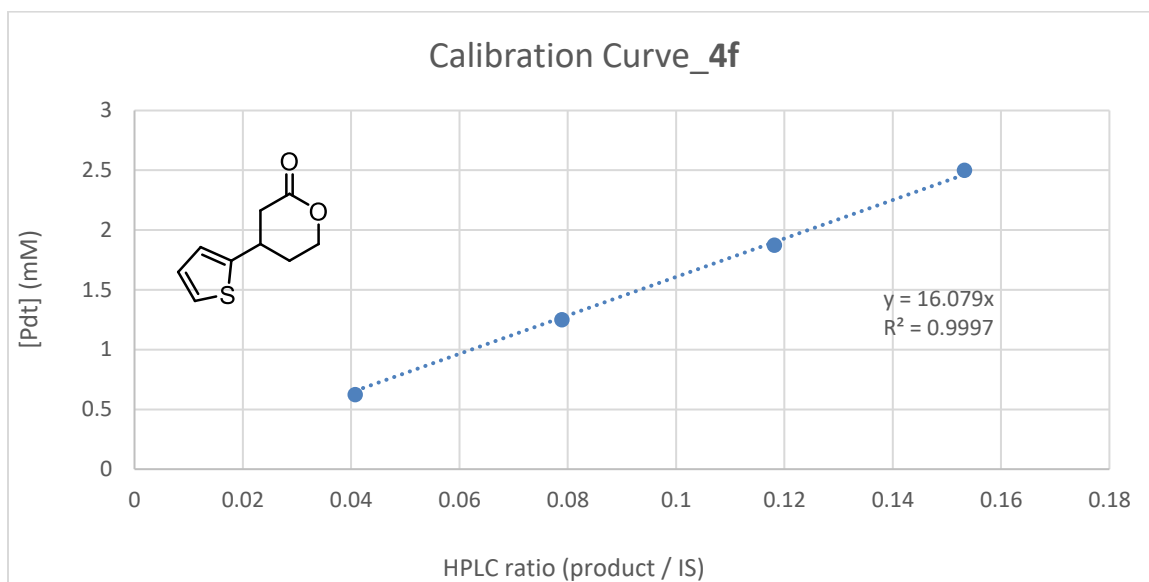


**4-(4-Methoxyphenyl)oxepan-2-one (6a)**
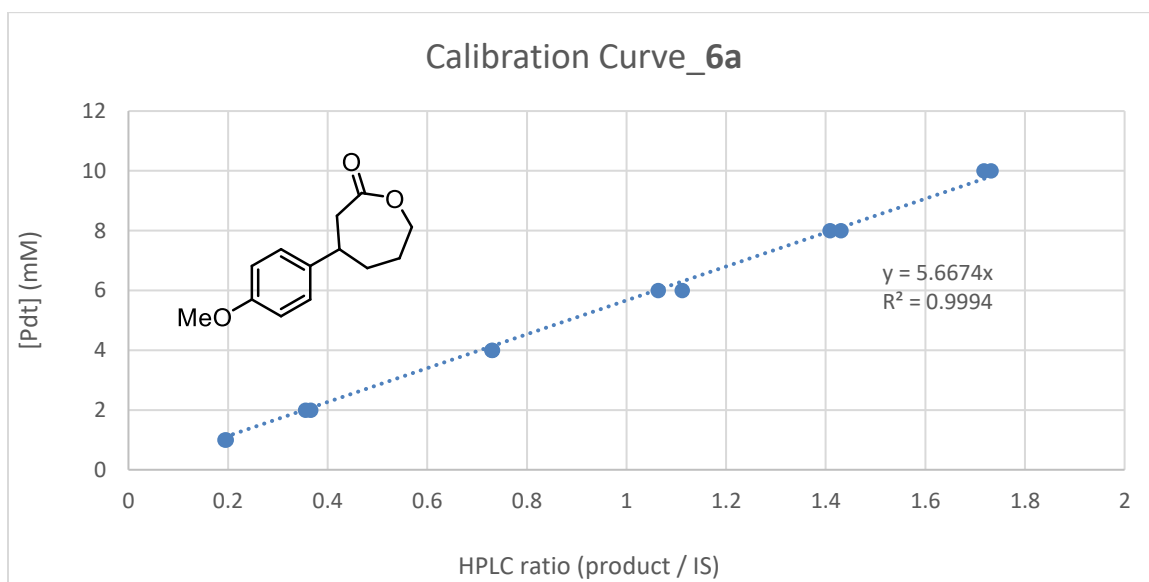
**(*E*)-4-(4-Methoxystyryl)oxepan-2-one (6b)**
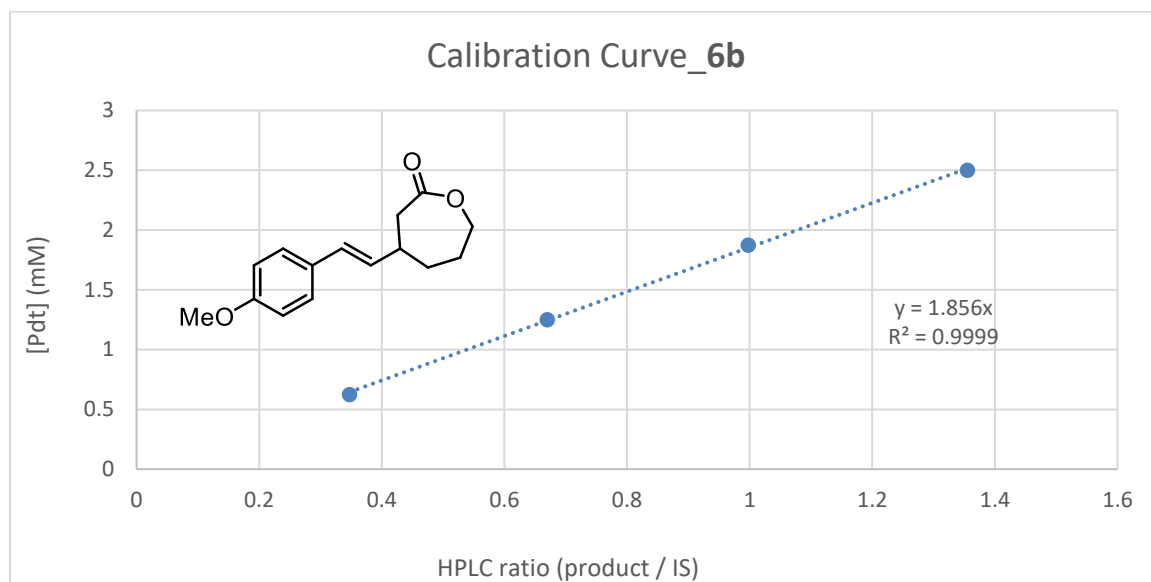


**4-(1*H*-Indol-3-yl)oxepan-2-one (6c)**
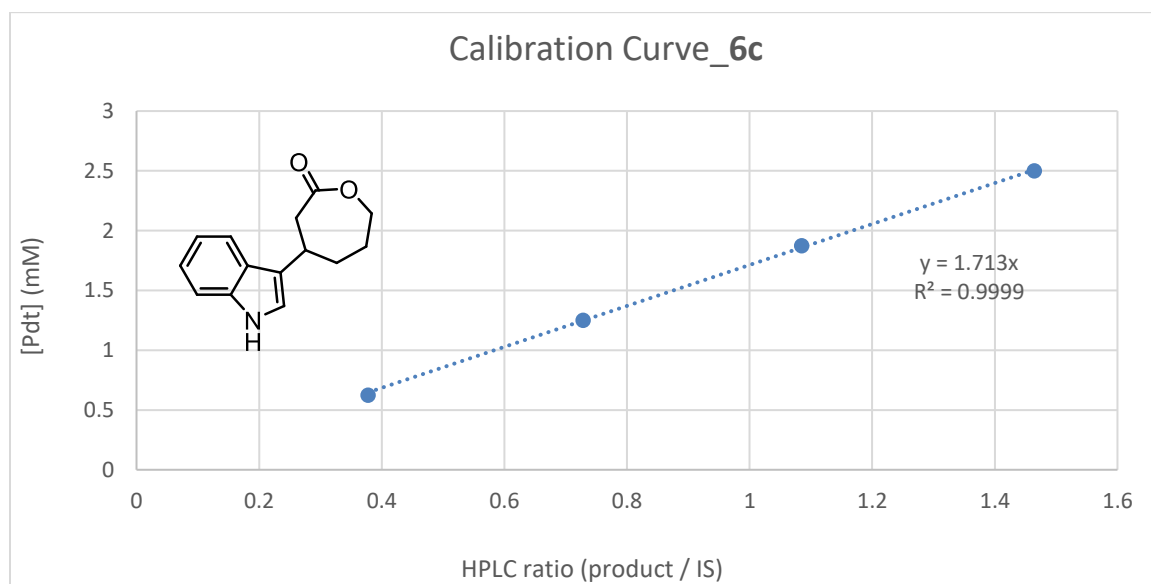
**3,3a,8,8a-Tetrahydro-2*H*-indeno[2,1-*b*]furan-2-one (8a)**
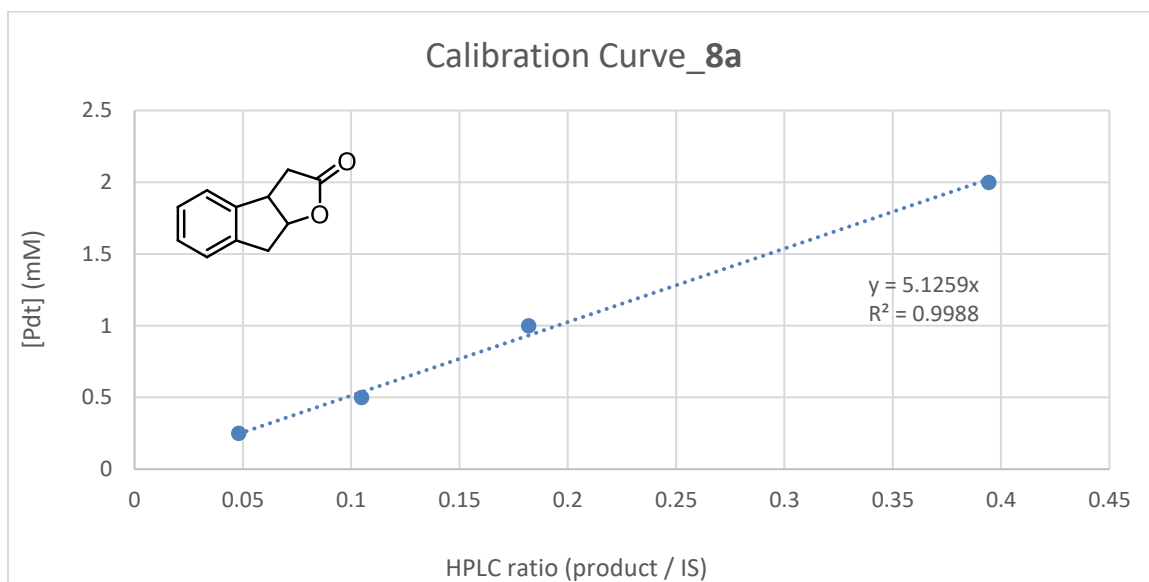


**4,4a,9,9a-Tetrahydroindeno[2,1-*c*]pyran-3(1*H*)-one (8b)**
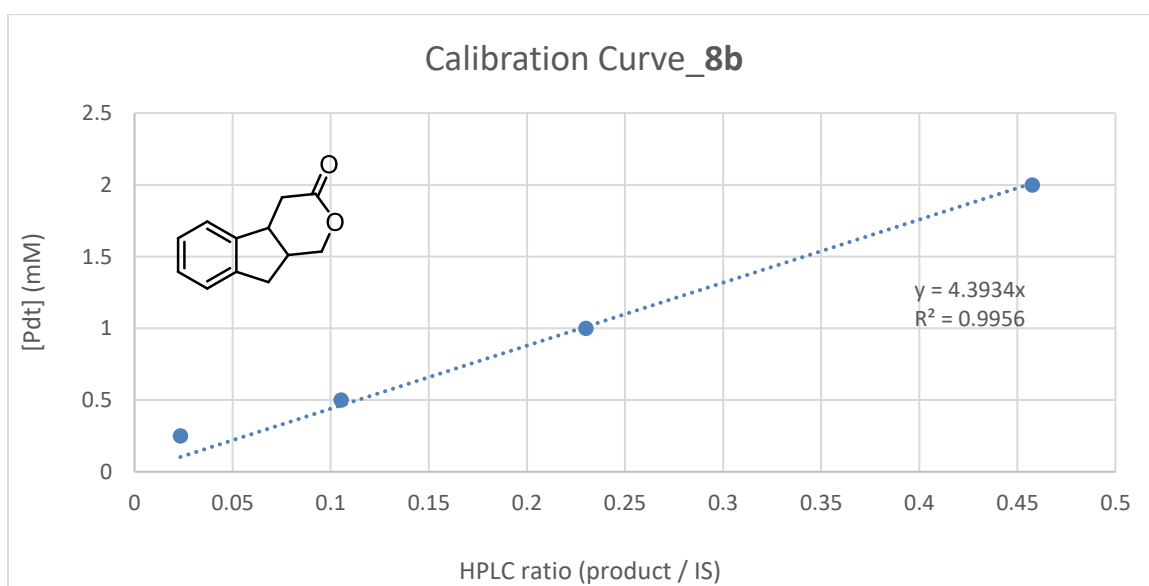
**3,4,5',6'-Tetrahydro-2*H*-spiro[naphthalene-1,4'-pyran]-2'(3'*H*)-one (8c)**
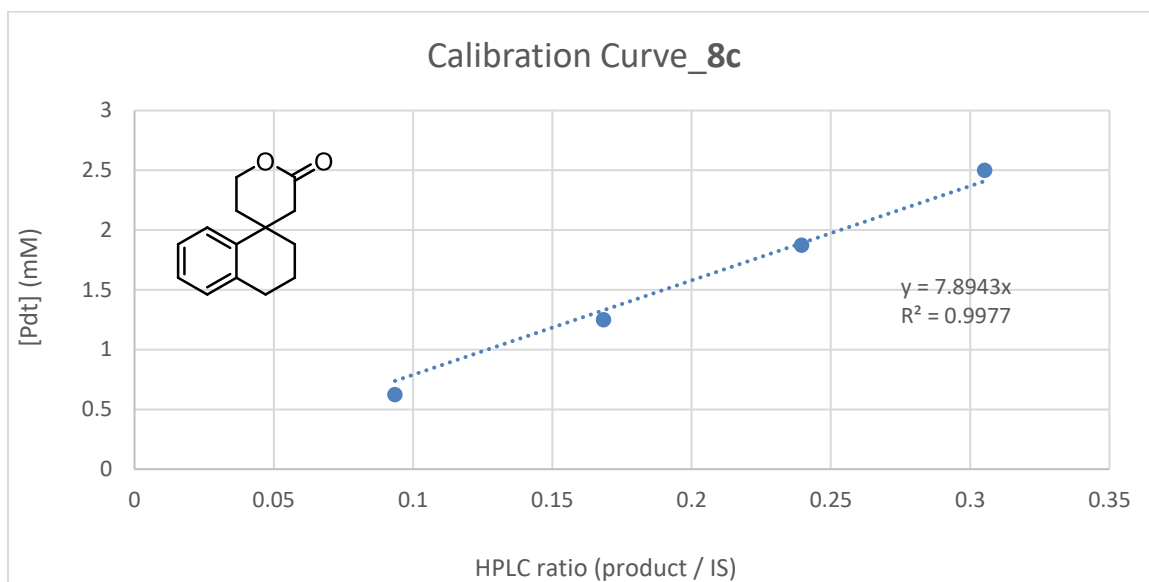


Calibration Curve_**8c**

$y = 7.8943x$
$R^2 = 0.9977$

[Pdt] (mM) vs. HPLC ratio (product / IS)

**5-Methoxy-2,3,5',6'-tetrahydrospiro[indene-1,4'-pyran]-2'(3'*H*)-one (8d)**



Calibration Curve_**8d**

$y = 2.1425x$
$R^2 = 0.9989$

[Pdt] (mM) vs. HPLC ratio (product / IS)
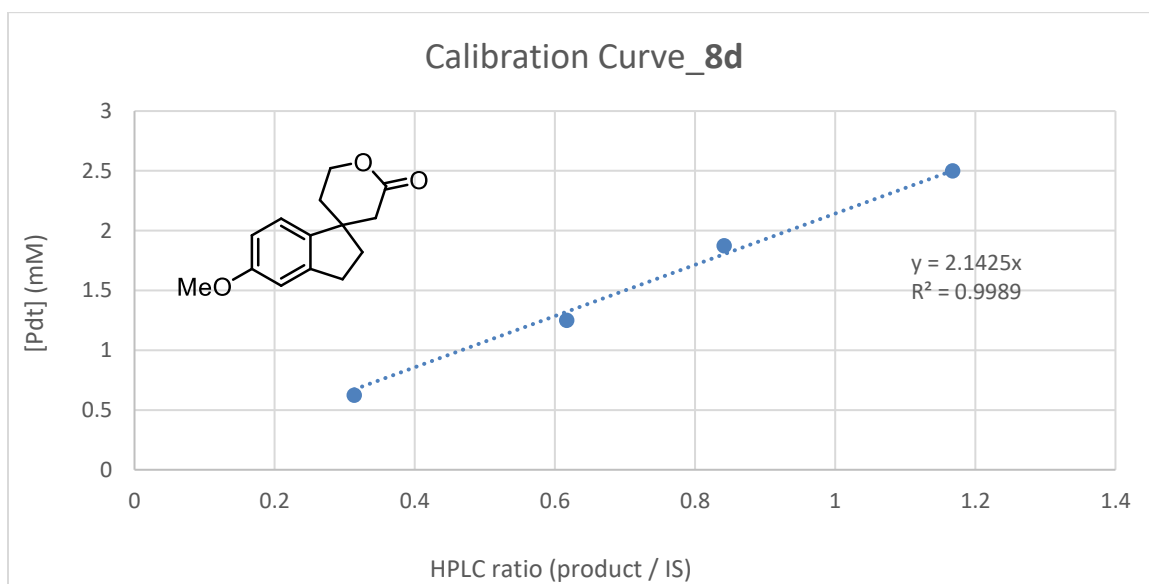
**3a,4,5,9b-Tetrahydronaphtho[2,1-*b*]furan-2(1*H*)-one (8e)**
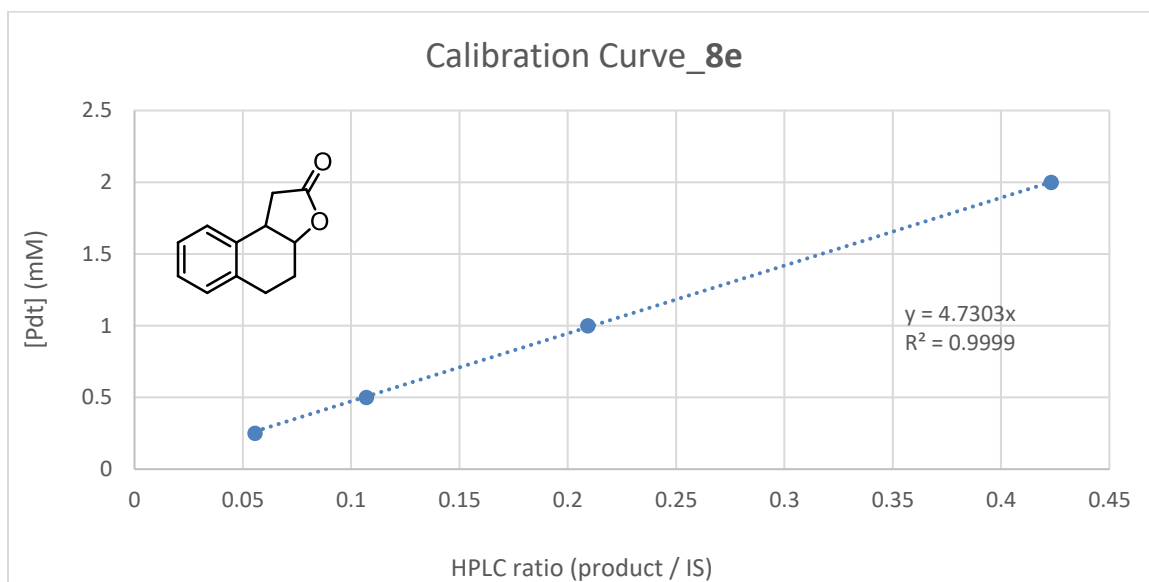


**(6,6a)-1,2,5,6,6a,7-Hexahydro-4*H*-2,6-methanobenzo[*d*]oxocin-4-one (8e')**
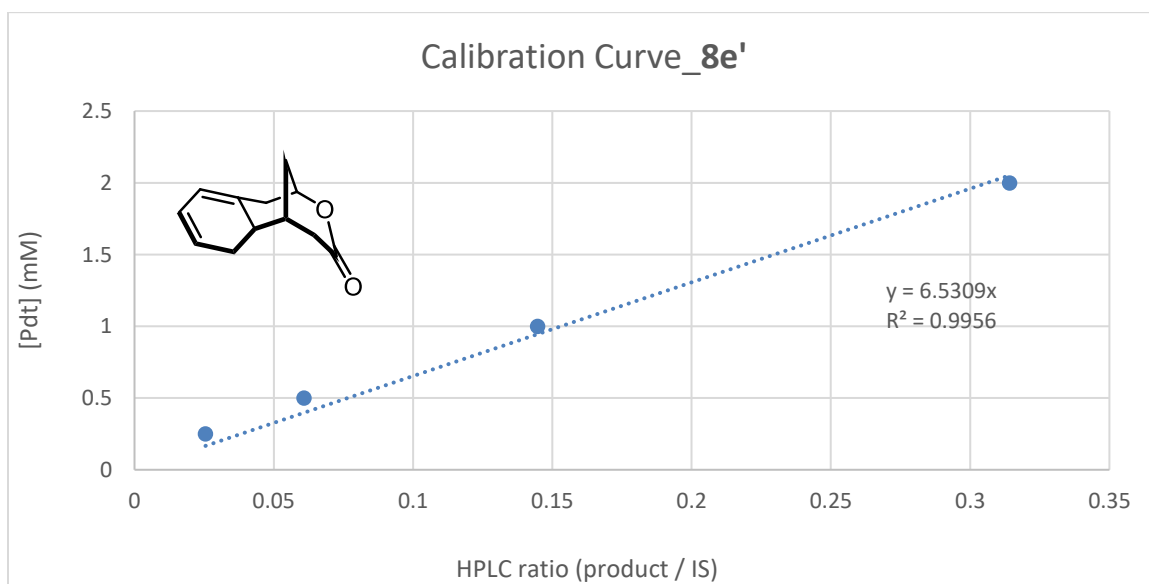
**2,3,5',6'-Tetrahydrospiro[indene-1,4'-pyran]-2'(3'*H*)-one (8f)**
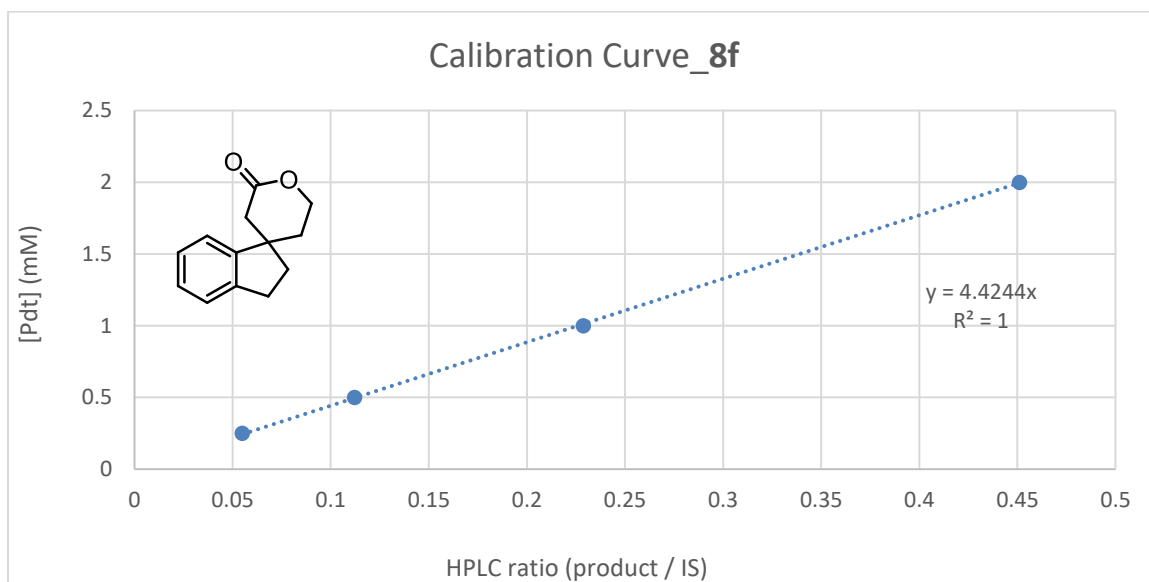


**(7)-1,5,6,7-Tetrahydro-1,7-methanobenzo[*e*]oxonin-3(2*H*)-one (8f')**
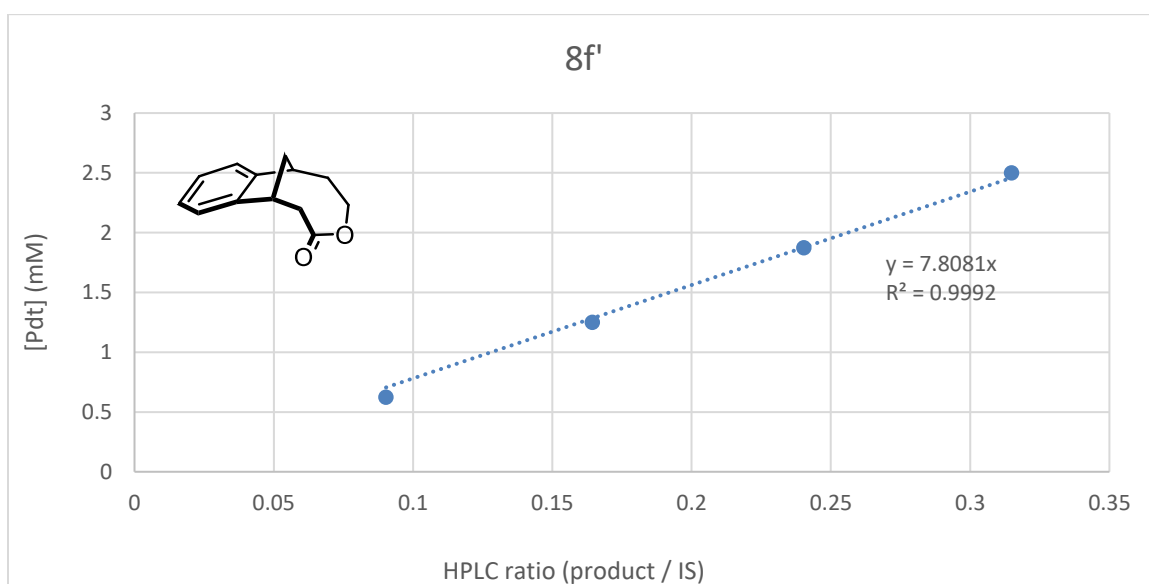
**Enantioselectivity Determination**

The absolute configuration of enzymatic product **2c** was assigned to be *S* by comparing the elution order of two enantiomers with a literature report under same elution conditions using the same column (Chiralpak IB).[18] The other *γ*-lactone products **2a**–**2b**, **2d**–**2m** were assigned by analogy.

The absolute configuration of enzymatic product **4b** was assigned to be *R* by comparing the elution order of two enantiomers with a literature report under same elution conditions using the same column (Chiralpak AS-H).[18] The other *δ*-lactone products **4a**, **4c**–**4f** were assigned by analogy.

The absolute stereochemistry for enzymatic product **6a** was assigned as *S* through X-ray crystallography (see **Section 8**). The other *ε*-lactone products **6b** and **6c** were assigned by analogy.

**4-Phenyldihydrofuran-2(3*H*)-one (2a)**
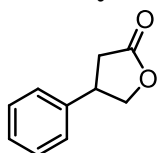


**HPLC conditions:** Chiralpak IA, 2.5% *i*-PrOH in hexane, 1.2 mL/min, 32 ℃, 220 nm



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 14.53 | 11606.2 | 543 | 0.3563 | 49.635 | 0.545 |
| 2 | 15.648 | 11776.7 | 493.9 | 0.3974 | 50.365 | 0.483 |
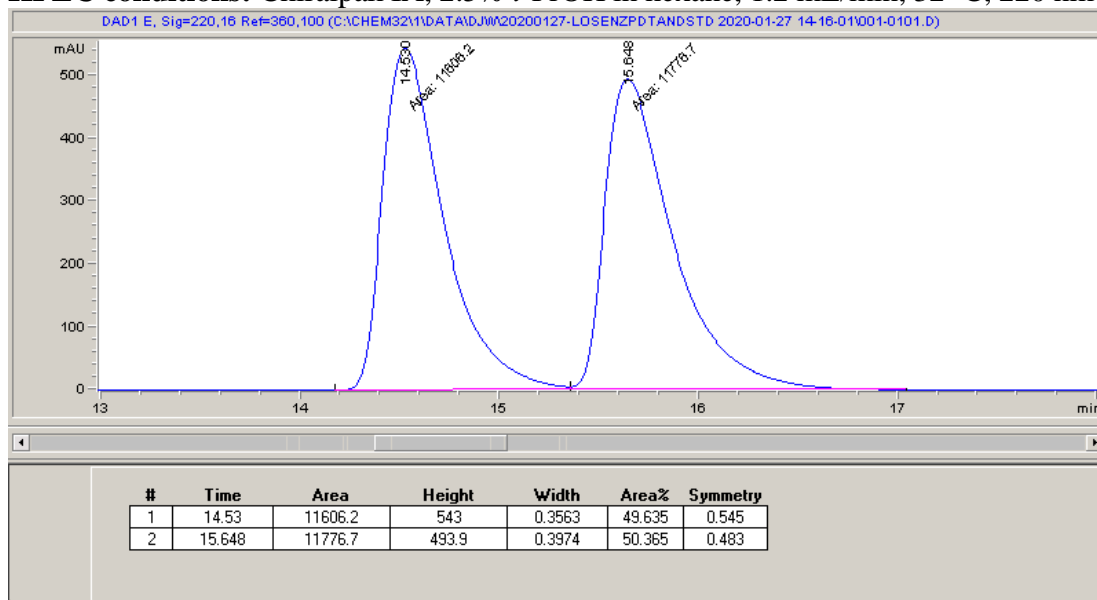
**Enzymatic preparation of 2a with P411-LAS-5247: >99% e.e.**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 15.813 | 6771.9 | 290.4 | 0.3886 | 100.000 | 0.492 |
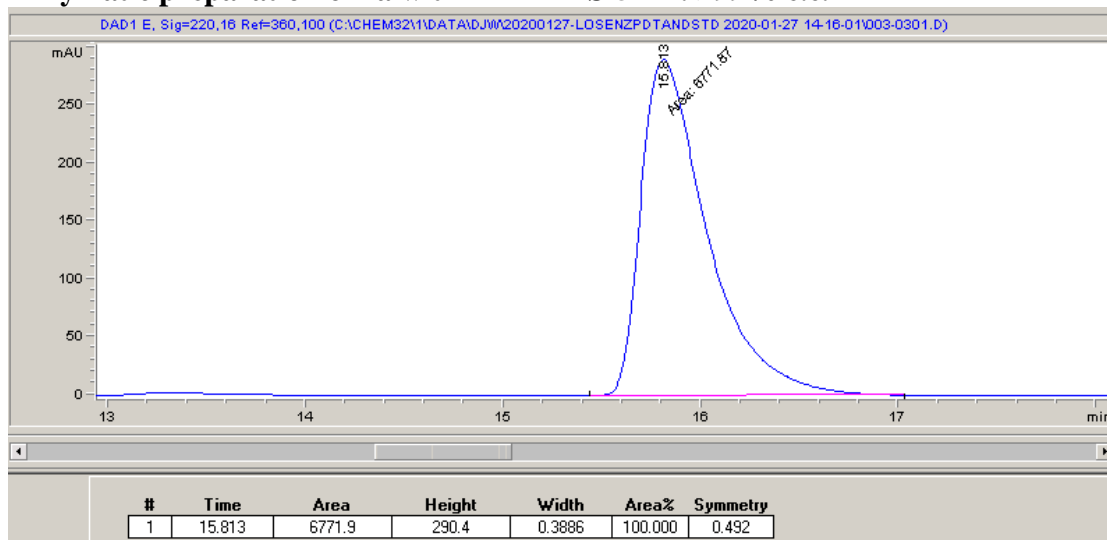
**4-(*p*-Tolyl)dihydrofuran-2(3*H*)-one (2b)**



**SFC conditions:** Chiralpak AD-H, 5% *i*-PrOH in supercritical $CO_2$, 2.5 mL/min, 210 nm



```
Peak  RetTime  Type   Width       Area        Height       Area
 #     [min]          [min]     [mAU*s]        [mAU]         %
----|-------|----|-------|----------|----------|--------|
  1    4.856  BB    0.1019  1146.25952   171.62752   49.6720
  2    5.296  BB    0.1147  1161.39697   156.22182   50.3280
```

**Enzymatic preparation of 2b with P411-LAS-5247: 80% e.e.**



```
Peak  RetTime  Type   Width       Area        Height       Area
 #     [min]          [min]     [mAU*s]        [mAU]         %
----|-------|----|-------|----------|----------|--------|
  1    4.860  BV    0.1360  6139.50293   688.93280   90.1516
  2    5.364  VB    0.1209   670.69690    84.23247    9.8484
```

**4-(4-Methoxyphenyl)dihydrofuran-2(3*H*)-one (2c)**
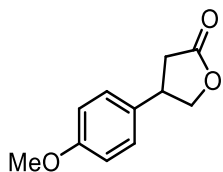


**HPLC conditions:** Chiralpak ODH, 2.5% *i*-PrOH in hexane, 1.2 mL/min, 32 ºC, 220 nm
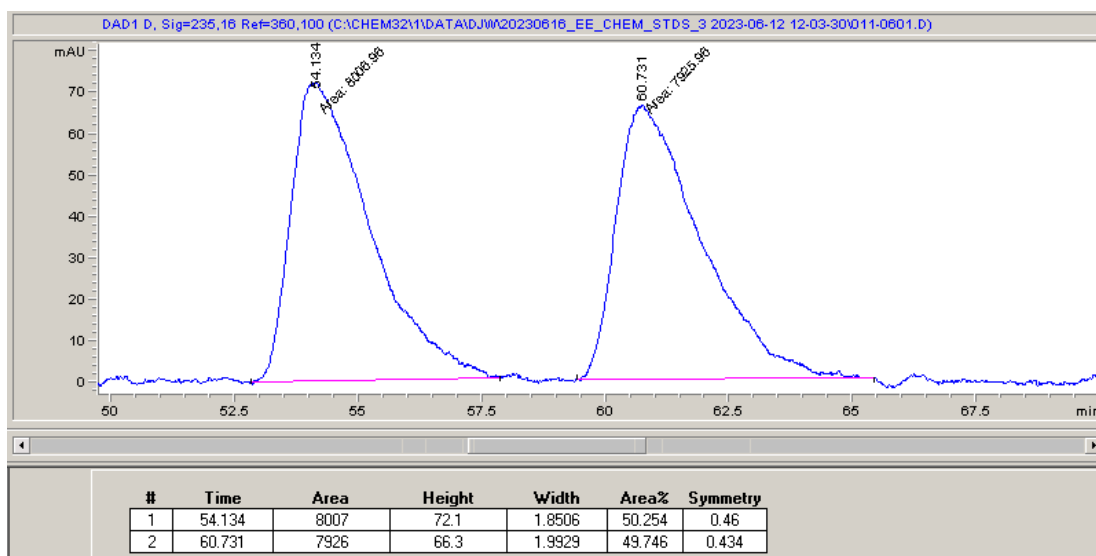


| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 54.134 | 8007 | 72.1 | 1.8506 | 50.254 | 0.46 |
| 2 | 60.731 | 7926 | 66.3 | 1.9929 | 49.746 | 0.434 |

**Enzymatic preparation of 2c with P411-LAS-5247: 88% e.e.**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 55.793 | 33903.8 | 293.5 | 1.9252 | 94.012 | 0.371 |
| 2 | 61.851 | 2159.5 | 21.9 | 1.6406 | 5.988 | 0.618 |
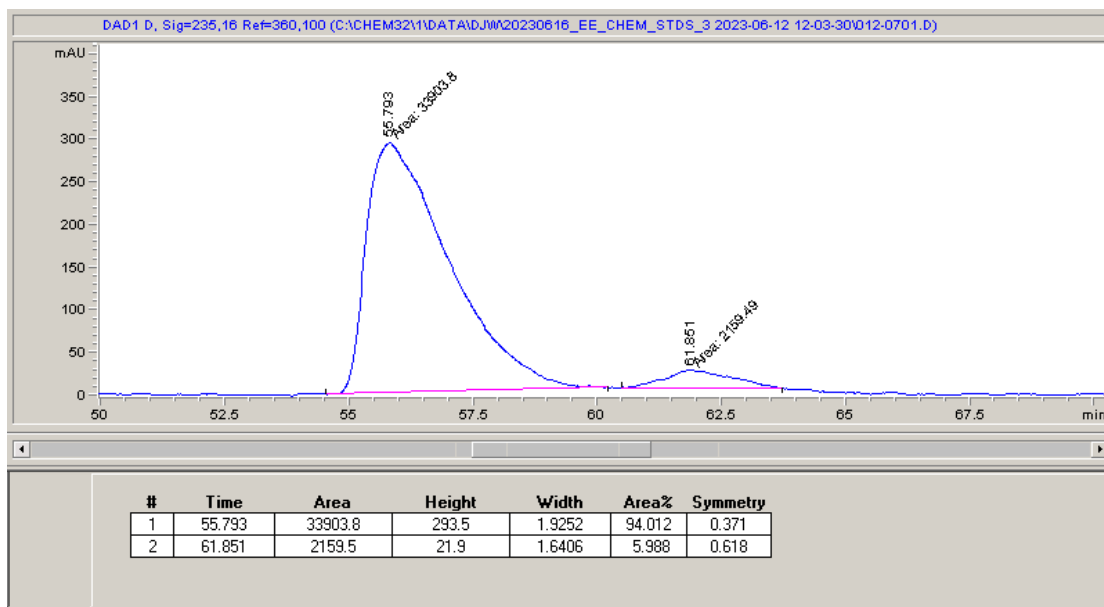
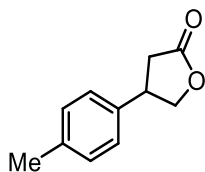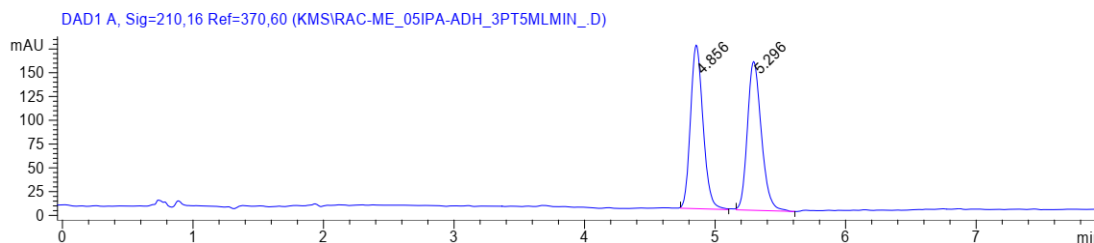**HPLC conditions:** Chiralpak IB, 15% *i*-PrOH in hexane, 1 mL/min, 220 nm
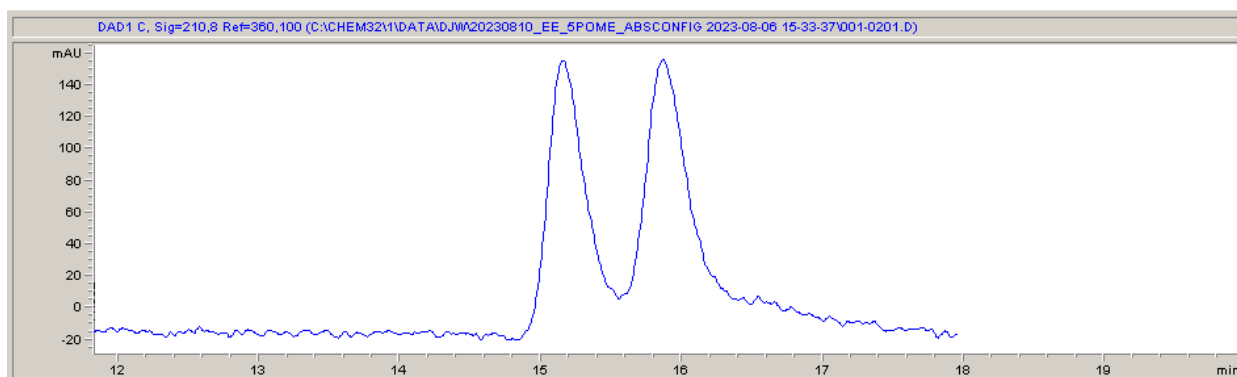


**Enzymatic preparation of 2c with P411-LAS-5247:**
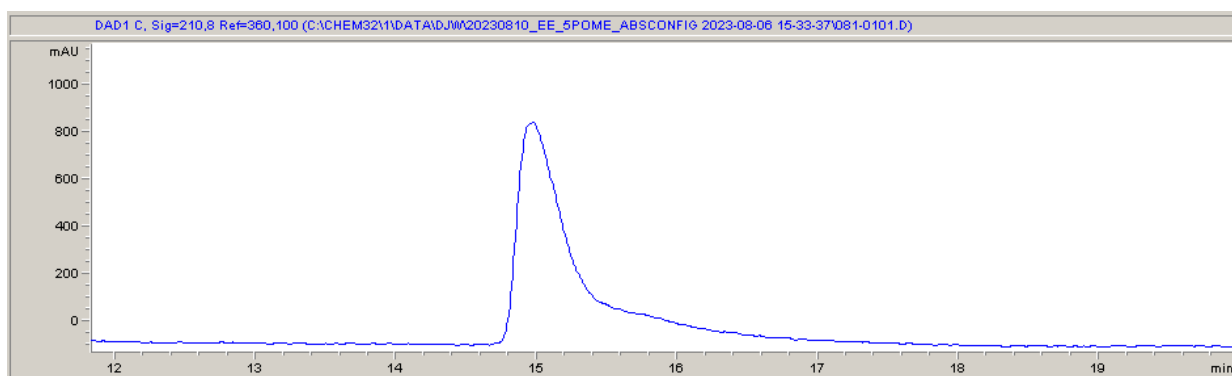
**4-(4-Fluorophenyl)dihydrofuran-2(3H)-one (2d)**



**SFC conditions:** Chiralpak AD-H, 5% *i*-PrOH in supercritical $CO_2$, 2.5 mL/min, 210 nm
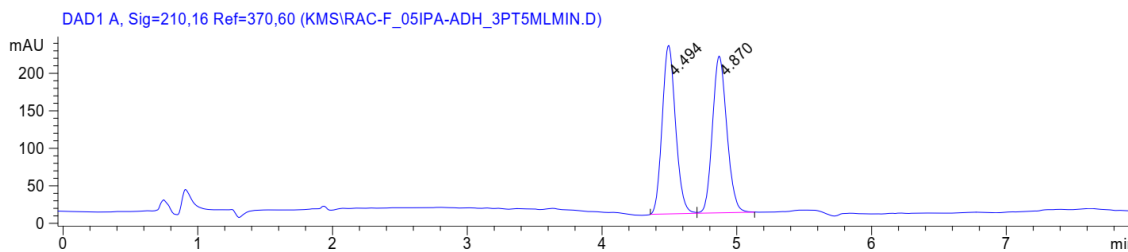


```
Peak  RetTime  Type   Width      Area      Height      Area
 #     [min]          [min]    [mAU*s]      [mAU]        %
----|-------|----|-------|----------|----------|--------|
  1    4.494  BV    0.1086  1548.69226  224.07057  50.0417
  2    4.870  VB    0.1145  1546.11169  208.46843  49.9583
```

**Enzymatic preparation of 2d with P411-LAS-5247: 77% e.e.**
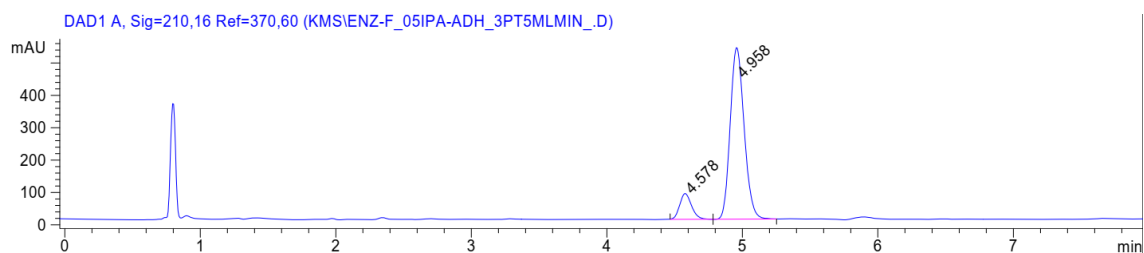


```
Peak  RetTime  Type   Width      Area      Height      Area
 #     [min]          [min]    [mAU*s]      [mAU]        %
----|-------|----|-------|----------|----------|--------|
  1    4.578  BV    0.0951   482.87766   79.23338  11.3347
  2    4.958  VB    0.1113  3777.29224  528.91382  88.6653
```

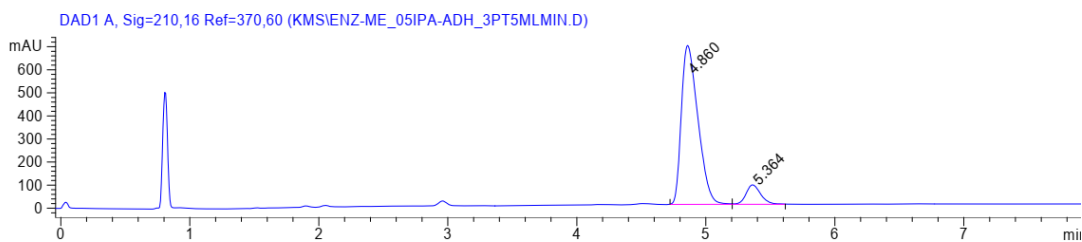**4-(4-Chlorophenyl)dihydrofuran-2(3*H*)-one (2e)**
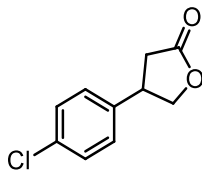


**SFC conditions:** Chiralpak AD-H, 5% *i*-PrOH in supercritical $CO_2$, 2.5 mL/min, 210 nm
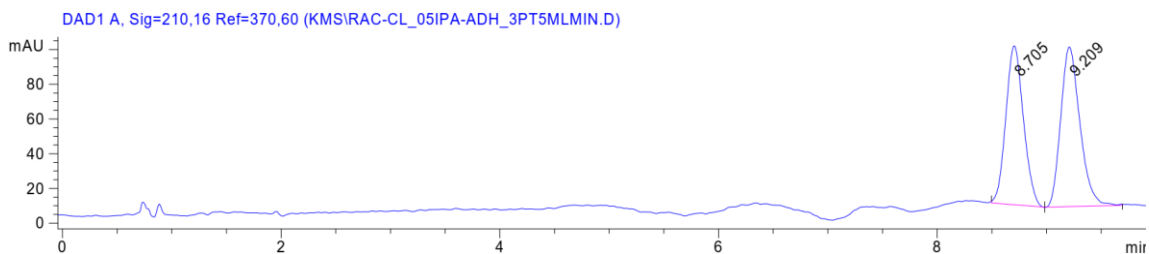


| Peak # | RetTime [min] | Type | Width [min] | Area [mAU*s] | Height [mAU] | Area % |
|----|-------|----|-------|----------|---------|---------|
| 1 | 8.705 | BV | 0.1705 | 1023.10614 | 91.49740 | 47.7289 |
| 2 | 9.209 | VB | 0.1863 | 1120.47363 | 91.84954 | 52.2711 |

**Enzymatic preparation of 2e with P411-LAS-5247: 16% e.e.**
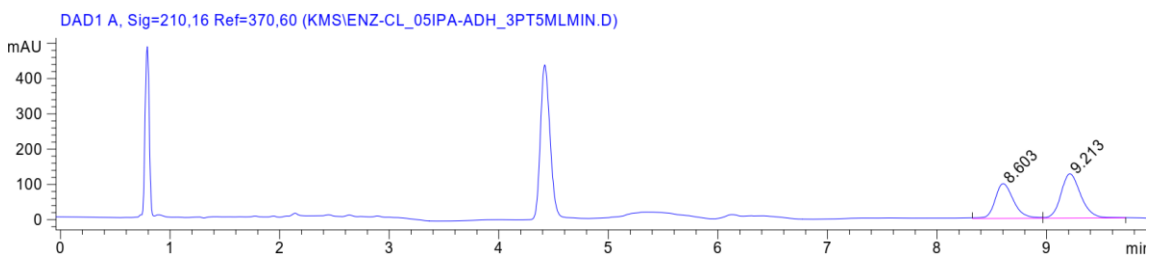


| Peak # | RetTime [min] | Type | Width [min] | Area [mAU*s] | Height [mAU] | Area % |
|----|-------|----|-------|----------|---------|---------|
| 1 | 8.603 | BV | 0.1806 | 1152.80786 | 98.45173 | 42.0722 |
| 2 | 9.213 | VB | 0.1914 | 1587.26416 | 125.60918 | 57.9278 |

**4-Phenyldihydrofuran-2(3*H*)-one (2f)**
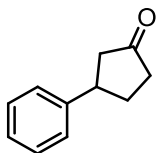


**SFC conditions:** Chiralpak AD-H, 5% *i*-PrOH in supercritical $CO_2$, 2.5 mL/min, 210 nm
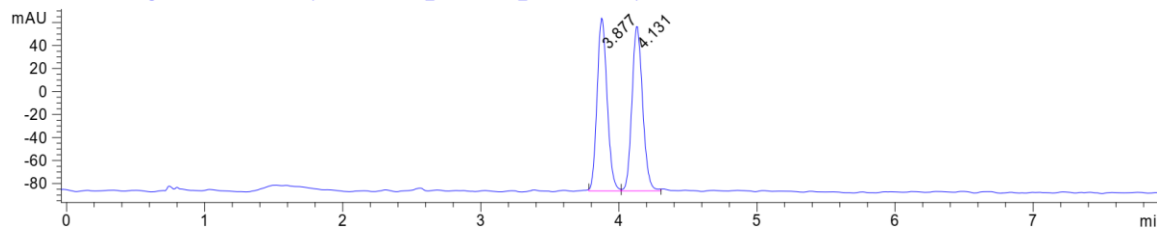


```
Peak  RetTime  Type   Width      Area       Height      Area
 #     [min]          [min]     [mAU*s]     [mAU]         %
----|-------|----|-------|----------|----------|--------|
  1   3.877  BV    0.0794  761.17230  149.62044  49.7216
  2   4.131  VB    0.0829  769.69629  142.76801  50.2784
```

**Enzymatic preparation of 2f with P411-LAS-5247: 97% e.e.**
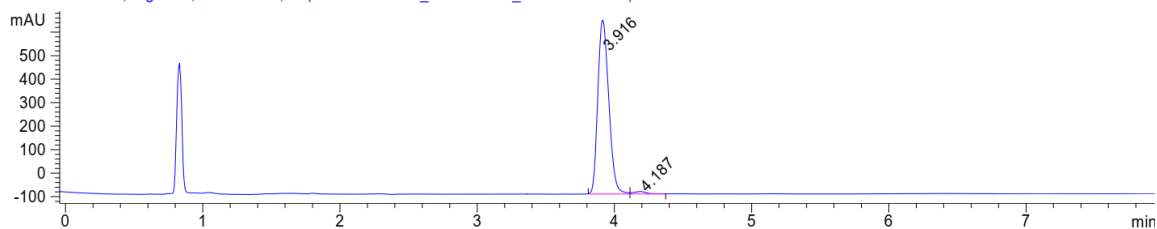


```
Peak  RetTime  Type   Width      Area       Height      Area
 #     [min]          [min]     [mAU*s]     [mAU]         %
----|-------|----|-------|----------|----------|--------|
  1   3.916  BV    0.0890  4094.99829  735.15277  98.3037
  2   4.187  VB    0.0982    70.66048   10.54307   1.6963
```

**4-Phenylpyrrolidin-2-one (2g)**



**HPLC conditions:** Chiralpak IC, 25% *i*-PrOH in hexane, 1.2 mL/min, 32 ℃, 220 nm



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 24.074 | 13278.8 | 323.3 | 0.6846 | 49.628 | 0.753 |
| 2 | 27.877 | 13477.6 | 273.3 | 0.8219 | 50.372 | 0.613 |
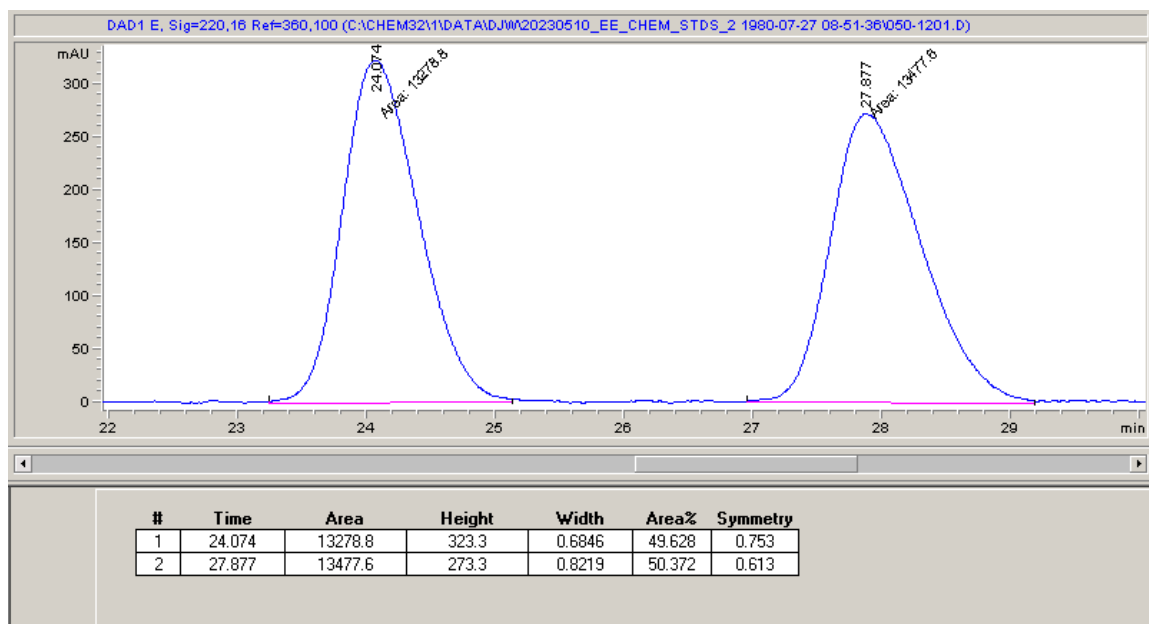
**Enzymatic preparation of 2g with P411-LAS-5247**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 22.932 | 3629.1 | 99.8 | 0.6062 | 52.911 | 0.908 |
| 2 | 26.482 | 3229.8 | 79.4 | 0.678 | 47.089 | 0.908 |
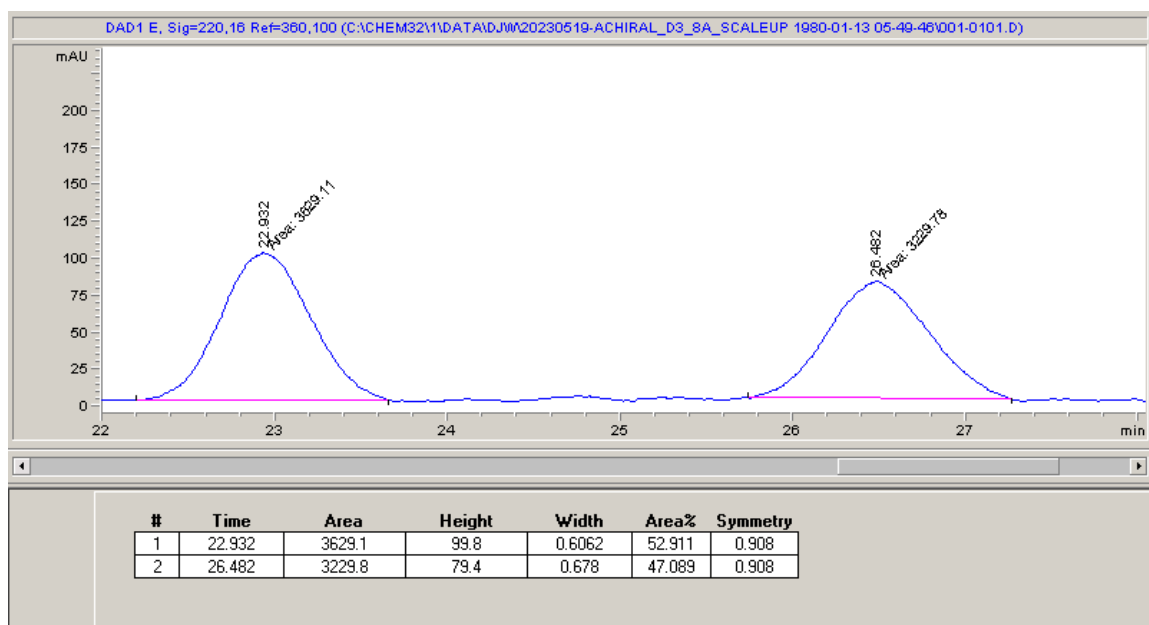
**4-(*p*-Tolyl)tetrahydro-2*H*-pyran-2-one (4a)**



**HPLC conditions:** Chiralpak IC, 15% *i*-PrOH in hexane, 1.2 mL/min, 28 ℃, 220 nm
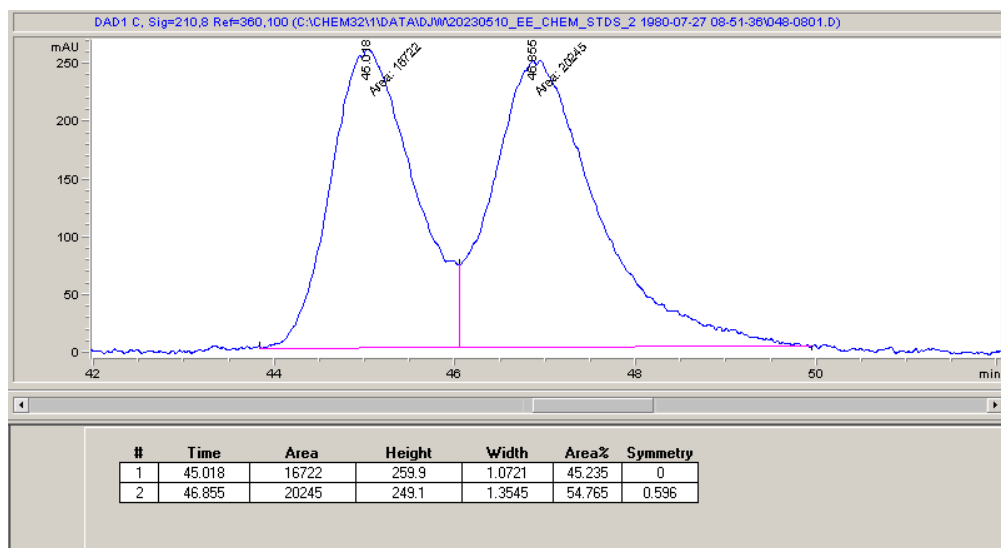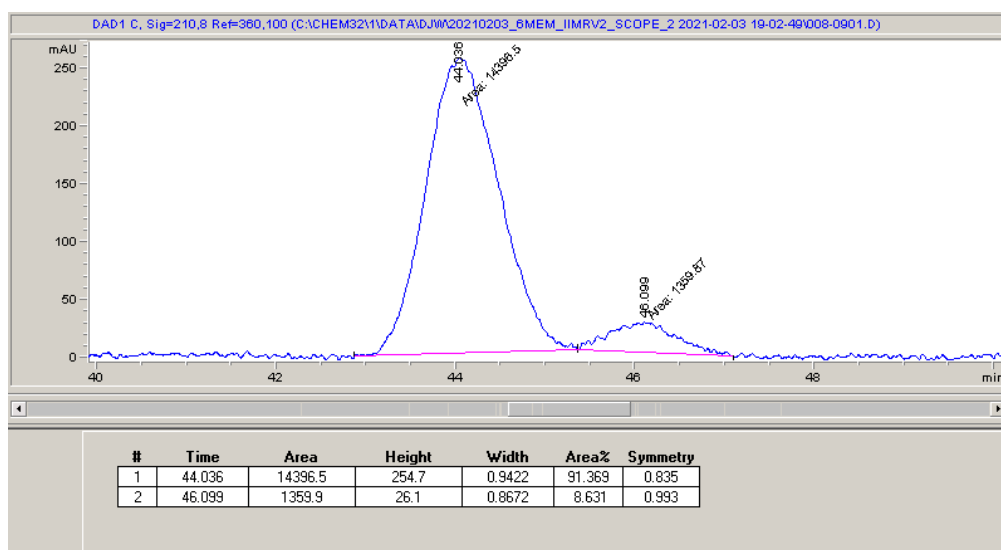


| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 45.018 | 16722 | 259.9 | 1.0721 | 45.235 | 0 |
| 2 | 46.855 | 20245 | 249.1 | 1.3545 | 54.765 | 0.596 |

**Enzymatic preparation of 4b with P411-LAS-5249: 83% e.e.**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 44.036 | 14396.5 | 254.7 | 0.9422 | 91.369 | 0.835 |
| 2 | 46.099 | 1359.9 | 26.1 | 0.8672 | 8.631 | 0.993 |

**HPLC conditions:** Chiralpak IC, 25% *i*-PrOH in hexane, 1.2 mL/min, 28 ℃, 220 nm

DAD1 E, Sig=220,16 Ref=360,100 (C:\CHEM32\1\DATA\DJW\20230810_EE_6PME_FOROLD_2 2023-08-06 17-01-11\001-0101.D)

| # | Time | Area | Height | Width | Area% | Symmetry |
|---|---|---|---|---|---|---|
| 1 | 26.304 | 4972.1 | 100.7 | 0.8232 | 50.944 | 1.035 |
| 2 | 27.254 | 4787.8 | 93 | 0.8579 | 49.056 | 1.027 |
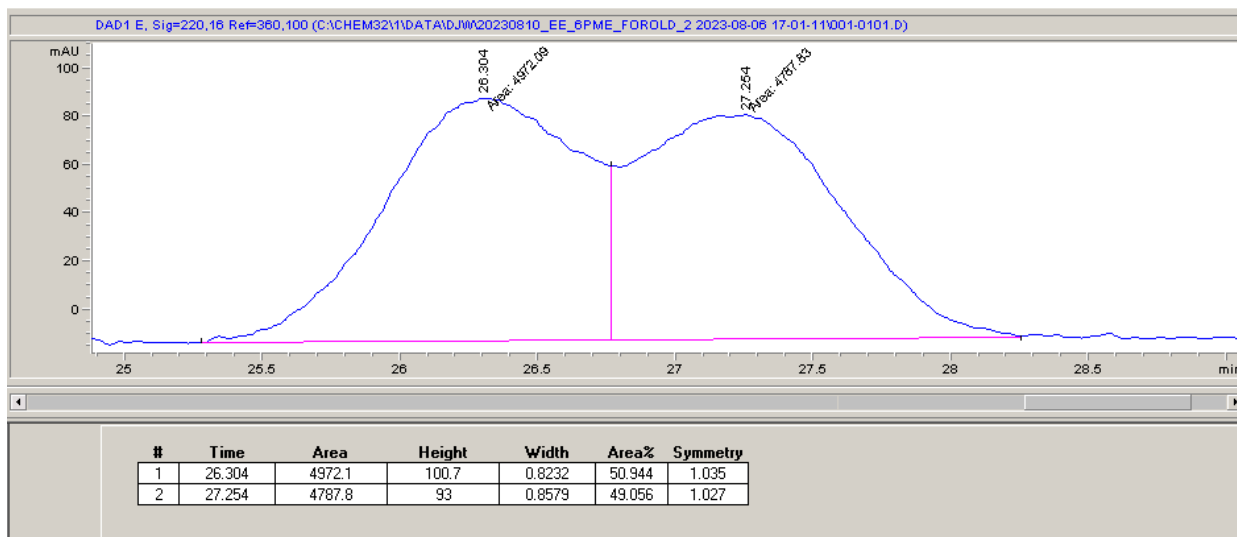
## Enzymatic preparation of 4a with P411-LAS-5244: 46% e.e.

DAD1 E, Sig=220,16 Ref=360,100 (C:\CHEM32\1\DATA\DJW\20200313-6EETHIO_FURAN_PME_IMRV 2020-03-13 15-07-02\073-0301.D)

| # | Time | Area | Height | Width | Area% | Symmetry |
|---|---|---|---|---|---|---|
| 1 | 26.213 | 2070.8 | 55.4 | 0.6225 | 73.452 | 0.95 |
| 2 | 27.126 | 748.5 | 20.6 | 0.6046 | 26.548 | 0.889 |

**4-Phenyltetrahydro-2*H*-pyran-2-one (4b)**



**HPLC conditions:** Chiralpak IC, 25% *i*-PrOH in hexane, 1.2 mL/min, 28 ℃, 220 nm
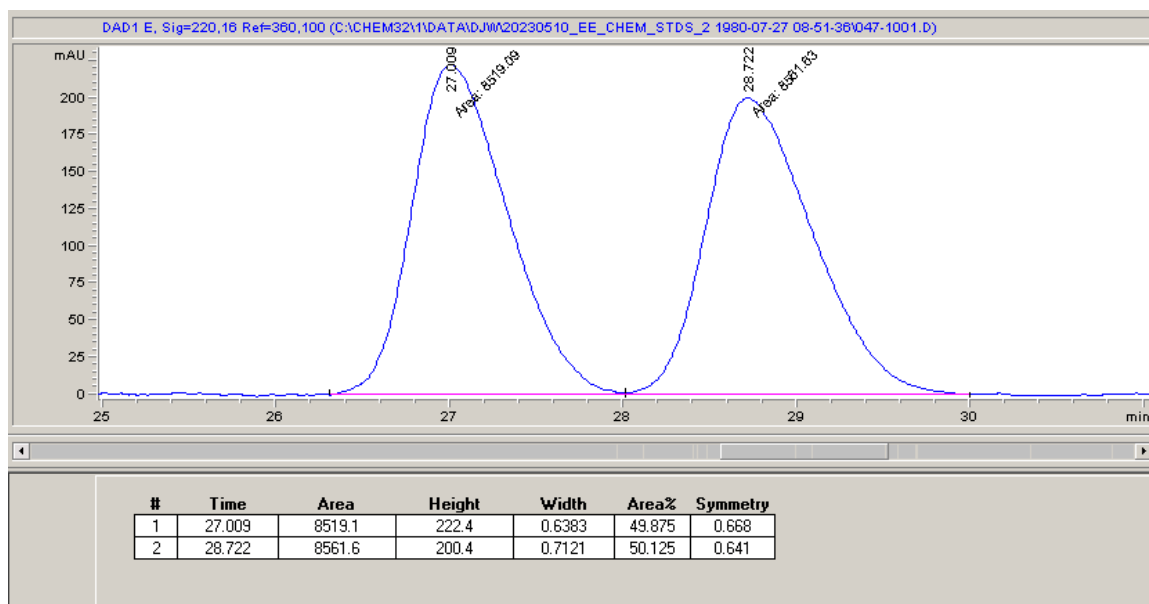


| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 27.009 | 8519.1 | 222.4 | 0.6383 | 49.875 | 0.668 |
| 2 | 28.722 | 8561.6 | 200.4 | 0.7121 | 50.125 | 0.641 |

**Enzymatic preparation of 4b with P411-LAS-5244: 90% e.e.**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 26.87 | 2593.5 | 68.9 | 0.627 | 95.155 | 1.014 |
| 2 | 28.542 | 132.1 | 3.6 | 0.6113 | 4.845 | 0.87 |

**Enzymatic preparation of 4b with P411-LAS-5249: 93% e.e.**



DAD1 E, Sig=220,16 Ref=360,100 (C:\CHEM32\1\DATA\DJW\20210203_6MEM_IIMRV2_SCOPE_2 2021-02-03 19-02-49\002-0201.D)

| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 27.017 | 1238.9 | 34.5 | 0.599 | 96.380 | 0.909 |
| 2 | 28.671 | 46.5 | 1.7 | 0.4554 | 3.620 | 0.751 |
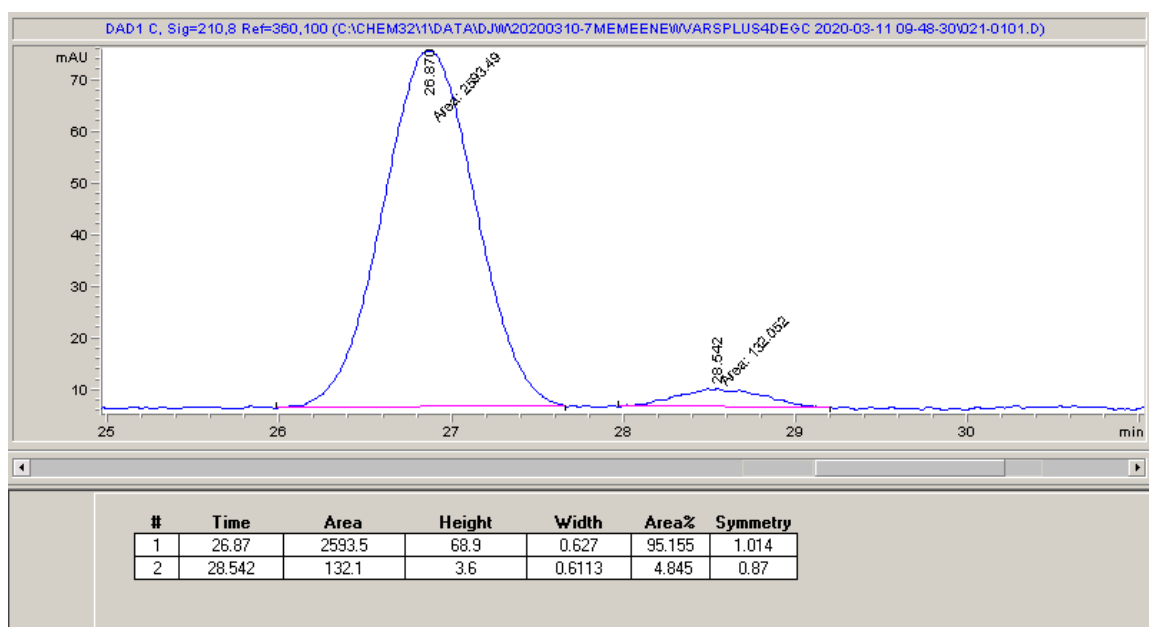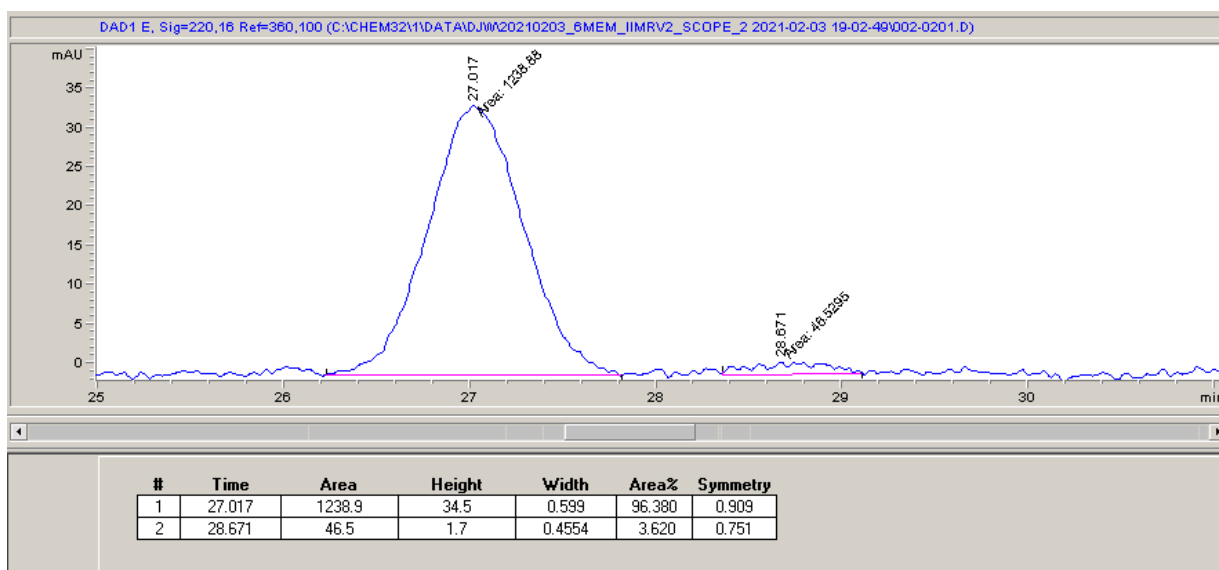
**HPLC conditions:** Chiralpak AS-H, 40% *i*-PrOH in hexane, 0.6 mL/min, 210 nm
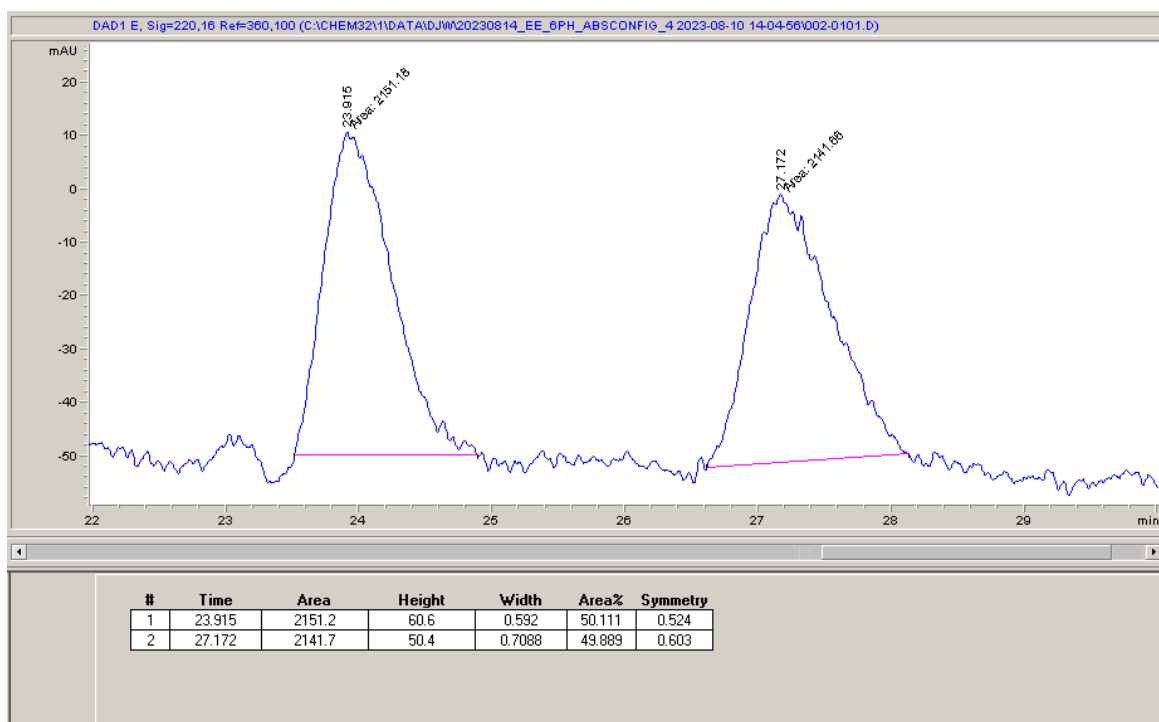


DAD1 E, Sig=220,16 Ref=360,100 (C:\CHEM32\1\DATA\DJW\20230814_EE_6PH_ABSCONFIG_4 2023-08-10 14-04-56\002-0101.D)

| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 23.915 | 2151.2 | 60.6 | 0.592 | 50.111 | 0.524 |
| 2 | 27.172 | 2141.7 | 50.4 | 0.7088 | 49.889 | 0.603 |

**Enzymatic preparation of 2c with P411-LAS-5249:**



DAD1 E, Sig=220,16 Ref=360,100 (C:\CHEM32\1\DATA\DJW\20230814_EE_6PH_ABSCONFIG_5 2023-08-10 15-35-26\001-0201.D)

| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
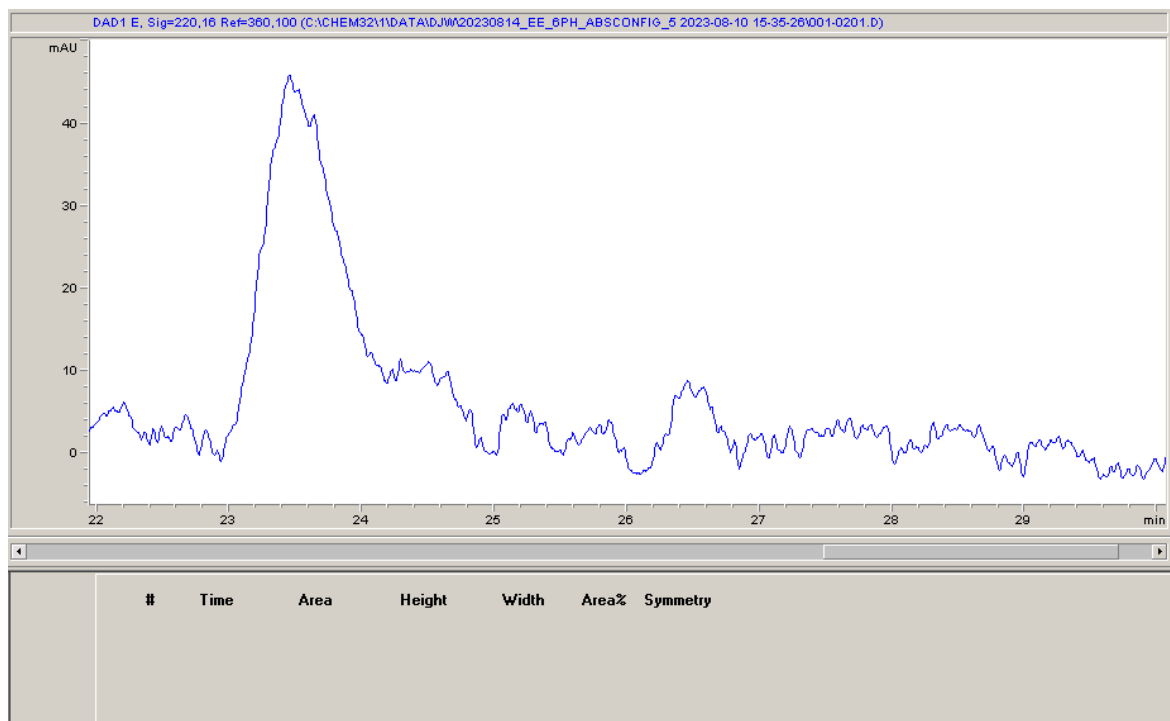
**4-(4-Methoxyphenyl)tetrahydro-2*H*-pyran-2-one (4c)**



**HPLC conditions:** Chiralpak IC, 25% *i*-PrOH in hexane, 1.2 mL/min, 32 ℃, 220 nm



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 40.78 | 14973.6 | 255.9 | 0.9752 | 49.977 | 0.733 |
| 2 | 49.779 | 14987.5 | 207.7 | 1.2027 | 50.023 | 0.773 |
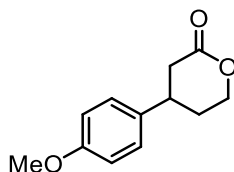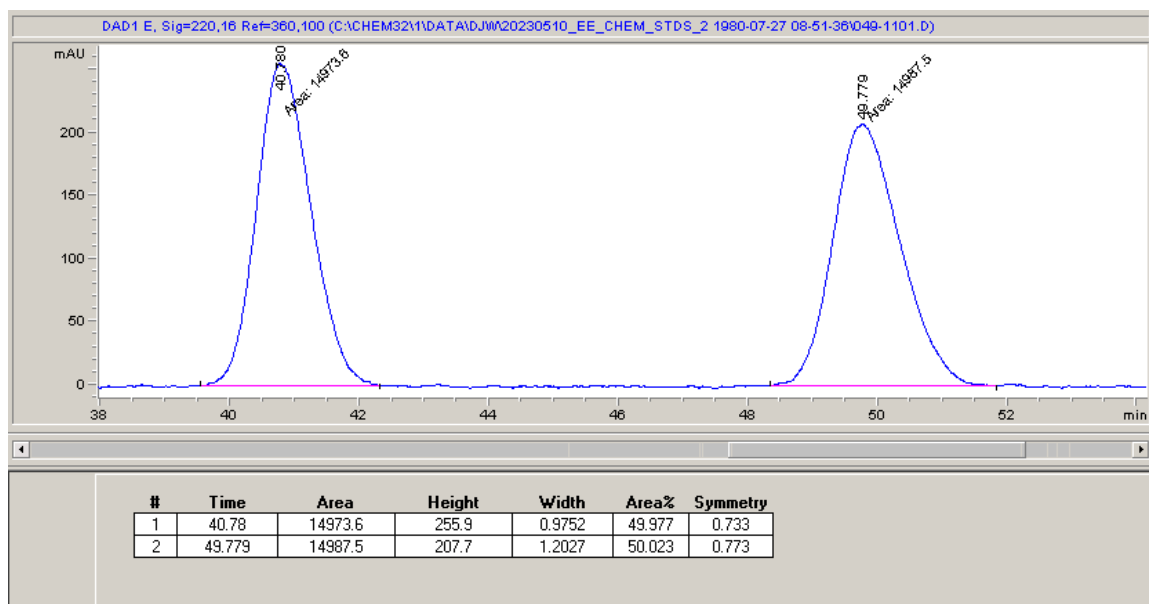
**Enzymatic preparation of 4c with P411-LAS-5249: 75% e.e.**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 40.051 | 6391 | 112.9 | 0.9433 | 87.740 | 0.806 |
| 2 | 48.968 | 893 | 14 | 1.0639 | 12.260 | 0.804 |
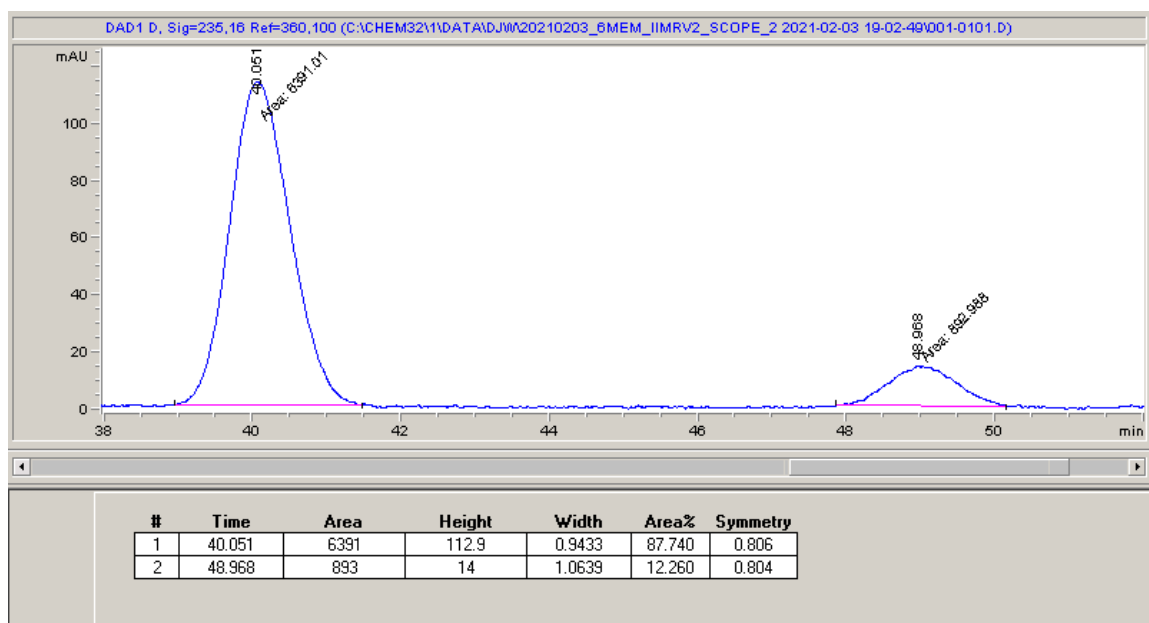
**4-(4-Fluorophenyl)tetrahydro-2*H*-pyran-2-one (4d)**
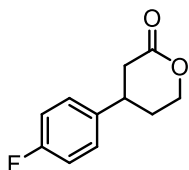
**HPLC conditions:** Chiralpak IC, 25% *i*-PrOH in hexane, 1.2 mL/min, 32 ℃, 220 nm
**4d isolated from P411-LAS-5244-catalyzed transformation: 65% e.e.**
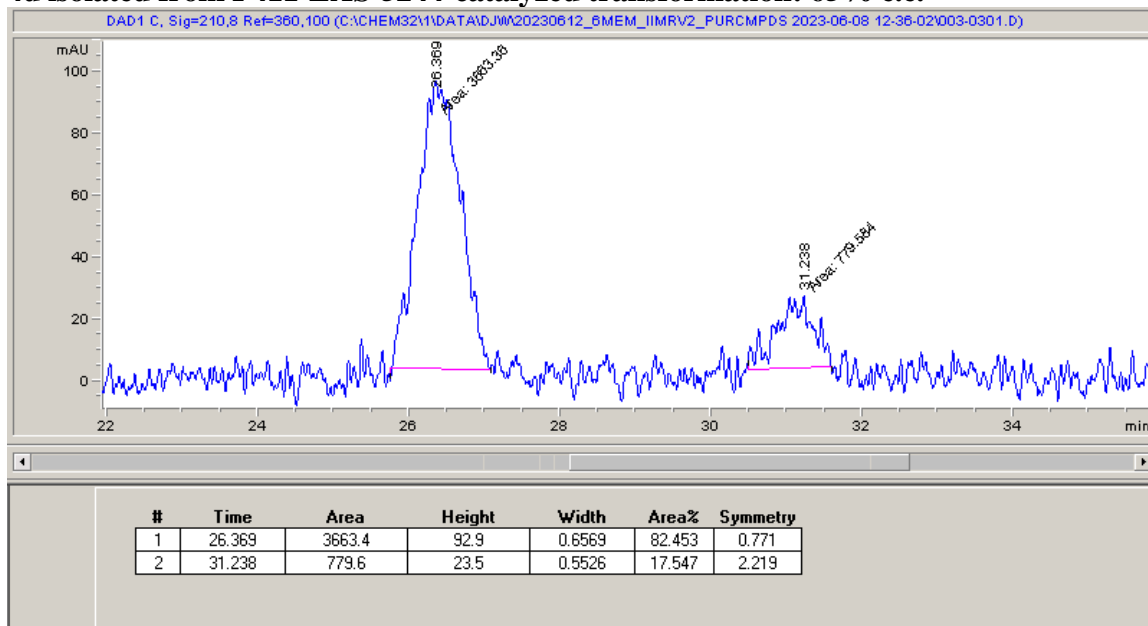
DAD1 C, Sig=210,8 Ref=360,100 (C:\CHEM32\1\DATA\DJW\20230612_6MEM_IIMRV2_PURCMPDS 2023-06-08 12-36-02\003-0301.D)

| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 26.369 | 3663.4 | 92.9 | 0.6569 | 82.453 | 0.771 |
| 2 | 31.238 | 779.6 | 23.5 | 0.5526 | 17.547 | 2.219 |

**Enzymatic preparation of 4d with P411-LAS-5249: 80% e.e.**

DAD1 C, Sig=210,8 Ref=360,100 (C:\CHEM32\1\DATA\DJW\20210203_6MEM_IIMRV2_SCOPE_2 2021-02-03 19-02-49\003-0301.D)

| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 26.121 | 3590.9 | 101.1 | 0.5918 | 90.160 | 0.847 |
| 2 | 30.819 | 391.9 | 9.9 | 0.6623 | 9.840 | 2.471 |
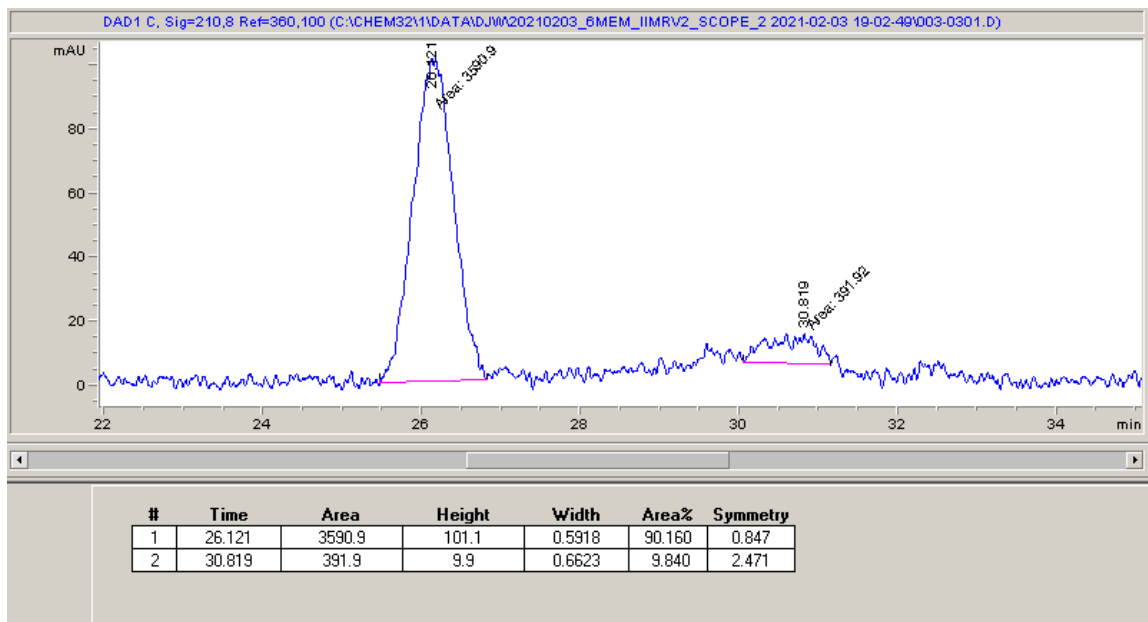
**4-(Thiophen-2-yl)tetrahydro-2H-pyran-2-one (4f)**

**HPLC conditions:** Chiralpak IC, 25% *i*-PrOH in hexane, 1.2 mL/min, 32 ℃, 220 nm
**4f isolated from P411-LAS-5244-catalyzed transformation: 76% e.e.**



DAD1 D, Sig=235,16 Ref=360,100 (C:\CHEM32\1\DATA\DJW\20230612_6MEM_IIMRV2_PURCMPDS 2023-06-08 12-36-02\001-0101.D)

| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 28.412 | 4878.7 | 133 | 0.6113 | 12.245 | 1.059 |
| 2 | 29.981 | 34962.3 | 816.2 | 0.7139 | 87.755 | 0.633 |

**Enzymatic preparation of 4d with P411-LAS-5249: 86% e.e.**



DAD1 D, Sig=235,16 Ref=360,100 (C:\CHEM32\1\DATA\DJW\20210203_6MEM_IIMRV2_SCOPE_2 2021-02-03 19-02-49\006-0601.D)

| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 28.051 | 401.6 | 12.1 | 0.5549 | 6.968 | 0.96 |
| 2 | 29.886 | 5361.3 | 134.9 | 0.6625 | 93.032 | 0.998 |
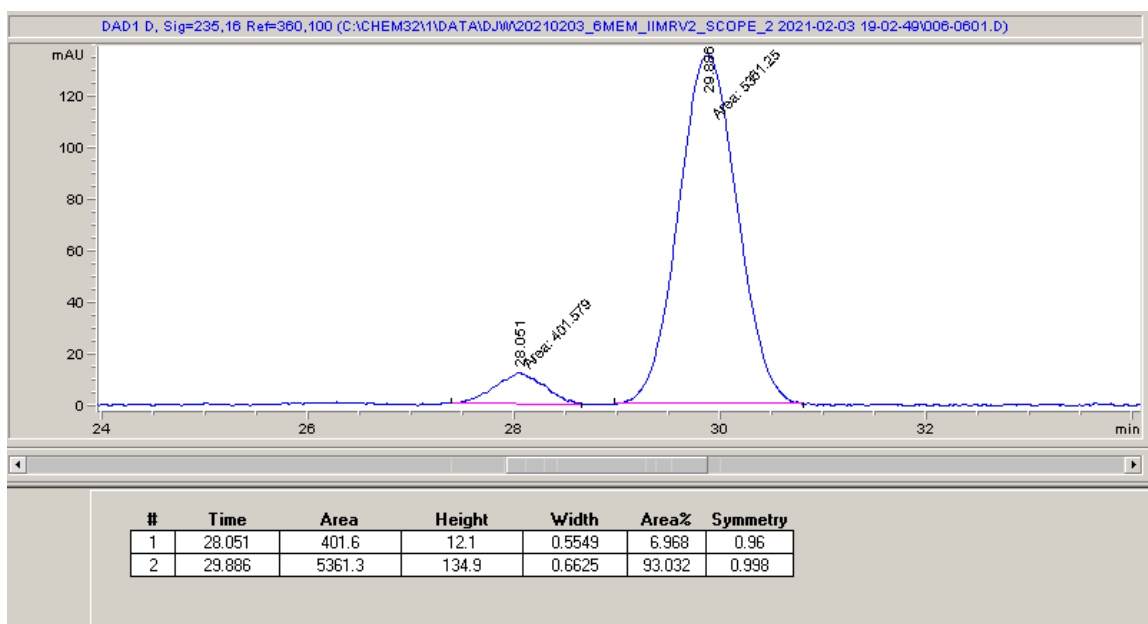
# 4-(4-Methoxyphenyl)oxepan-2-one (6a)



**HPLC conditions:** Chiralpak IC, 25% *i*-PrOH in hexane, 1.2 mL/min, 32 ℃, 220 nm

**6a isolated from P411-C10 variant-catalyzed transformation: 32% e.e.**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 28.709 | 30163.8 | 699.9 | 0.7183 | 33.850 | 0.754 |
| 2 | 36.576 | 58946.8 | 945.6 | 1.0389 | 66.150 | 0.467 |
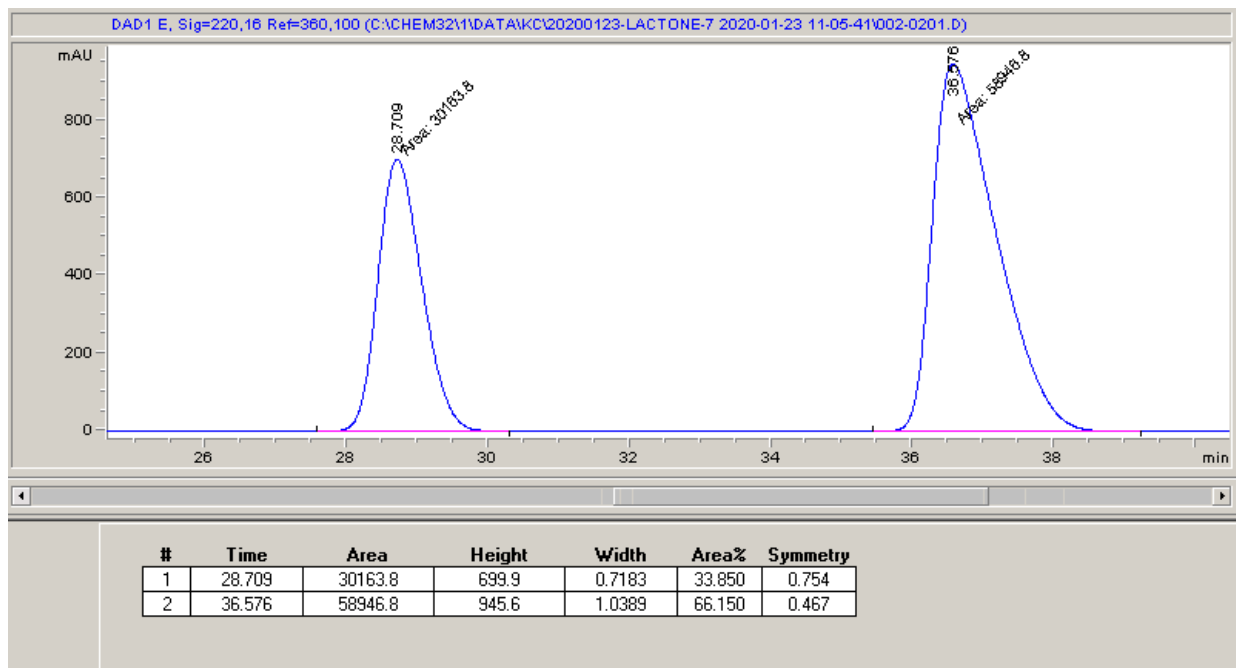
**Enzymatic preparation of 6a with P411-LAS-5259: 51% e.e.**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 28.649 | 610.6 | 14.2 | 0.7154 | 24.214 | 0.96 |
| 2 | 37.01 | 1911 | 34.7 | 0.919 | 75.786 | 0.981 |

**Enzymatic preparation of 6a with P411-LAS-5260: 96% e.e.**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 28.588 | 16.8 | 5.4E-1 | 0.5194 | 2.195 | 0.589 |
| 2 | 37.085 | 750.1 | 13.3 | 0.9386 | 97.805 | 1.009 |

**Enzymatic preparation of 6a with P411-LAS-5264: 95% e.e.**



| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 28.656 | 179 | 3.9 | 0.7647 | 2.606 | 0.959 |
| 2 | 36.938 | 6691.5 | 124.1 | 0.8987 | 97.394 | 0.79 |

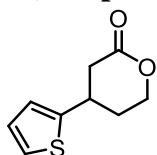**(*E*)-4-(4-Methoxystyryl)oxepan-2-one (6b)**
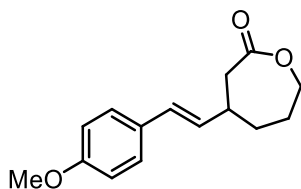


**HPLC conditions:** Chiralpak IC, 35% *i*-PrOH in hexane, 1.2 mL/min, 32 ºC, 220 nm

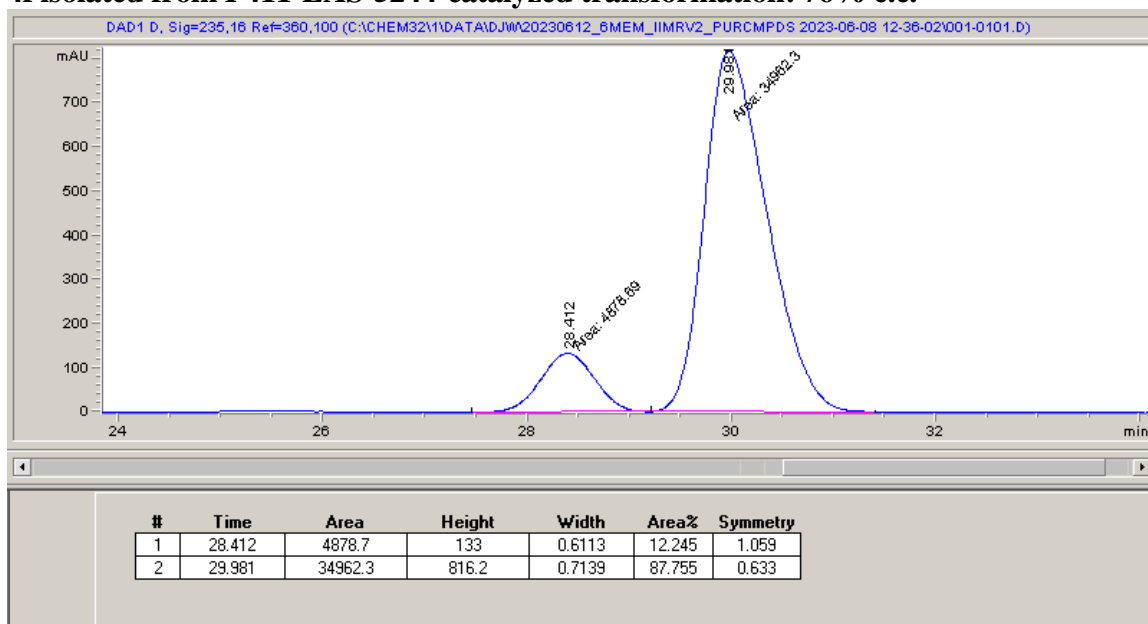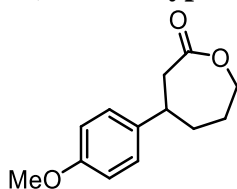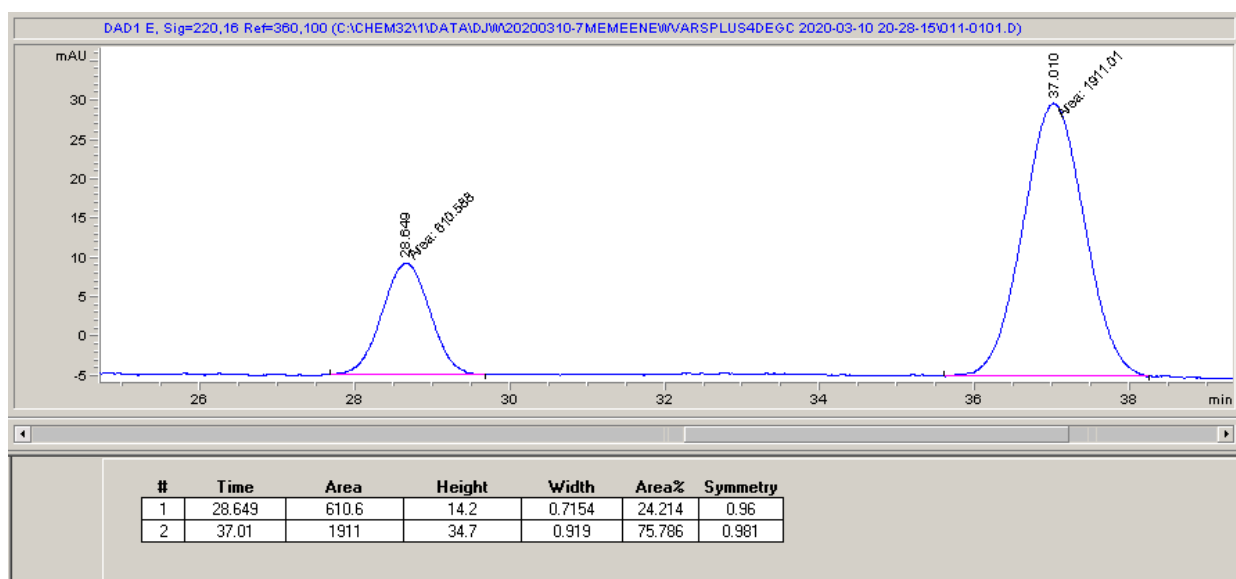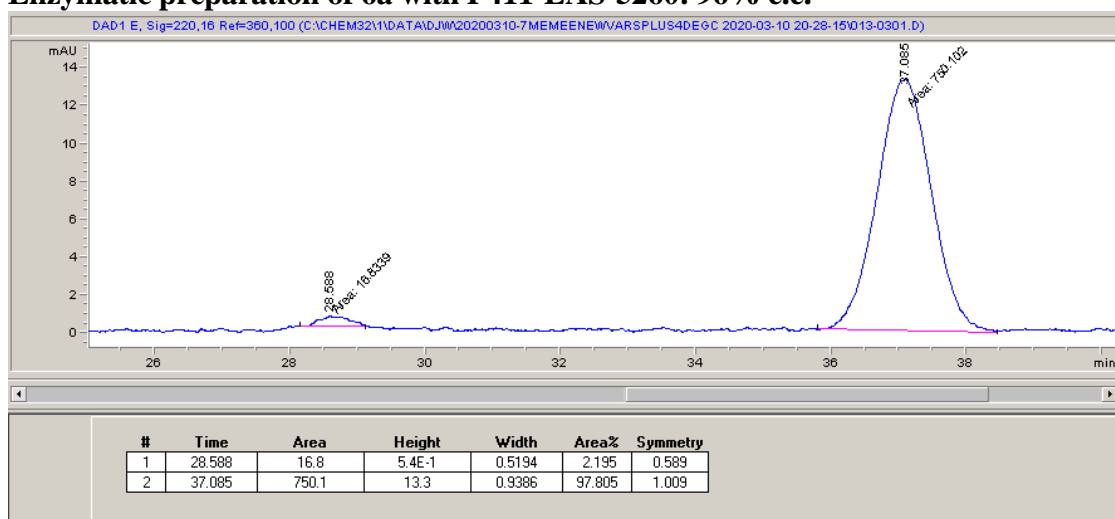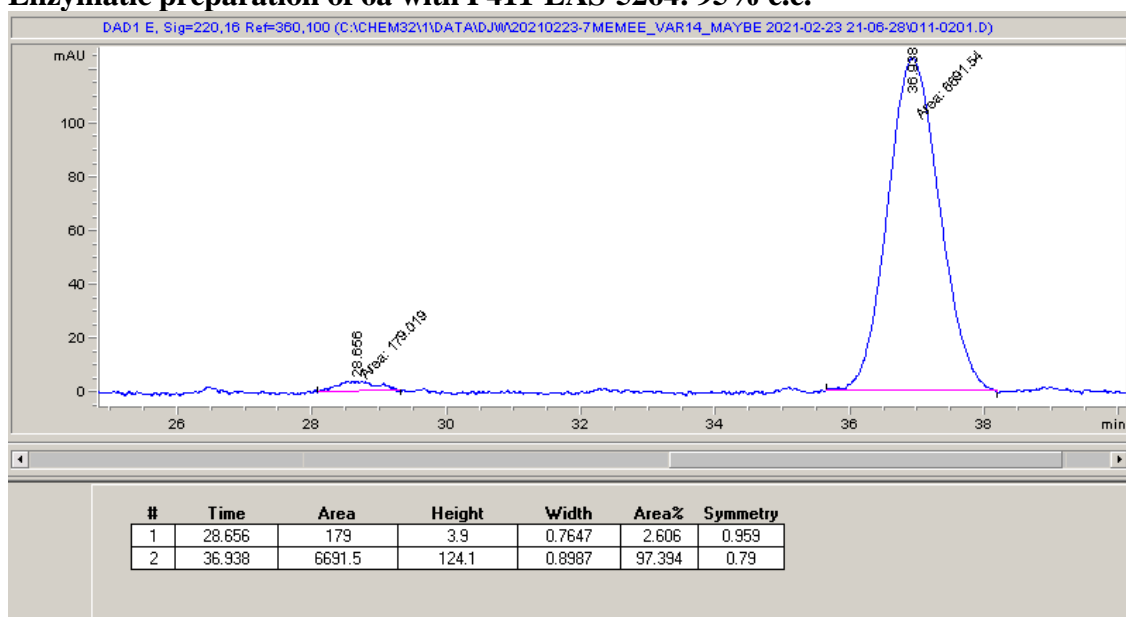**6b isolated from P411-LAS-5264-catalyzed transformation: 90% e.e.**
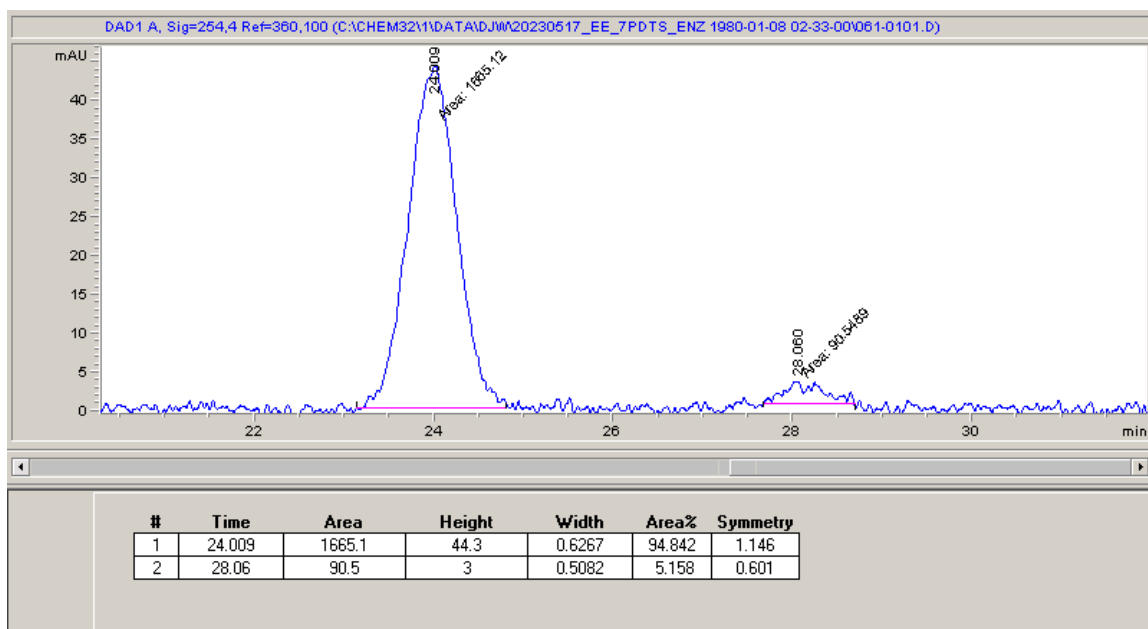


| # | Time | Area | Height | Width | Area% | Symmetry |
|---|------|------|--------|-------|-------|----------|
| 1 | 24.009 | 1665.1 | 44.3 | 0.6267 | 94.842 | 1.146 |
| 2 | 28.06 | 90.5 | 3 | 0.5082 | 5.158 | 0.601 |

**X-ray Crystallography**

Low-temperature diffraction data ($\phi$-and $\omega$-scans) were collected on a Bruker AXS D8 VENTURE KAPPA or Bruker APEX-II diffractometer coupled to a PHOTON II CPAD detector with Cu $K_\alpha$ radiation ($\lambda = 1.54178$ Å) from an I$\mu$S micro-source for the structure of compounds. The structure was solved by direct methods using SHELXS[19] and refined against $F^2$ on all data by full-matrix least squares with SHELXL-2019[20] using established refinement techniques.[21] All non-hydrogen atoms were refined anisotropically. Unless otherwise noted, all hydrogen atoms were included into the model at geometrically calculated positions and refined using a riding model. The isotropic displacement parameters of all hydrogen atoms were fixed to 1.2 times the $U$ value of the atoms they are linked to (1.5 times for methyl groups).

**Figure 4-S3.** Displacement ellipsoid plot for **6a** plotted at 50% probability. Single crystals were obtained from slow evaporation of **6a** dissolved in $CHCl_3$. Compound **6a** crystallizes in the orthorhombic space group $P2_12_12_1$ with one molecule in the asymmetric unit.

**Table 4-S6.** X-ray experimental details of **6a** (CCDC 2288858).

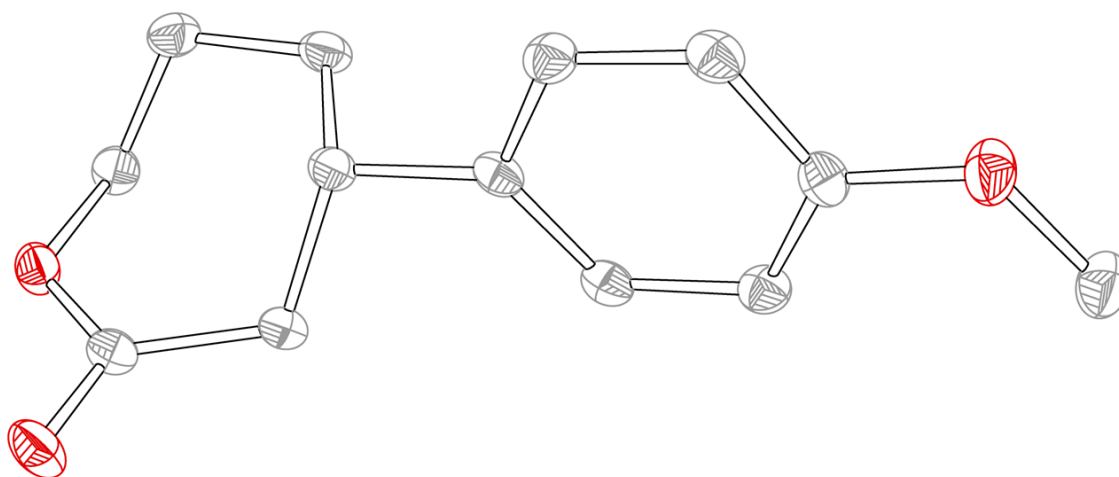| Crystal data | |
| --- | --- |
| Chemical formula | $C_{13}H_{16}O_3$ |
| $M_r$ | 220.26 |
| Crystal system, space group | Orthorhombic, $P2_12_12_1$ |
| Temperature (K) | 100 |
| $a$, $b$, $c$ (Å) | 6.4683 (7), 9.5514 (11), 18.433 (2) |
| $V$ (Å$^3$) | 1138.8 (2) |
| $Z$ | 4 |
| Radiation type | Cu $K\alpha$ |
| $\mu$ (mm$^{-1}$) | 0.74 |
| Crystal size (mm) | $0.25 \times 0.2 \times 0.1$ |
| **Data collection** | |
| Diffractometer | Bruker *APEX*-II CCD |
| Absorption correction | |
| $T_{min}$, $T_{max}$ | 0.662, 0.754 |
| No. of measured, independent and observed [$I > 2\sigma(I)$] reflections | 14831, 2298, 2272 |
| $R_{int}$ | 0.036 |
| $(\sin \theta/\lambda)_{max}$ (Å$^{-1}$) | 0.625 |
| **Refinement** | |
| $R[F^2 > 2\sigma(F^2)]$, $wR(F^2)$, $S$ | 0.026, 0.066, 1.05 |
| No. of reflections | 2298 |
| No. of parameters | 146 |
| H-atom treatment | H-atom parameters constrained |
| $\Gamma_{max}$, $\Gamma_{min}$ (e Å$^{-3}$) | 0.14, -0.17 |
| Absolute structure | Flack x determined using 932 quotients [(I+)-(I-)]/[(I+)+(I-)]  (Parsons, Flack and Wagner, Acta Cryst. B69 (2013) 249-259). |
| Absolute structure parameter | -0.09 (6) |

**Figure 4-S4.** Displacement ellipsoid plot for **8a** plotted at 50% probability. Single crystals were obtained from slow evaporation of **8a** dissolved in CHCl$_3$. Compound **8a** crystallizes in the orthorhombic space group $P2_12_12_1$ with one molecule in the asymmetric unit.

**Table 4-S7.** X-ray experimental details of **8a** (CCDC 2287785).

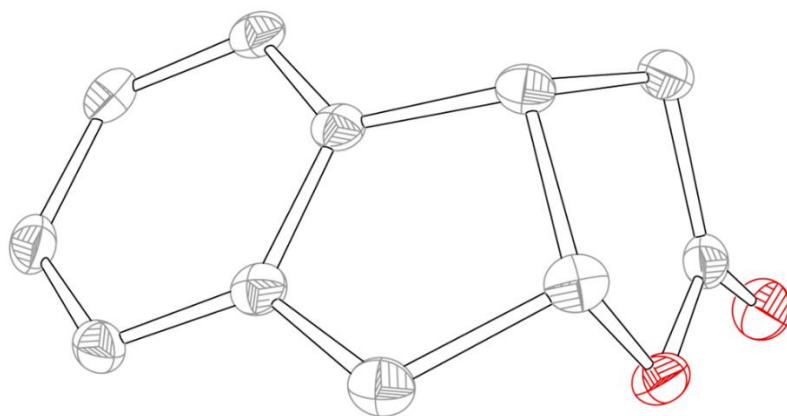| Crystal data | |
|---|---|
| Chemical formula | $C_{11}H_{10}O_2$ |
| $M_r$ | 179.19 |
| Crystal system, space group | Orthorhombic, $P2_12_12_1$ |
| Temperature (K) | 100 |
| $a$, $b$, $c$ (Å) | 4.7240 (4), 12.2962 (14), 14.3829 (12) |
| $V$ (Å$^3$) | 835.46 (14) |
| $Z$ | 4 |
| Radiation type | Cu $K\alpha$ |
| $\mu$ (mm$^{-1}$) | 0.768 |
| Crystal size (mm) | $0.15 \times 0.1 \times 0.1$ |
| **Data collection** | |
| Diffractometer | Bruker D8 VENTURE Kappa Duo PHOTON II CPAD |
| Absorption correction | Multi-scan SADABS2016/2 (Sheldrick, 2014) |
| $T_{min}$, $T_{max}$ | 0.6762, 0.7538 |
| No. of measured, independent and observed [$I > 2\sigma(I)$] reflections | 15040, 1711, 1670 |
| $R_{int}$ | 0.0367 |
| $(\sin \theta/\lambda)_{max}$ (Å$^{-1}$) | 0.625 |
| **Refinement** | |
| $R[F^2 > 2\sigma(F^2)]$, $wR(F^2)$, $S$ | 0.0261, 0.0668, 1.079 |
| No. of reflections | 1711 |
| No. of parameters | 118 |
| H-atom treatment | H-atom parameters constrained |
| $\Gamma_{max}$, $\Gamma_{min}$ (e Å$^{-3}$) | 0.15, -0.15 |
| Absolute structure | Flack x determined using 667 quotients [(I+)-(I-)]/[(I+)+(I-)]   (Parsons, Flack and Wagner, Acta Cryst. B69 (2013) 249-259). |
| Absolute structure parameter | 0.05 (7) |
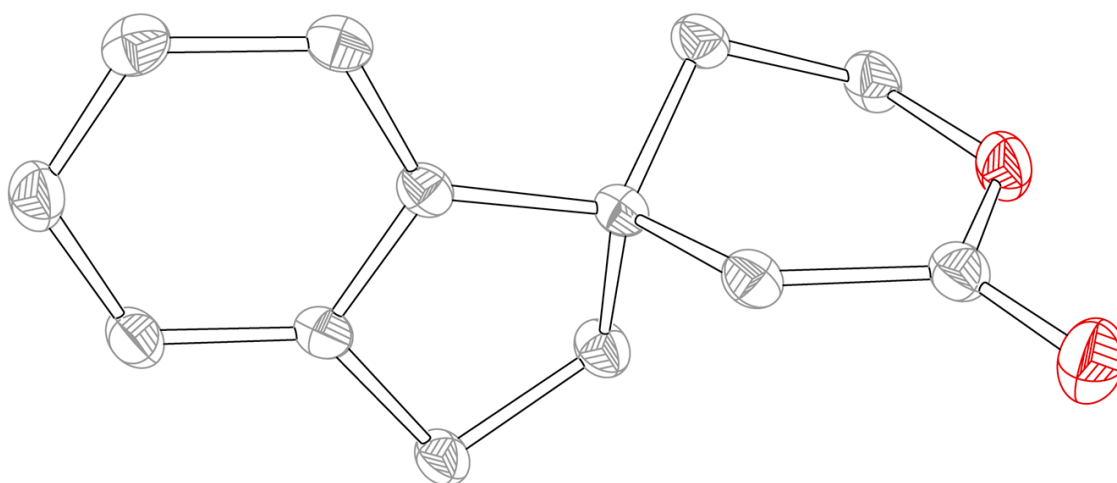
**Figure 4-S5.** Displacement ellipsoid plot for **8f** plotted at 50% probability. Single crystals were obtained from slow evaporation of **8f** dissolved in CHCl$_3$. Compound **8f** crystallizes in the monoclinic space group $P2_1$/c with one molecule in the asymmetric unit.

**Table 4-S8.** X-ray experimental details of **8f** (CCDC 2287784).

| Crystal data | |
| --- | --- |
| Chemical formula | $C_{13}H_{14}O_2$ |
| $M_r$ | 202.24 |
| Crystal system, space group | Monoclinic, $P2_1/c$ |
| Temperature (K) | 100 |
| $a$, $b$, $c$ (Å) | 13.0013 (8), 6.6891 (6), 12.8888 (10) |
| $\alpha$, $\beta$, $\gamma$ (°) | 90, 113.787 (3), 90 |
| $V$ (Å$^3$) | 1025.68 (14) |
| $Z$ | 4 |
| Radiation type | Cu $K\alpha$ |
| $\mu$ (mm$^{-1}$) | 0.696 |
| Crystal size (mm) | $0.2 \times 0.15 \times 0.05$ |
| **Data collection** | |
| Diffractometer | Bruker D8 VENTURE Kappa Duo PHOTON II CPAD |
| Absorption correction | Multi-scan SADABS2016/2 (Sheldrick, 2014) |
| $T_{min}$, $T_{max}$ | 0.5155, 0.7538 |
| No. of measured, independent and observed [$I > 2\sigma(I)$] reflections | 13396, 2098, 1884 |
| $R_{int}$ | 0.0632 |
| $(\sin \theta/\lambda)_{max}$ (Å$^{-1}$) | 0.625 |
| **Refinement** | |
| $R[F^2 > 2\sigma(F^2)]$, $wR(F^2)$, $S$ | 0.0449, 0.1248, 1.07 |
| No. of reflections | 2098 |
| No. of parameters | 136 |
| H-atom treatment | H-atom parameters constrained |
| $\Gamma_{max}$, $\Gamma_{min}$ (e Å$^{-3}$) | 0.41, -0.23 |

^{1}H, ^{13}C, and ^{19}F NMR spectra of the compounds can be found in the Supporting Information of the published paper.

## 4.7 References for Chapter 4

References for Sections 4.1-4.6:

1. Sartori, S. K., Diaz, M. A. N. & Diaz-Muñoz, G. Lactones: Classification, Synthesis, Biological Activities, and Industrial Applications. *Tetrahedron* **84**, 132001 (2021).

2. Hur, J., Jang, J. & Sim, J. A Review of the Pharmacological Activities and Recent Synthetic Advances of gamma-Butyrolactones. *Int. J. Mol. Sci*. **22**, 2769 (2021).

3. Delgove, M. A. F., Elford, M. T., Bernaerts, K. V. & Wildeman, S. M. A. Toward Upscaled Biocatalytic Preparation of Lactone Building Blocks for Polymer Applications. *Org. Process Res. Dev*. **22**, 803–812 (2018).

4. Hollmann, F., Kara, S., Opperman, D. J. & Wang, Y. Biocatalytic Synthesis of Lactones and Lactams. *Asian J. Chem*. **13**, 3601–3610 (2018).

5. Doyle, M. P. & Forbes, D. C. Recent Advances in Asymmetric Catalytic Metal Carbene Transformations. *Chem. Rev*. **98**, 911–935 (1998).

6. Davies, H. M. & Manning, J. R. Catalytic C–H Functionalization by Metal Carbenoid and Nitrenoid Insertion. *Nature* **451**, 417–424 (2008).

7. Davies, H. M. L. & Beckwith, R. E. J. Catalytic Enantioselective C−H Activation by Means of Metal−Carbenoid-Induced C−H Insertion. *Chem. Rev*. **103**, 2861−2903 (2003).

8. Santiago, J. V. & Machado, A. H. Enantioselective Carbenoid Insertion Into C(sp3)-H Bonds. Beilstein *J. Org. Chem*. **12**, 882–902 (2016).

9. Doyle, M. P., Duffy, R., Ratnikov, M. & Zhou, L. Catalytic Carbene Insertion into C−H Bonds. *Chem. Rev*. **110**, 704–724 (2010).

10. Doyle, M. P. & Hu, W. Enantioselective Carbon–Hydrogen Insertion is an Effective and Efficient Methodology for the Synthesis of (R)-(-)-Baclofen. *Chirality* **14**, 169–172 (2002).

11. Doyle, M. P., Oeveren, A. v., Westrum, L. J., Protopopova, M. N. & Thomas W. Clayton, J. Asymmetric Synthesis of Lactones with High Enantioselectivity by Intramolecular Carbon–Hydrogen Insertion Reactions of Alkyl Diazoacetates

Catalyzed by Chiral Rhodium(II) Carboxamides. *J. Am. Chem. Soc*. **113**, 8982–8984 (1991).

12. Doyle, M. P., Kalinin, A. V. & Ene, D. G. Chiral Catalyst Controlled Diastereoselection and Regioselection in Intramolecular Carbon–Hydrogen Insertion Reactions of Diazoacetates. *J. Am. Chem. Soc*. **118**, 8837–8846 (1996).

13. Doyle, M. P. & Protopopova, M. N. Macrocyclic Lactones from Dirhodium(II)-Catalyzed Intramolecular Cyclopropanation and Carbon-Hydrogen Insertion. *J. Am. Chem. Soc*. **117**, 7281–7282 (1995).

14. Fürst, M. J. L. J., Gran-Scheuch, A., Aalbers, F. S. & Fraaije, M. W. Baeyer–Villiger Monooxygenases: Tunable Oxidative Biocatalysts. *ACS Catal*. **9**, 11207–11241 (2019).

15. Dong, J. et al. Biocatalytic Oxidation Reactions: A Chemist's Perspective. *Angew. Chem., Int. Ed*. **57**, 9238-9261 (2018).

16. Manning, J. et al. Regio- and Enantio-selective Chemo-enzymatic C–H Lactonization of Decanoic Acid to (S)-delta-Decalactone. *Angew. Chem., Int. Ed*. **58**, 5668–5671 (2019).

17. Borowiecki, P. et al. Biocatalytic Asymmetric Reduction of γ-Keto Esters to Access Optically Active γ-Aryl-γ-butyrolactones. *Adv. Synth. Catal*. **362**, 2012–2029 (2020).

18. Poulos, T. L. Cytochrome P450 Flexibility. *Proc. Natl. Acad. Sci*. **100**, 13121–13122 (2003).

19. Brandenberg, O. F., Fasan, R. & Arnold, F. H. Exploiting and Engineering Hemoproteins for Abiological Carbene and Nitrene Transfer Reactions. *Curr. Opin. Biotechnol.* **47**, 102–111 (2017).

20. Kaur, P. & Tyagi, V. Recent Advances in Iron-Catalyzed Chemical and Enzymatic Carbene-Transfer Reactions. *Adv. Synth. Catal*. **363**, 877–905 (2021).

21. Yang, Y. & Arnold, F. H. Navigating the Unnatural Reaction Space: Directed Evolution of Heme Proteins for Selective Carbene and Nitrene Transfer. *Acc. Chem. Res*. **54**, 1209–1225 (2021).

22. Chen, K., Huang, X., Kan, S. B. J., Zhang, R. K. & Arnold, F. H. Enzymatic Construction of Highly Strained Carbocycles. *Science* **360**, 71–75 (2018).

23. Illuminati, G., Mandolini, L. & Masci, B. Ring-Closure Reactions. 9. Kinetics of Ring Formation from omicron- omega-Bromoalkoxy Phenoxides and omicron-omega-Bromoalkyl Phenoxides in the Range of 11- To 24-Membered Rings. A Comparison With Related Cyclization Series. *J. Am. Chem. Soc*. **99**, 6308–6312 (1977).

24. Casadei, M. A., Galli, C. & Mandolini, L. Ring-Closure Reactions. 22. Kinetics of Cyclization of Diethyl (omega-Bromoalkyl) Malonates in the Range of 4- To 21-Membered Rings. Role of Ring Strain. *J. Am. Chem. Soc*. **106**, 1051–1056 (1984).

25. Chen, K. & Arnold, F. H. Engineering Cytochrome P450s for Enantioselective Cyclopropenation of Internal Alkynes. *J. Am. Chem. Soc*. **142**, 6891–6895 (2020).

26. Zhou, A. Z., Chen, K. & Arnold, F. H. Enzymatic Lactone-Carbene C–H Insertion to Build Contiguous Chiral Centers. *ACS Catal*. **10**, 5393–5398 (2020).

27. Roccatano, D. Structure, Dynamics, and Function of the Monooxygenase P450BM3: Insights From Computer Simulations Studies. *J. Condens. Matter Phys*. **27**, 273102 (2015).

28. Chang, Y. T. & Loew, G. H. Molecular Dynamics Simulations of P450BM3 Examination of Substrate-Induced Conformational Change. *J. Biomol. Struct. Dyn*. **16**, 1189–1203 (1999).

29. Zhou, Y. et al. Facile Access to Chiral Gamma-Butyrolactones via Rhodium-Catalysed Asymmetric Hydrogenation of Gamma-Butenolides and Gamma-Hydroxybutenolides. *Chem. Sci*. **14**, 4888-4892 (2023).

30. Bode, J. W., Doyle, M. P., Protopopova, M. N. & Zhou, Q.-L. Intramolecular Regioselective Insertion into Unactivated Prochiral Carbon−Hydrogen Bonds with Diazoacetates of Primary Alcohols Catalyzed by Chiral Dirhodium(II) Carboxamidates. Highly Enantioselective Total Synthesis of Natural Lignan Lactones. *J. Org. Chem*. **61**, 9146–9155 (1996).

31. Doyle, M. P. et al. Optimization of Enantiocontrol for Carbon–Hydrogen Insertion with Chiral Dirhodium(II) Carboxamidates. Synthesis of Natural Dibenzylbutyrolactone Lignans from 3-Aryl-1-propyl Diazoacetates in High Optical Purity. *J. Org. Chem*. **60**, 6654–6655 (1995).

32. Wang, Y., Hu, X. & Du, H. Vicinal-Diamine-Based Chiral Chain Dienes as Ligands for Rhodium(I)-Catalyzed Highly Enantioselective Conjugated Additions. *Org. Lett*. **12**, 5482–5485 (2010).

33. Knight, A. M. et al. Diverse Engineered Heme Proteins Enable Stereodivergent Cyclopropanation of Unactivated Alkenes. *ACS Cent. Sci*. **4**, 372–377 (2018).

34. Brandenberg, O. F. et al. Stereoselective Enzymatic Synthesis of Heteroatom-Substituted Cyclopropanes. *ACS Catal*. **8**, 2629–2634 (2018).

35. Dehli, J. R. & Gotor, V. Parallel Kinetic Resolution of Racemic Mixtures: A New Strategy for the Preparation of Enantiopure Compounds? *Chem. Soc. Rev*. **31**, 365–370 (2002).

36. Eames, J. Parallel Kinetic Resolutions. *Angew. Chem., Int. Ed*. **39**, 885–888 (2000).

37. Mao, R. et al. Enantio- and Diastereoenriched Enzymatic Synthesis of 1,2,3-Polysubstituted Cyclopropanes from (Z/E)-Trisubstituted Enol Acetates. Preprint at https://doi.org/10.21203/rs.3.rs-2802333/v1 (2023).

References for Section 4.7:

1. Gibson, D. G. et al. Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases. *Nat. Methods* **6**, 343-345 (2009).

2. Kille, S. et al. Reducing Codon Redundancy and Screening Effort of Combinatorial Protein Libraries Created by Saturation Mutagenesis. *ACS Synth. Biol*. **2**, 83–92 (2013).

3. Barr, I. & Guo, F. Pyridine Hemochromagen Assay for Determining the Concentration of Heme in Purified Protein Solutions. *Bio. Protoc*. **5**, e1594 (2015).

4. Zhou, A. Z., Chen, K. & Arnold, F. H. Enzymatic Lactone-Carbene C–H Insertion to Build Contiguous Chiral Centers. *ACS Catal*. **10**, 5393-5398 (2020).

5. Chen, K. & Arnold, F. H. Engineering Cytochrome P450s for Enantioselective Cyclopropenation of Internal Alkynes. *J. Am. Chem. Soc*. **142**, 6891-6895 (2020).

6. Jia, Z. J., Gao, S. & Arnold, F. H. Enzymatic Primary Amination of Benzylic and Allylic C(sp3)–H Bonds. *J. Am. Chem. Soc*. **142**, 10279–10283 (2020).

7.  Feng, C. G., Wang, Z. Q., Tian, P., Xu, M. H. & Lin, G. Q. Easily accessible C2-symmetric chiral bicyclo[3.3.0] dienes as ligands for rhodium-catalyzed asymmetric 1,4-addition. *Asian J. Chem*. **3**, 1511-1516 (2008).

8.  Khiar, N. et al. Asymmetric Rhodium-Catalyzed 1,4- and 1,2-Additions of Arylboronic Acids to Activated Ketones in Water at Room Temperature Using a Mixed Sulfur-Olefin Ligand. *Adv. Synth. Catal*. **355**, 1303-1307 (2013).

9.  Díaz-Rodríguez, A. et al. From Diols to Lactones under Aerobic Conditions using a Laccase/TEMPO Catalytic System in Aqueous Medium. *Adv. Synth. Catal*. **354**, 3405-3408 (2012).

10. Smith, A. J., Abbott, L. K. & Martin, S. F. Enantioselective Conjugate Addition Employing 2-Heteroaryl Titanates and Zinc Reagents. *Org. Lett*. **11**, 4200-4203 (2009).

11. Le Notre, J., Allen, J. C. & Frost, C. G. Enantioselective rhodium-catalysed 1,4-additions of 2-heteroarylzinc donors using Me-DUPHOS. *Chem. Commun*. 3795-3797 (2008).

12. Wu, W. et al. Asymmetric Baeyer-Villiger oxidation: classical and parallel kinetic resolution of 3-substituted cyclohexanones and desymmetrization of meso-disubstituted cycloketones. *Chem. Sci*. **10**, 7003-7008 (2019).

13. Smith, A. B. et al. Design, Synthesis, and Biological Evaluation of Monopyrrolinone-Based HIV-1 Protease Inhibitors. *J. Med. Chem*. **46**, 1831–1844 (2003).

14. Touge, T. et al. Multiple Absolute Stereocontrol in Cascade Lactone Formation via Dynamic Kinetic Resolution Driven by the Asymmetric Transfer Hydrogenation of Keto Acids with Oxo-Tethered Ruthenium Catalysts. *J. Am. Chem. Soc*. **141**, 16354-16361 (2019).

15. Sheldrick, G. M. Phase Annealing in SHELX-90: Direct Methods for Larger Structures. *Acta Cryst*. **A46**, 467-473 (1990).

16. Sheldrick, G. M. Crystal Structure Refinement With SHELXL. *Acta Cryst*. **C71**, 3-8 (2015).

17. Müller, P. Practical Suggestions for Better Crystal Structures. *Crystallogr. Rev*. **15**, 57-83 (2009).

18. Zhou, Y. et al. Facile access to chiral gamma-butyrolactones via rhodium-catalysed asymmetric hydrogenation of gamma-butenolides and gamma-hydroxybutenolides. *Chem. Sci*. **14**, 4888-4892 (2023).

19. Wang, Y., Hu, X. & Du, H. Vicinal-Diamine-Based Chiral Chain Dienes as Ligands for Rhodium(I)-Catalyzed Highly Enantioselective Conjugated Additions. *Org. Lett*. **12**, 5482-5485 (2010).

20. Sheldrick, G. M. Phase Annealing in SHELX-90: Direct Methods for Larger Structures. *Acta Cryst*. **A46**, 467-473 (1990).

21. Sheldrick, G. M. Crystal Structure Refinement With SHELXL. *Acta Cryst*. **C71**, 3-8 (2015).

22. Müller, P. Practical Suggestions for Better Crystal Structures. *Crystallogr. Rev*. **15**, 57-83 (2009).