Engineering Heme Proteins for C(sp^3)–H Primary Amination

Thesis by

Shilong Gao

In Partial Fulfillment of the Requirements for

the degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2024

(Defended August 21, 2023)

Shilong Gao

ORCID: 0000-0003-2808-6283

ACKNOWLEDGEMENTS

I am immensely grateful to have reached this significant milestone in my academic journey by completing this Ph.D. thesis. I am indebted to the participants of my work who generously contributed their time and expertise. This accomplishment would not have been possible without the support, encouragement, and contributions of numerous individuals to whom I extend my heartfelt appreciation.

First and foremost, I express my deepest gratitude to my advisor, Prof. Frances H. Arnold, for her invaluable guidance, mentorship, and encouragement throughout the entire process. Trained as a chemist during college, I did not have protein engineering as an option for graduate school on my mind when I first arrived at Caltech. But Frances "converted" me with one persuasive lecture. I am really grateful that she gave me a chance to learn everything about protein engineering from scratch in the past five years. Thank you, Frances, for your constructive feedback, scholarly insights, and belief in my abilities to improve, which have been instrumental in shaping this thesis and my personal growth.

I would like to thank my committee members, Prof. Brian M. Stoltz, Prof. Ryan G. Hadt, and Prof. Harry B. Gray, for their precious time, expertise, and counsel, which have undoubtedly strengthened the quality and rigor of this work. And I extend my thanks to the faculty members and staff at Caltech, especially Dr. Scott C. Virgil and Dr. Mona Shahgholi, for creating an enriching academic environment and providing access to resources essential for carrying out this research. It is comforting to see the Caltech campus's vitality slowly returning after the pandemic.

My sincere appreciation also goes to my colleagues and fellow researchers in the Arnold group, with whom I have had stimulating discussions and shared ideas. I want to express my gratitude to Dr. Sabine Brinkmann-Chen for her constant accessibility to discussions, offering feedback on nearly every written work I created during my time at Caltech. Moreover, your tireless support has been tremendous throughout the rollercoaster journey of graduate school. To my mentor, Zhijun Jia, thank you for starting the primary amination work in the lab and guiding me through the beginning of my journey as a researcher. To my

close collaborators, Soumitra Athavale, Ziyang Qin, and Runze Mao, thank you for your endeavor that enriched this thesis and contributed to advancing knowledge in the field.

To the friends in office 335, Dave Miller, Patrick Almjell, Kadina Johnston, Lucas Schaus, Jenni Kennemur, and Tyler Fulton, thank you for providing emotional support and encouragement when it was most needed. Your presence in my life has been a constant reminder of the importance of work-life balance and the joy found in sharing both personal and professional accomplishments. To Edwin Alfonzo and Anuvab Das, thank you for being excellent friends and co-workers. I will always remember our daily routine of walking home together and our fun discussions along the way. To many past members of the Arnold group, Nicholas Sarai, Bruce Wittman, Nicholas Porter, Noah Dunham, Benjamin Levin, and Dominic Rieger, I cherish every moment that we shared in and outside of work, and I hope our paths will cross again in the future.

My undergraduate advisor, Prof. John K. Snyder, has been an essential figure in my research career. His passion for teaching and science profoundly influenced me to pursue a career in research. As an undergraduate researcher, I appreciated the freedom to explore organic chemistry in John's lab.

Last but not least, to my parents, Wenhong Long and Zhijun Gao, I cannot express enough gratitude for their unconditional support, encouragement, and belief in me throughout my pursuits in science and research. Their unwavering support during both the highs and lows of this journey has been a source of strength and motivation. As I reflect on this transformative odyssey, I am acutely aware that their sacrifices, both big and small, have played a pivotal role in shaping my aspirations and molding me into the person I am today.

ABSTRACT

Primary amine is one of the most prevalent moieties in synthetic intermediates and pharmaceutical compounds. The preparation of aliphatic primary amines via C−H functionalization would provide direct access to the nitrogen-containing compounds from hydrocarbon substrates. While the enzymatic oxyfunctionalization of C–H bonds is well established, the analogous strategy for nitrogen incorporation is unknown in Nature. Likewise, a synthetic method for selective primary amination of aliphatic C–H bonds remains elusive. Combining chemical intuition and inspiration from Nature, chemists and protein engineers have created new heme-containing enzymes for the C(sp^3)–H primary amination via directed evolution. This thesis describes some of the efforts in the continued pursuit of these new-to-nature reactions. Chapter I discusses directed evolution in the context of biocatalysis, the strategies for introducing new-to-nature chemistry in enzymes, the discovery of nitrene transferases from the cytochrome P450 monooxygenase, and finally, the development of C(sp^3)–H primary aminases. Chapter II details the discovery and engineering of serine-ligated cytochrome P411 enzymes that catalyze the first primary amination of C(sp^3)–H bonds with excellent selectivity, affording a broad scope of enantioenriched primary amines. Chapter III demonstrates that these new-to-nature nitrene transferases were engineered to aminate and amidate unactivated, unbiased C(sp^3)–H bonds with unprecedented selectivity. In Chapter IV, engineered protoglobins are shown to utilize hydroxylamine (NH_2OH) for nitrene transfer reactions, including benzylic C–H primary amination and styrene aminohydroxylation. Overall, these new-to-nature reactions can be considered the nitrogen analogs to the C–H oxidation chemistry performed by monooxygenases and peroxygenases. By offering a direct path from saturated precursors, these enzymes present a new biochemical logic for accessing nitrogen-containing compounds. Finally, this work hints at the possible future discovery of natural enzymes that use hydroxylamine precursors for amination chemistry.

PUBLISHED CONTENT AND CONTRIBUTIONS

(* corresponding author; † equal contribution)

1. Jia, Z. J.; **Gao, S.**; Arnold, F. H.* Enzymatic primary amination of benzylic and allylic C(sp^3)–H bonds. *J. Am. Chem. Soc.* **2020**, *142*(23), 10279–10283.

S. G. participated in the execution of the research, including enzymatic reactions, substrate scope study and synthetic applications. S. G. participated in reviewing and editing the manuscript.

2. Athavale, S.V.; *,† **Gao, S.;** † Das, A.; † Mallojjala, S. C.; Alfonzo, E.; Long, Y.; Hirschi, J. S.;* Arnold, F. H.* Enzymatic nitrogen insertion into unactivated C–H bonds. *J. Am. Chem. Soc.* **2022**, *144*(41), 19097–19105.

S. G. discovered the initial activity for unactivated C–H primary amination. S. G. participated in the design and execution of the research. S. G. participated in the preparation of the manuscript.

3. **Gao, S.**; Das, A.; Alfonzo, E.; Sicinski, K. M.; Rieger, D.; Arnold, F. H.* Enzymatic Nitrogen Incorporation Using Hydroxylamine. Submitted.

S. G. discovered the initial activity. S. G. participated in the design and execution of the research. S. G. prepared the manuscript.

PUBLISHED CONTENT NOT INCLUDED IN THESIS

4. Athavale, S. V.; † **Gao, S.;** † Liu, Z.; Mallojjala, S. C.; Hirschi, J. S.;* Arnold, F. H.* Biocatalytic, Intermolecular C−H Bond Functionalization for the Synthesis of Enantioenriched Amides. *Angew. Chem., Int. Ed.* **2021**, *60*, 24864−24869.

S. V. A and S. G. discovered the initial activity. S. G. participated in the design and execution of the research. S. G. participated in the preparation of the manuscript.

5. Qin, Z. Y.; † **Gao, S.;** † Zou, Y.; Liu, Z.; Wang, J. B.; Houk, K. N.; Arnold, F. H.* Biocatalytic Construction of Chiral Pyrrolidines and Indolines via Intramolecular C(sp^3)–H Amination. *ACS Cent. Sci.* **2023**, Manuscript in revision.

S. G. discover the initial activity for the alkynitrene transfer. S. G. participated in the design and execution of the research. S. G. participated in the preparation of the manuscript.

6. 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***, ASAP.* doi.org/10.1021/jacs.3c04870.

S. G. participated in the execution of the research, including substrate scope study.

7. Mao, R.; † **Gao, S.;** † Qin, Z. Y.; † Rogge, T.; Wu, S. J.; Li, M.; Ma, P.; Li, Z. Q.; Houk, K. N.; Arnold, F. H.* Biocatalytic, Enantioenriched Primary Amination of Tertiary C−H Bonds. Manuscript in preparation.

R. M. and S. G. discovered initial activity. S. G. participated in the design and execution of the research. S. G. participated in the preparation of the manuscript.

TABLE OF CONTENTS

LIST OF FIGURES, TABLES, AND SCHEMES

Scheme Number *Page*

ABBREVIATIONS

Å	Ångstrom
ACN	acetonitrile
Ac	acetyl
AcOH	acetic acid
ALA	aminolevulinic acid
Ar	aryl
BSA	bovine serum albumin
BCA	bicinchoninic acid assay
BM3	cytochrome P450 from *Bacillus megaterium* (CYP102A1)
Cys	cystine
d.r.	diastereomeric ratio
ee	enantiomeric excess
E. coli	*Escherichia coli*
epPCR	error-prone PCR
e.r.	enantiomeric ratio
Et	ethyl
EtOH	ethanol
GC	gas chromatography
GC-MS	gas chromatography with mass spectrometry
h	hour(s)
HB	Hyper Broth
HPLC	high performance liquid chromatography
HPLC-MS	high performance liquid chromatography with mass spectrometry
HRMS	high-resolution mass spectrometry
IPGT	isopropyl β-D-1-thiogalactopyranoside
k	rate constant
K_m	Michaelis constant
Kpi	potassium phosphate
LB	Lysogeny Broth (Luria-Bertani medium)
Me	methyl
LC-MS	liquid chromatography with mass spectrometry
MeOH	methanol
min	minute(s)
M9-N	M9 minimum media minus nitrogen sources
OD_{600}	optical density at 600 nm
P450	cystine-ligated cytochrome P450 monooxygenase
P411	serine-ligated cytochrome P450 enzymes
PCR	polymerase chain reaction
PDB	Protein Data Bank
RT	room temperature or retention time

r.r.	regiomeric ratio
s	second(s)
Ser	serine
StEP	staggered extension process
SSM	site saturation mutagenesis
TB	Terrific Broth
TTN	total turnover number
Ts	para-toluenesulfonyl
UV-vis	ultraviolet visible
V_{max}	maximum rate
WT	wild type

Chapter I

DISCOVERY AND DIRECTED EVOLUTION OF HEME ENZYMES FOR C–H AMINATION

1.1 Introduction

Nitrogen-containing compounds such as amines and amides are ubiquitous in biomolecules, pharmaceuticals, and functional materials.[1-2] Traditionally, selective amine synthesis relies on manipulating the inherent reactivity of functional groups, as seen in hydroamination and aminofunctionalization with olefins, reductive amination with ketones, dehydrogenative coupling with alcohol, and amination of aryl halides and boronates.[3-6] However, the majority of readily accessible chemicals acquired from petroleum, plants, and other natural sources often lack these functional groups at the desired position and have mostly carbon-hydrogen (C−H) bonds. To address the need for functionalizing the abundant C–H bonds, various synthetic methods for C−H amination have been explored in recent years, precluding the necessity for pre-functionalized coupling partners.[7-8] However, most C−H amination methods are far from ideal: poor atom economy, harsh reaction conditions, and the need for rare metal catalysts led to increased waste, making the direct synthesis of amines from hydrocarbons a laborious and energy-intensive process.[9] More importantly, the primary amine moiety, especially privileged in valuable building blocks for synthesizing other nitrogen-containing groups, are limited to the synthetic methods via aromatic C–H functionalization. C(sp^3)–H primary amination has remained elusive in synthetic chemistry.

Enzymes are highly efficient, low-cost, and non-toxic, traits crucial for industrial applications.[10] The genetically encoded and fully tunable protein scaffolds provide exquisite control over reactive intermediates with chemo-, regio-, and enantioselectivity

unparalleled by small-molecule catalysts. In particular, C–H functionalization is ubiquitous in biology for decorating natural products and installing chemical handles for downstream functionalization.[11] Yet, such biochemical logic is restricted to oxygen and halogen introduction.[12-13] Enzymatic carbon-nitrogen bond formation often relies on a preoxygenated carbon center, typically through an enzymatic reductive amination route.[14] Enzymatic C–H primary amination, which would provide access to a plethora of aminated products, has not yet been discovered in Nature.

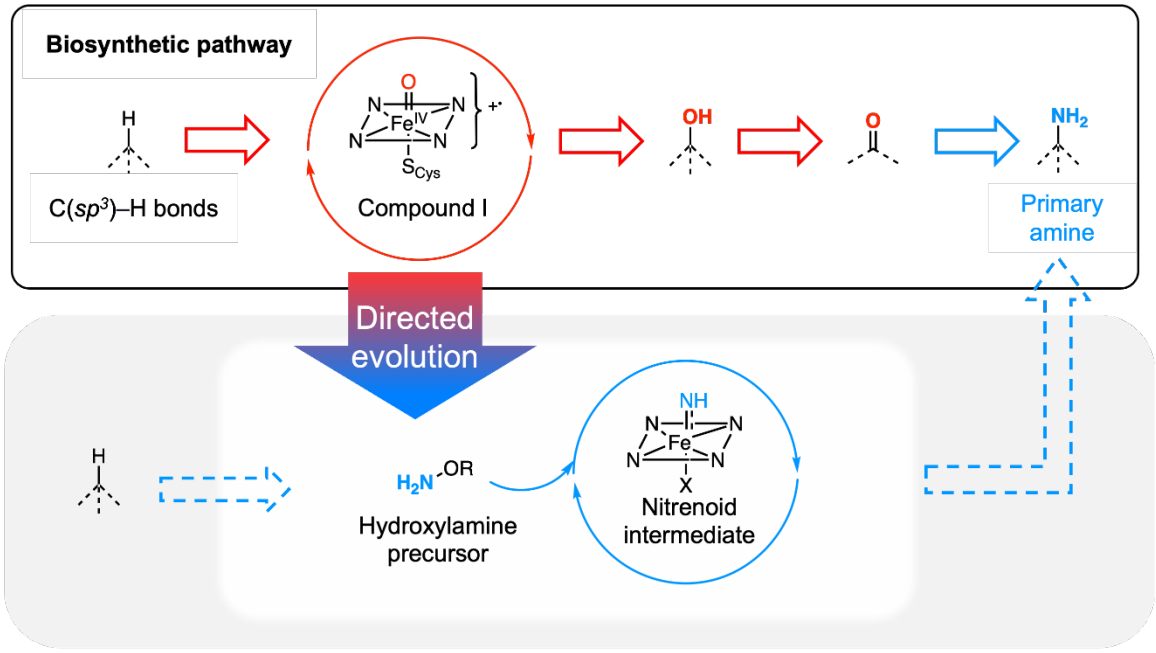


Figure 1-1. Biosynthetic pathway of primary amine moieties and the proposed pathway via direct C–H primary amination.

We seek to revise the logic of biomolecular nitrogen incorporation by engineering heme enzymes to catalyze direct C–H primary amination (Figure 1-1). In the Arnold group, this long-term effort was initiated almost a decade ago with the discovery of the abiotic "nitrene

transferase" activity of heme enzymes, realized via a nitrogen analog of the Fe-oxo species-an active site heme-nitrenoid.[15] Powered by directed evolution, a diverse enzymatic platform for the benzylic, allylic, propargylic, and unactivated C–H primary amination has been developed.[16-18] This introductory chapter will discuss directed evolution in the context of biocatalysis, the strategies for introducing new-to-nature chemistry in enzymes, the discovery of nitrene transferases from the cytochrome P450 monooxygenase, and finally, the development of C–H primary aminases.

1.2 Directed Evolution: An Uphill Walk in the Protein Fitness Landscape

Enzymes are Nature's powerful catalysts that enable all chemical transformations of life. This highly efficient biological machinery is finely tuned by evolution and often unparalleled by small-molecule catalysts created by chemists.[19-20] The protein scaffold is encoded as a linear arrangement of twenty canonical amino acids, folding into dynamic three-dimensional structures that accomplish various functions. The optimization of the protein scaffold by evolution is enabled by introducing genetic diversity via mutations in the DNA encoding the enzyme. In his influential 1970 work titled "Natural Selection and the Concept of a Protein Space", John Maynard Smith established a theoretical framework explaining how protein evolution through natural selection is achievable within a vast sequence space predominantly lacking functional sequences.[21] He suggested that for evolution to occur, a daughter variant that at least retains functionality must exist among the potential single mutations of a functional protein. The theory suggests that the sequence space of functional proteins forms a connected network that can be navigated via single

amino acid mutations. Eventually, mutations that benefit the organism's survival become enriched in the gene pool.

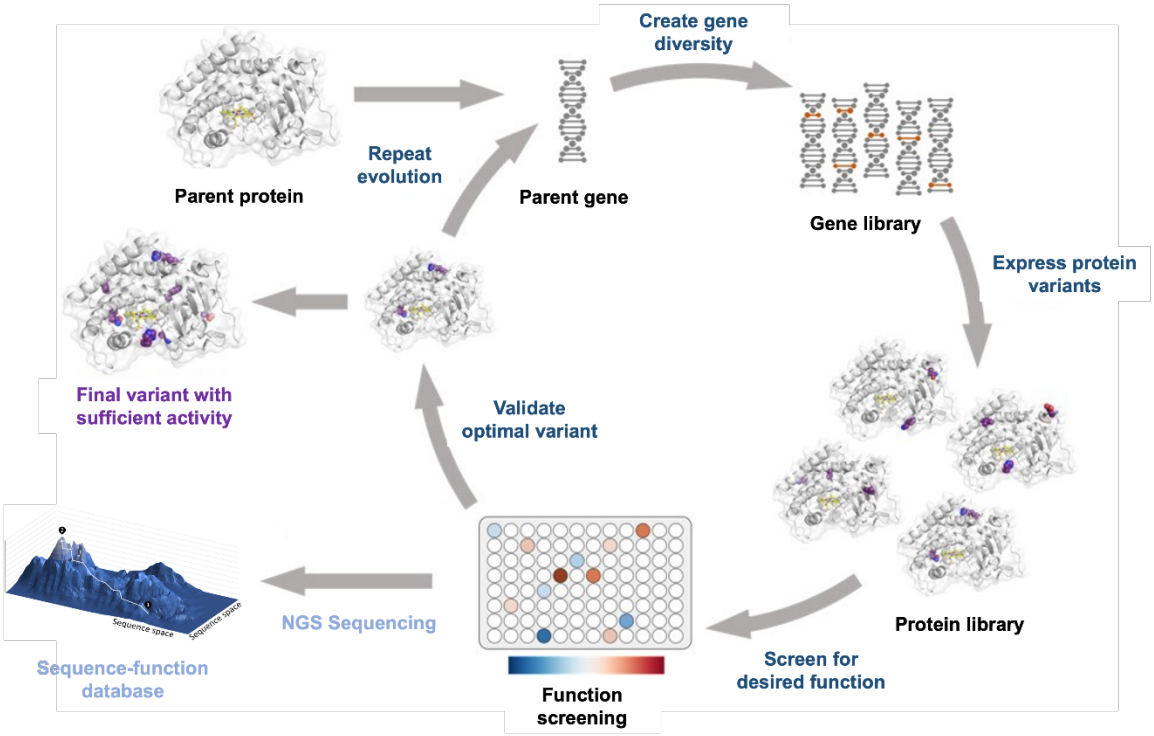


Figure 1-2. Overview of directed evolution.

The development of screening technology and genetic engineering, particularly the error-prone polymerase chain reaction (PCR) in the 1980s, facilitated the emergence of "directed evolution" as a technique for protein engineering in a human-friendly timescale. The Arnold group and others demonstrated that enzymes could be evolved toward a specific objective by gradually accumulating predominantly single amino acid mutations in a stochastic uphill progression on a high-dimensional fitness landscape in sequence space, where fitness is defined by the experimenter.[22-23] Unlike natural evolution, directed

evolution involves much more aggressive mutagenesis to traverse the sequence landscape faster and speed up the evolution process. The approach requires sequence diversification by mutagenesis to generate a library of modified DNA sequences followed by screening to identify enzyme variants with improved properties, often in a high-throughput manner, with iterative rounds until the desired fitness is achieved (Figure 1-2).[24] New methods for directed evolution, such as machine-learning-guided directed evolution, can amplify the power of evolution by a faster search of the local optimum on this landscape.[25-26] This revolutionary technique has been widely used to improve intrinsic properties such as stability under harsh conditions and enhance activity and specificity on non-native substrates, significantly impacting medicine, agriculture, and industrial processes.

1.3 Engineering Enzymes for New-to-Nature Chemistry

Finding the initial activity is crucial for a directed evolution campaign since an optimization process cannot start from zero. A desirable initial activity closely related to the existing natural chemistries of enzymes can often be found by screening libraries of natural and engineered enzymes. In addition, the development of metagenomic mining allows researchers to interrogate *in silico* libraries of sequence data across different species to identify even more starting points for enzyme engineering.[27] However, despite the astonishing diversity of naturally occurring enzymatic processes, natural enzymes often do not directly catalyze transformations preferred by chemists. Therefore, engineering an enzyme to catalyze a new-to-nature reaction, especially breaking and forging bonds

through new-to-nature mechanisms, was widely regarded as a tremendous challenge in biocatalysis.

Fortunately, enzymes are catalytically promiscuous: they catalyze fortuitous side reactions in addition to their primary functions.[28] Chemists often overlook this versatility when considering enzymes as catalysts geared toward one specific type of transformation.[29] The natural evolution process takes advantage of these genetically encoded promiscuities to create new enzymes with novel functions in response to changing environments. The desirable "side reactions" would be further optimized and diversified by evolution and become new families of enzymes (Figure 1-3). Indeed, when circumstances arise where a promiscuous activity confers advantages, the natural selection process can generate a novel enzyme with that function as its predominant feature. In particular, new enzymes can emerge when a novel food source or the capacity to break down a new antibiotic or pesticide presents an opportunity to surpass neighboring organisms in adaptation. Through millennia of evolution, nature has fashioned a plethora of enzymes through this process to catalyze diverse reactivities that make life possible.

Although highly enabling, the native reaction pathways in enzymes are still restricted to the natural biosynthetic logics; non-natural catalytic cycles and intermediates that can be desirable in synthetic routes are not accessible from natural precursors. Alternatively, one could enable new-to-nature chemistries by introducing new functionalities redirecting the native catalytic cycles or introducing new intermediates using synthetic precursors. Combining chemical intuition and biochemical knowledge, chemists and protein engineers

can introduce new-to-nature reactivities to enzymes and create unprecedented reactions in both worlds.

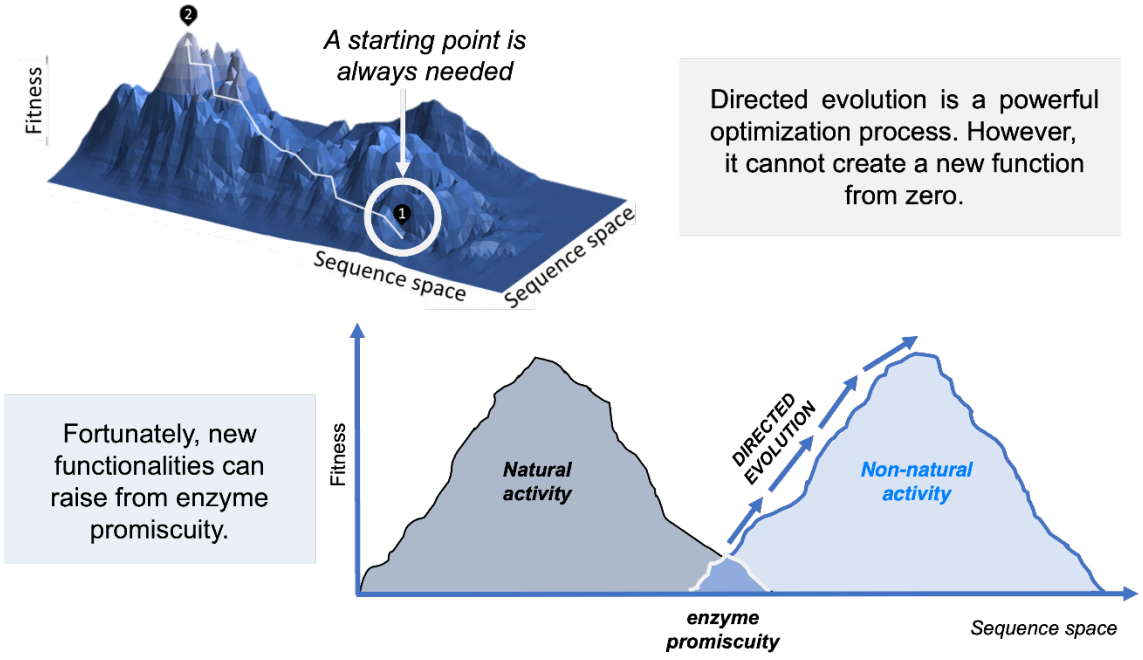


Figure 1-3. A cartoon illustration depicts catalytic promiscuity, which refers to an enzyme's capability to exhibit additional reactivities alongside its primary reactivity. These supplementary reactivities may or may not be present in the enzyme. Through directed evolution, these reactivities can be improved and adjusted accordingly.

Enzyme cofactors, such as nicotinamide, flavin, thiamine, and pyridoxal phosphate, are of particular interest as many possess versatile catalytic potential for chemical transformations.[30-33] Compared with the diversity of small-molecule catalysts invented by chemists, nature uses a more limited set of organocofactors for catalysis. However, a given enzyme family can encompass diverse functions often as broad as and sometimes exceeds that of their counterparts in small-molecule catalysts, thanks to the exquisite control of the

enzyme scaffold over the reactive intermediates. This chemical flexibility provides opportunities to use the diverse properties of cofactors to repurpose key intermediates in the catalytic cycles of enzymes into different reaction pathways to access new enzymatic activities. For instance, Hyster and coworkers created various new-to-nature biocatalysts by using photoexcitation of nicotinamide-dependent enzymes.[30] The nicotinamide cofactor of ketoreductases with the native function of carbonyl reduction was photoexcited to transfer and generate a prochiral radical intermediate upon dehalogenation or nitrogen-oxygen bond cleavage. Accessing radical intermediates within a tunable protein scaffold opens possibilities of novel reductase-photoredox systems for highly selective carbon-carbon and carbon-nitrogen formations.[34]

Similarly, metalloenzymes carry out a variety of complex reactions in nature. The versatility of transition metal electronic states and coordination modes provides the basis for diverse two-electron and single-electron redox/non-redox chemistry.[35-38] In particular, metalloporphyrin complexes have been investigated for more than half a century as structural analogues to heme cofactors in proteins.[39] Many metalloporphyrins were developed to mimic the oxo-transfer activities of cytochrome P450s through the formation of high-valent metal–oxo intermediates.[40] Conversely, carbene and nitrene chemistry, non-natural intermediates that resemble compound I, have been introduced to the enzymatic system for new-to-nature chemistry.[15] The following chapter will use metallonitrene in nitrene transferases as a case to explain this approach in depth.

Finally, other approaches to introduce non-native functions in protein scaffolds include the installation of artificial cofactors known in small molecular catalysis or genetic encoding of non-canonical amino acids that function as reaction centers. Some excellent examples of these approaches include incorporating Ir-porphyrin in myoglobin for carbene reactivities by Hartwig and Clark and genetic encoding of photosensitizer non-canonical amino acid in beta-barrel protein for [2+2] cycloaddition by Green.[41-42] These approaches can also serve as starting points for directed evolution.[43]

1.4 From Monooxygenases to Nitrene Transferases

Oxyfunctionalization of C–H bonds is ubiquitous in biology.[44-45] Cytochromes P450 (P450s), among other dioxygen- and peroxide-dependent heme enzymes, can form a high-valent iron oxo species (compound I) to catalyze hydroxylation reactions.[46] The astonishing ability of these heme-containing enzymes to selectively oxygenate specific C–H bonds with molecular oxygen has captured the attention of the broader chemical community since their discovery in the 1960s. Pioneering studies by Groves and collaborators using iron tetraphenylporphyrin complexes as synthetic model systems elucidated the 'radical rebound' mechanism.[47]

Inspired by these studies and related work for transition metal-catalyzed amination, Breslow and Gellman first demonstrated in 1982 that iron tetraphenylporphyrin complexes generate a nitrene intermediate (Fe═NR), the isoelectronic and structural analog of compound I (Fe═O), to perform C–H amination reactions (Figure 1-4).[48-49] Since the iron nitrene species involved in these amination reactions closely resemble compound I in

native P450s, Breslow and Gellman next asked whether a cytochrome P450 itself could perform amidation instead of oxygenation using appropriate precursors.[50] Indeed, the rabbit liver microsomal cytochrome P450-LM3 was shown to catalyze intramolecular C–H amidation with a total turnover number (TTN) of 2.2. Unfortunately, limited protein engineering methods at the time restrained the interest of the catalysis community in this non-natural enzymatic reaction. The rapid degradation of iminoiodinanes also made it challenging to optimize the reaction further. Nevertheless, this outcome illustrated that the enzymes could carry out reactions beyond their native repertoire, and it showcased the possibility of chemists leveraging their knowledge to elicit novel chemical reactions from nature's catalytic apparatus.

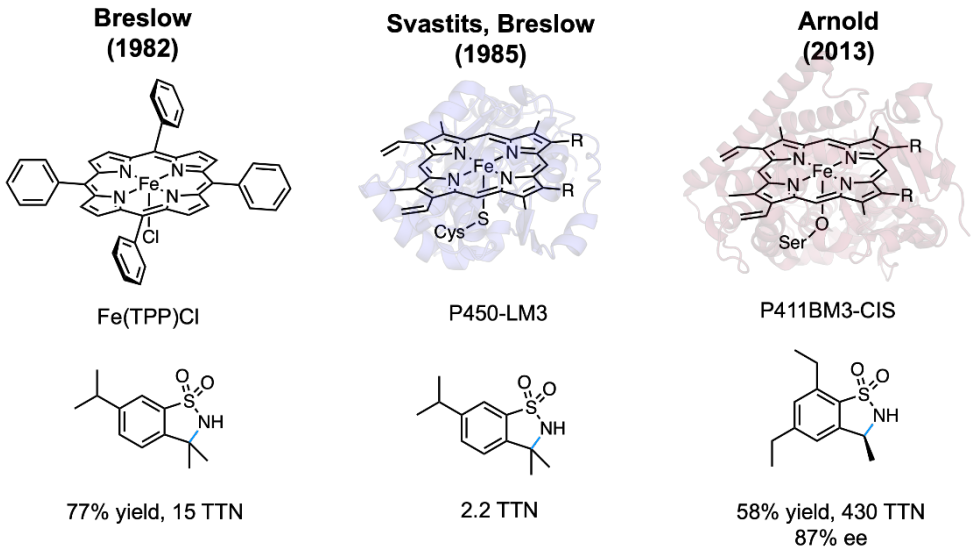
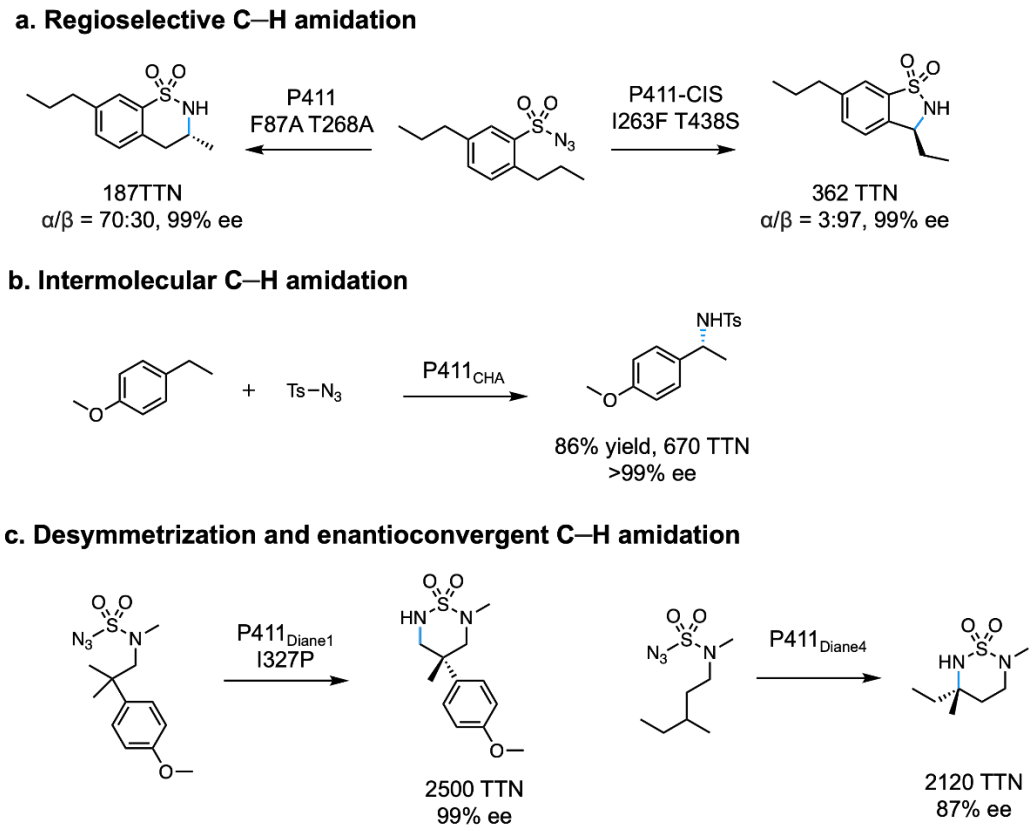


Figure 1-4. In 1982, iron (III) tetraphenylporphyrin chloride (Fe(TPP)Cl) was first utilized as a catalyst for biomimetic intramolecular amidation.[48] In 1985, enzyme-catalyzed intramolecular amidation can be achieved using rabbit liver P450-LM3.[50] In 2013, engineered P411BM3-CIS catalyzed asymmetric intramolecular amidation.[51] R, –CH2CH2CO2H; Cys, cysteine; Ser, serine.

After nearly three decades, McIntosh et al. revisited the enzymatic reaction proposed by Breslow and Gellman, employing a serine-ligated cytochrome P450BM3, namely cytochromes P411, and a sulfonyl azide precursor.[51] This work achieved a significant breakthrough with the help of directed evolution by introducing the first nitrene transferase of catalytic relevance, providing improved reactivity and enantioselectivity for sultam products. In a subsequent study, Hyster et al. demonstrated that under the control of a protein scaffold, a five-membered or a six-membered amidation product could be prepared with exceptional site selectivity (Scheme 1-1.a.).[52] This regiodivergent amidation exemplified the ability of enzymes to override inherent regiochemical preferences through precise substrate binding within the active site. Furthermore, Prier et al. expanded this biocatalytic enantioselective C–H amidation to intermolecular processes (Scheme 1-1.b.).[53] More recently, Yang et al. reaffirmed the exquisite control exerted by protein scaffolds in achieving regioselective C–H functionalization, desymmetrization, and enantioconvergent C–H functionalization within the same research (Scheme 1-1.c.).[54] Overall, these studies emphasized the potential to attain site selectivity that was previously inaccessible with conventional small-molecule catalysts.

Today, enzymatic nitrene transfer chemistry has enjoyed significant success as one example of learning from nature's mechanisms, exploiting enzyme catalytic promiscuity, and exploring new-to-nature functions to discover novel biocatalysts starting from nature's vast repertoire. Notably, nitrene transfer reactions were previously thought to have no equivalent process in biology. Nevertheless, the Ohnishi group discovered a cytochrome P450 enzyme (BezE) from *Streptomyces* sp. RI18 that plays a crucial role in the

biosynthesis of benzastatin natural products.[55] This enzyme facilitates an intramolecular



Scheme 1-1. a. Enzymatic regiodivergent C–H amidation by Hyster et al. (2014).[52] **b.** Enzymatic intermolecular benzylic C–H amidation by Prier et al. (2017).[53] **c.** Asymmetric amidation of various C–H Bonds by Yang et al. (2019).[54]

aziridination process through a nitrene transfer mechanism, utilizing N-acetoxyaniline as the nitrene source. We were delighted that mechanistic insights gained through new-to-nature reactions could inspire the discovery of a novel biosynthesis pathway. This discovery in biochemistry also precluded that specific hydroxylamine-derivative reagents can function as nitrene precursors in nature.

1.5 Enzymatic C(sp^3)–H Primary Amination

Earlier developed nitrene transfer reactions typically employ iminoiodinanes or organic azides as nitrene precursors.[56] While these precursors are well studied for their nitrene generation pathway, they also pose risks due to their explosiveness and hazards on a large scale. More importantly, the resulting product of amidation reactions bears protecting groups such as sulfonyl groups that are difficult to remove. Primary amines, prevalent in natural products, pharmaceuticals, and functional materials, cannot be easily obtained via nitrene chemistry enabled by azide precursors. Therefore, the direct synthesis of unprotected, enantioenriched primary amines via practical chemistry remains challenging with azide reagents.

In this regard, hydroxylamine-derived reagents emerged as alternative nitrene sources. A key reagent, *O*-pivaloylated hydroxylammonium reagent (PivONH3OTf), initially reported by Guimond and Fagnou as a surrogate to install a hydroxamic acid directing group,[57] has been particularly enabling in this field. The *O*-electrophilic substitution lowers the bond dissociation energy of the N–O bond, and its protonated form prevents a facile rearrangement to the more stable *N*-hydroxy amide and thus makes the reagent an ideal NH2+ surrogate.[58] This reagent and other *O*-electrophilic substituted hydroxylamine reagents have enabled diverse amination reactions, including aziridination, aminohalogenation, diamination and aminohydroxylation of olefins, amination of sulfur compounds, and aromatic C–H primary amination.[59] Particularly, this reagent had been employed in alkene aminohydroxylation reactions catalyzed by iron phthalocyanine

developed by the Morandi group.[60] However, due to the high energy and small size of non-substituted nitrene species, achieving asymmetric transformations involving this intermediate has proven to be highly challenging with small-molecule catalysts. Drawing inspiration from Nature's nitrene transferase BezE and Morandi's iron (II) phthalocyanine catalytic system, we engineered *Rhodothermus marinus* cytochrome *c* variants to catalyze the enantioselective aminohydroxylation of styrene.[61] This effort has resulted in the successful incorporation of *O*-pivaloylated hydroxylammonium reagents in hemoprotein-catalyzed amination reactions.

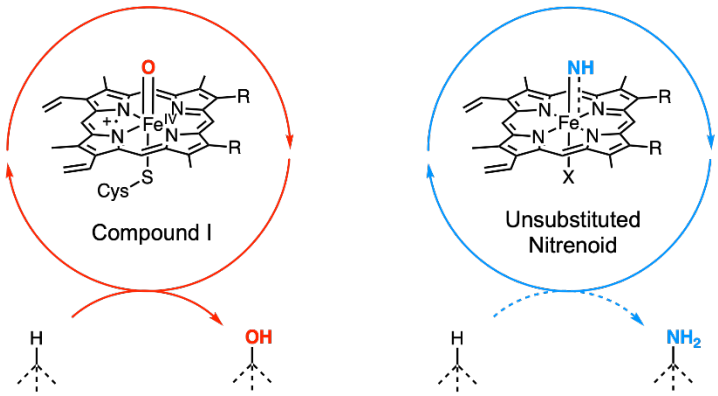


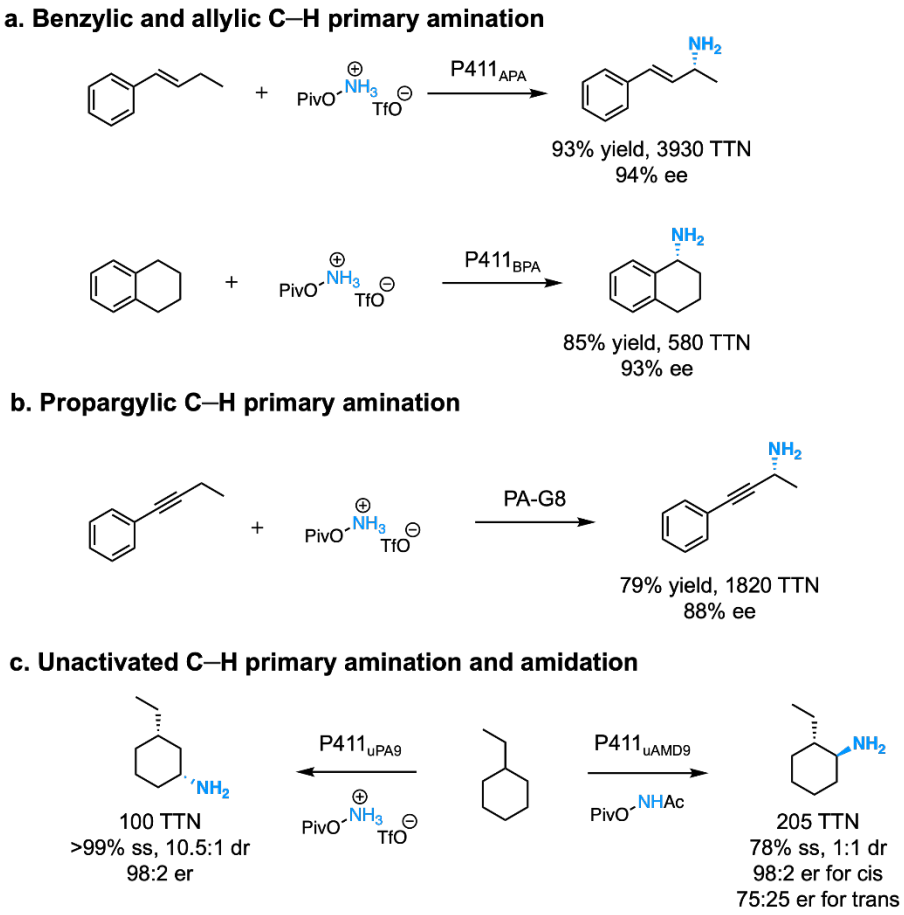
Figure 1-5. Comparison between compound I and unsubstituted nitrenoid intermediate. The proposed C–H primary amination is the analogous reaction of enzymatic C–H hydroxylation.

From an enzymologist's perspective, the putative iron nitrene species (Fe═NH) involved in these amination reactions resemble the closest structural mimic of compound I (Fe═O) in native P450 catalysis (Figure 1-5). And the activation pathway of hydroxylamine that involves N–O bond cleavage is likely to resemble the O–O bond cleavage steps (from compound 0 to compound I) in monooxygenases and peroxygenases. However, compared

to the rich oxidation chemistry enabled by compound I, the amination reactions enabled by nitrene chemistry are missing an important segment: the conversion of C(sp^3)–H bonds to C–OH functional groups is iconic among the oxidation chemistry performed by P450 monooxygenases,[62-63] while its nitrene analog reaction, C(sp^3)–H primary amination, remained elusive.

Based on the precedence of enzymatic aminohydroxylation, we hypothesized that an engineered protein scaffold could direct the heme nitrene intermediate for the C(sp^3)–H primary amination. In 2020, we first reported this biocatalytic reaction using P411 variants (Scheme 1-2.a.).[16] Variant P411BPA engineered for the primary amination of benzylic C–H bonds exhibits high activity and excellent enantiocontrol. By bypassing the reliance on oxygenated precursors, this single-step, enzymatic C(sp^3)–H functionalization by "primary aminases" provides a new route for biology to synthesize nitrogen-containing molecules. More intriguingly, variant $P411_{APA}$ engineered for the primary amination of allylic C–H bonds can chemoselectively aminate the allylic C–H bond, leaving the olefin moiety completely untouched. Most notably, no small-molecule catalysts are known to promote the primary amination of C(sp^3)–H bonds despite extensive previous efforts dedicated to studying Fe-catalyzed amination. The successful development of this biocatalytic primary amination of C(sp^3)–H bonds illustrates the enormous potential of enzymes to enable desirable reactivities elusive to small-molecule catalysts. In subsequent work, Liu et al. expanded this enzymatic reaction to aminate propargylic C–H bonds with PA variants selectively (Scheme 1-2.b.).[17] With this diverse enzyme platform, an array of readily

available hydrocarbon starting materials could be converted into value-added primary amines.



Scheme 1-2. a. Enzymatic primary amination of benzylic and allylic C–H bonds by Jia et al (2020).[16] **b.** Enzymatic primary amination of propargylic C–H bonds by Liu et al (2021).[17] **c.** Enzymatic primary amination and amidation of unactivated C–H bonds by Athavale et al. (2022).[18]

In nature, P450 monooxygenases are capable of hydroxylating unactivated C(sp^3)–H bonds using a highly reactive Fe=O intermediate with unparalleled site- and enantioselectivities. Previously, despite almost a decade of research, the intermolecular C(sp^3)–H amination via

enzymatic nitrene intermediates was limited to electronically activated C(sp^3)–H systems. And the small-molecule transition-metal catalysts capable of catalyzing the primary amination of unactivated C–H bonds remained elusive. Most recently, directed evolution of a nitrene transferase has finally unveiled enzymes that selectively introduce primary amine and amide at the unbiased, unactivated C(sp^3)–H bonds under mild conditions (Scheme 1-2.c.).[18] The desymmetrization of methyl- and ethylcyclohexane serve as remarkable examples of the excellent site- and enantioselectivity of the enzymes. Furthermore, the ability of evolved biocatalysts to convert abundant hydrocarbon feedstocks into enantioenriched amines will pave the way for the rapid preparation of valuable molecular entities in pharmaceuticals and agrochemicals.

1.6 Conclusion and outlook

From monooxygenases to primary aminases, directed evolution, in combination with chemical intuition, has enabled several unprecedented biocatalytic nitrene transfer reactions. Nevertheless, many challenges remain to be addressed. Future work in this field should focus on engineering primary aminases for a broader scope of substrates, such as prodrug candidates and functional molecules, and bringing these biocatalytic transformations to a practical level by exploring other robust heme enzyme systems along with practical amination reagents. Further exploration of the activation pathway and characterization of the unprotected nitrene intermediate would be valuable for generalizing this biocatalytic process as well.

Overall, the field of enzymatic nitrene transfer will remain an active area of investigation for the foreseeable future. Beyond installing primary amines, the next challenge for the biocatalytic nitrene transfer reaction would be installing *N*-alkyl groups since moieties such as methylamine are also prevalent in natural products and drug molecules. Numerous iron catalysts have been demonstrated to use azide and hydroxylamine reagents for intermolecular and intramolecular *N*-alkyl group transfer reactions.[64-65] Filling this gap in the nitrene transferases platform would further enrich the biocatalytic toolbox.

Chapter I Bibliography

1. Ricci, A. Amino Group Chemistry: From Synthesis to the Life Sciences (Wiley-VCH: Weinheim, 2008).
2. Legnani, L.; Bhawal, N. B.; Morandi, B. Recent Developments in the Direct Synthesis of Unprotected Primary Amines. *Synthesis* **2017**, *49*, 776-789.
3. Gross, T.; Seayad, M. A.; Ahmad, M.; Beller, M. Synthesis of Primary Amines: First Homogeneously Catalyzed Reductive Amination with Ammonia. *Org. Lett*. **2002**, *4*, 2055-2058.
4. Hartwig, F. J. Evolution of a Fourth Generation Catalyst for the Amination and Thioetherification of Aryl Halides. *Acc. Chem. Res*. **2008**, *41*, 1534-1544.
5. Surry, S. D.; Buchwald, L. S. Biaryl Phosphane Ligands in Palladium-Catalyzed Amination. *Angew. Chem, Int. Ed*. **2008**, *47*, 6338-6361.
6. Chan, D.; Monaco, L. K.; Wang, -P. R.; Winters, P. M. New N- and O-arylations with phenylboronic acids and cupric acetate. *Tetrahedron Lett.* **1998**, *39*, 2933-2936.
7. Hartwig, J. F. *Evolution of C–H bond Functionalization from Methane to Methodology*. *J. Am. Chem. Soc*. **2016** 138, 2-24.
8. Yamaguchi, J.; Yamaguchi, A. D. & Itami, K. *C–H bond functionalization: emerging synthetic tools for natural products and pharmaceuticals. Angew. Chem. Int. Ed*. **2012** *51*, 8960–9009.
9. Legnani, L.; Prina Cerai, G.; Morandi, B. Direct and Practical Synthesis of Primary Anilines through Iron-Catalyzed C–H Bond Amination. *ACS Catal*. **6**, 8162-8165 (2016).
10. Brandenberg, F. O.; Fasan, R.; Arnold, H. F. Exploiting and Engineering Hemoproteins for Abiological Carbene and Nitrene Transfer Reactions. *Curr. Opin. Biotechnol*. **2017**, *47*, 102-111.
11. Lewis, C. J.; Coelho, S. P.; Arnold, H. F. Enzymatic functionalization of carbon–hydrogen bonds. *Chem. Soc. Rev*. **2011**, *40*, 2003-2021.
12. Ortiz de Montellano, R. P.; Hydrocarbon Hydroxylation by Cytochrome P450 Enzymes. *Chem. Rev*. **2010**, *110*, 932-948.
13. Agarwal, V.; Miles, Z. D.; Winter, J. M.; Eustáquio, A. S.; El Gamal, A. A.; Moore, B. S. Enzymatic Halogenation and Dehalogenation Reactions: Pervasive and Mechanistically Diverse. *Chem. Rev.* **2017**, *117*, 5619– 5674.
14. Tavanti, M.; Mangas-Sanchez, J.; Montgomery, S. L.; Thompson, M. P.; Turner, N. J. A Biocatalytic Cascade for the Amination of Unfunctionalised Cycloalkanes. *Org. Biomol. Chem.* **2017**, *15*, 9790– 9793.
15. 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*(5), 1209-1225.
16. Jia, Z. J.; Gao, S.; Arnold, F. H. Enzymatic primary amination of benzylic and allylic C (sp3)–H bonds. *J. Am. Chem. Soc.* **2020**, *142*(23), 10279–10283.
17. Liu, Z.; Qin, Z.-Y.; Zhu, L.; Athavale, S. V.; Sengupta, A.; Jia, Z. J.; Garcia-Borràs, M.; Houk, K. N. and Arnold, F. H. An enzymatic platform for primary amination of 1-aryl-2-alkyl alkynes. *J. Am. Chem. Soc.* **2021**, *144*(1), 80-85.

18. Athavale, S.V.; Gao, S.; Das, A.; Mallojjala, S. C.; Alfonzo, E.; Long, Y.; Hirschi, J. S. and Arnold, F. H. Enzymatic nitrogen insertion into unactivated C–H bonds. *J. Am. Chem. Soc.* **2022**, *144*(41), 19097–19105.
19. Chen, K.; Arnold, F. H. Engineering New Catalytic Activities in Enzymes. *Nat. Catal.* **2020**, *3*, 203– 213.
20. Nestl, B. M.; Hauer, B. Engineering of Flexible Loops in Enzymes. *ACS Catal.* **2014**, *4*, 3201– 3211.
21. Maynard Smith, J. Natural selection and the concept of a protein space. **1970**, *Nature*, *225*(5232), 563-564.
22. Fasan, R.; Jennifer Kan, S. B.; Zhao, H. A Continuing Career in Biocatalysis: Frances H. Arnold. *ACS Catal*. **2019**, *9*, 9775-9788.
23. Arnold, F. H. Innovation by evolution: bringing new chemistry to life (Nobel Lecture). *Angew. Chem. Int. Ed*. **2019**, *58*(41), 14420-14426.
24. Wang, Y.; Xue, P.; Cao, M.; Yu, T.; Lane, S. T.; & Zhao, H. Directed evolution: methodologies and applications. *Chem. rev* **2021**. *121*(20), 12384-12444.
25. Yang, K. K.; Wu, Z.; Arnold, F. H. Machine-learning-guided directed evolution for protein engineering. *Nat. Methods* **2019**, *16*(8), 687-694.
26. Wittmann, B. J.; Johnston, K. E.; Wu, Z.; Arnold, F. H. Advances in machine learning for directed evolution. *Curr. Opin. Struct. Biol.* **2021**, *69*, 11-18.
27. Robinson, S. L.; Piel, J.; Sunagawa, S. A roadmap for metagenomic enzyme discovery. *Nat. Prod. Rep.* **2021**, *38*(11), 1994-2023.
28. Bloom, J. D.; Arnold, F. H. In the light of directed evolution: pathways of adaptive protein evolution. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 9995-10000.
29. Khersonsky, O.; Tawfik, D. S. Enzyme promiscuity: a mechanistic and evolutionary perspective. *Annu. Rev. Biochem.* **2010**, *79*, 471–505.
30. Emmanuel, M. A.; Greenberg, N. R.; Oblinsky, D. G.; Hyster, T. K. Accessing non-natural reactivity by irradiating nicotinamide-dependent enzymes with light. *Nature* **2016**, *540*, 414–417.
31. Dockrey, S. A. B.; Narayan, A. R. H. Flavin-dependent biocatalysts in synthesis. *Tetrahedron* **2019**, *75*, 1115–1121.
32. Meyer, D.; Neumann, P.; Ficner, R.; Tittmann, K. Observation of a stable carbene at the active site of a thiamin enzyme. *Nat. Chem. Bio.* **2013**, *9*, 488–490.
33. Romney, D. K. et al. Unlocking reactivity of TrpB: a general biocatalytic platform for synthesis of tryptophan analogues. *J. Am. Chem. Soc.* **2017**, *139*, 10769–10776.
34. Biegasiewicz, K. F.; Cooper, S. J.; Emmanuel, M. A.; Miller, D. C.; Hyster, T. K. Catalytic promiscuity enabled by photoredox catalysis in nicotinamide-dependent oxidoreductases. *Nat. Chem.* **2018**, *10*(7), 770-775.
35. Solomon, E. I.; Decker, A.; Lehnert, N. Non-heme iron enzymes: contrasts to heme catalysis. *Proc. Natl. Acad. Sci. U.S.A.* **2003**,*100*(7), 3589-3594.
36. Lontie, R. *Copper Proteins and Copper Enzymes: Volume II* (Vol. 2). CRC press. **2018**.
37. Ragsdale, S. W. Nickel-based enzyme systems. *J. Biol. Chem.* **2009**. *284*(28), 18571-18575.

38. Andreini, C.; Bertini, I.; Cavallaro, G.; Holliday, G. L.; Thornton, J. M. Metal ions in biological catalysis: from enzyme databases to general principles. *J. Biol. Inorg. Chem.* **2008**, *13*, 1205-1218.
39. Poulos, T. L. Heme enzyme structure and function. *Chem. Rev.* **2014**, *114*(7), 3919-3962.
40. Gao, W. Y.; Chrzanowski, M.; Ma, S. Metal–metalloporphyrin frameworks: a resurging class of functional materials. *Chem. Soc. Rev.* **2014**, *43*(16), 5841-5866.
41. Key, H. M.; Dydio, P.; Liu, Z.; Rha, J. Y. E.; Nazarenko, A.; Seyedkazemi, V.; Clark, D. S.; Hartwig, J. F. Beyond iron: iridium-containing P450 enzymes for selective cyclopropanations of structurally diverse alkenes. *ACS Cent. Sci.* **2017**, *3*(4), 302-308.
42. Trimble, J.S.; Crawshaw, R.; Hardy, F.J.; Levy, C.W.; Brown, M.J.; Fuerst, D.E.; Heyes, D.J.; Obexer, R.; Green, A.P. A designed photoenzyme for enantioselective [2+ 2] cycloadditions. *Nature* **2022**, *611*(7937), 709-714.
43. Zhao, H. Directed evolution of artificial enzymes with new-to-nature functions. *Biophys. J.* **2023**, *122*(3), 146a.
44. Beltrán-Nogal, A.; Sánchez-Moreno, I.; Méndez-Sánchez, D.; de Santos, P. G.; Hollmann, F.; Alcalde, M. Surfing the wave of oxyfunctionalization chemistry by engineering fungal unspecific peroxygenases. *Curr. Opin. Struct. Biol.* **2022**, *73*, 102342.
45. Aranda, C.; Carro, J.; González-Benjumea, A.; Babot, E.D.; Olmedo, A.; Linde, D.; Martínez, A.T.; Gutiérrez, A.; Advances in enzymatic oxyfunctionalization of aliphatic compounds. *Biotechol. Adv.* **2021**, *51*, 107703.
46. Rittle, J.; Green, M. T. Cytochrome P450 compound I: capture, characterization, and CH bond activation kinetics. *Science* **2010**, *330*(6006), 933-937.
47. Groves, J. T.; Nemo, T. E.; Myers, R. S. Hydroxylation and epoxidation catalyzed by iron-porphine complexes. Oxygen transfer from iodosylbenzene. *J. Am. Chem. Soc.* **1979**, *101*(4), 1032-1033.
48. Breslow, R.; Gellman, S. H. Tosylamidation of cyclohexane by a cytochrome P-450 model. *J. Chem. Soc. Chem. Commun.* **1982**, *24*,1400–1401.
49. Breslow, R.; Gellman, S. H. Intramolecular nitrene carbon–hydrogen insertions mediated by transition-metal complexes as nitrogen analogs of cytochrome P-450 reactions. *J. Am. Chem. Soc.* **1983**, *105*, 6728–6729.
50. Svastits, E. W.; Dawson, J. H.; Breslow, R.; Gellman, S. H. Functionalized Nitrogen Atom Transfer Catalyzed by Cytochrome P-450. *J. Am. Chem. Soc.* **1985**, *107*, 6427–6428,
51. McIntosh, J. A.; Coelho, P. S.; Farwell, C. C.; Wang, Z. J.; Lewis, J. C.; Brown, T. R.; Arnold, F. H. Enantioselective Intramolecular C–H Amination Catalyzed by Engineered Cytochrome P450 Enzymes in Vitro and in Vivo. *Angew. Chem.; Int. Ed.* **2013**, *52*, 9309–9312.
52. Hyster, T. K.; Farwell, C. C.; Buller, A. R.; McIntosh, J. A.; Arnold, F. H. Enzyme-Controlled Nitrogen-Atom Transfer Enables Regiodivergent C–H Amination. *J. Am. Chem. Soc.* **2014**, *136*, 15505–15508.

53. 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
54. 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.
55. Tsutsumi, H.; Katsuyama, Y.; Izumikawa, M.; Takagi, M.; Fujie, M.; Satoh, N.; Shin-Ya, K.; Ohnishi, Y. Unprecedented cyclization catalyzed by a cytochrome P450 in benzastatin biosynthesis. *J. Am. Chem. Soc.* **2018**, *140*(21), 6631-6639.
56. Wang, Y.C.; Lai, X.J.; Huang, K.; Yadav, S.; Qiu, G.; Zhang, L.; Zhou, H.; Unravelling nitrene chemistry from acyclic precursors: recent advances and challenges. *Org. Chem. Front.* **2021**, *8*(7),1677-1693.
57. Guimond, N.; Gorelsky, S. I.; Fagnou, K. Rhodium (III)-catalyzed heterocycle synthesis using an internal oxidant: improved reactivity and mechanistic studies. *J. Am. Chem. Soc.* **2011**,*133*(16), 6449-6457.
58. Zinner, G. Über O-acylierte Hydroximsäure-ester und ihre Spaltung zu unsubstituierten O-Acylhydroxylaminen. X. Mitteilung über Hydroxylamin-Derivate. *Arch. Pharm.* **1960**, *293*(7), 657-661.
59. Gasser, V. C.; Makai, S.; Morandi, B. The advent of electrophilic hydroxylamine-derived reagents for the direct preparation of unprotected amines. *Chem. Comm.* **2022**. *58*(72), 9991-10003.
60. Legnani, L.; Morandi, B. Direct Catalytic Synthesis of Unprotected 2-Amino-1-Phenylethanols from Alkenes by Using Iron (II) Phthalocyanine. *Angew. Chem.; Int. Ed.* **2016**, *55*(6), 2248-2251.
61. Cho, I.; Prier, C. K.; Jia, Z. J.; Zhang, R. K.; Görbe, T.; Arnold, F. H. Enantioselective aminohydroxylation of styrenyl olefins catalyzed by an engineered hemoprotein. *Angew. Chem.; Int. Ed.* **2019**, *58*(10), 3138-3142
62. Liu, Y.; Wang, C.; Yan, J.; Zhang, W.; Guan, W.; Lu, X.; Li, S. Hydrogen peroxide-independent production of α-alkenes by OleTJE P450 fatty acid decarboxylase. *Biotechol. Biofuels*. **2014**, *7*(1), 1-12.
63. Greule, A.; Stok, J. E.; De Voss, J. J.; Cryle, M. J. Unrivalled Diversity: The Many Roles and Reactions of Bacterial Cytochromes P450 in Secondary Metabolism. *Nat. Prod. Rep.* **2018**, *35*, 757– 791
64. Falk, E.; Gasser, V. C.; Morandi, B. Synthesis of N-Alkyl Anilines from Arenes via Iron-Promoted Aromatic C–H Amination. *Org. Let.* **2021**, *23*(4), 1422-1426.
65. Falk, E.; Makai, S.; Delcaillau, T.; Gürtler, L.; Morandi, B. Design and Scalable Synthesis of N-Alkylhydroxylamine Reagents for the Direct Iron-Catalyzed Installation of Medicinally Relevant Amines. *Angew. Chem.; Int. Ed.* **2020**, *59*(47), 21064-21071.

Chapter II

ENZYMATIC PRIMARY AMINATION OF BENZYLIC AND ALLYLIC C(SP^3)–H BONDS

Material from this chapter appears in: "Jia, Z. J.; **Gao, S.**; Arnold, F. H. * Enzymatic primary amination of benzylic and allylic C(sp^3)–H bonds. *J. Am. Chem. Soc.* **2020**, *142*(23), 10279–10283."

S. G. participated in the execution of the research, including enzymatic reactions, substrate scope study and synthetic applications. S. G. participated in reviewing and editing the manuscript.

ABSTRACT

Aliphatic primary amines are prevalent in natural products, pharmaceuticals, and functional materials. While a plethora of processes are reported for their synthesis, methods that directly install a primary amine group into C(sp^3)–H bonds remain unprecedented. Here, we report a set of new-to-nature enzymes that catalyzes the direct primary amination of C(sp^3)–H bonds with excellent chemo-, regio-, and enantioselectivity, using a readily available hydroxylamine derivative as the nitrogen source. Directed evolution of genetically-encoded cytochrome P411 enzymes (P450s whose Cys axial ligand to the heme iron has been replaced with Ser) generated variants that selectively functionalize benzylic and allylic C–H bonds, affording a broad scope of enantioenriched primary amines. This biocatalytic process is efficient and selective (up to 3930 TTN and 96% *ee*), and can be performed on preparative scale.

2.1. Introduction

Nitrogen is a key component of functional molecules: 80% of small-molecule drugs contain at least one nitrogen atom.[1] Particularly privileged are primary amines that serve as essential intermediates in the construction of secondary amines, tertiary amines, and heterocycles.[2] A plethora of transformations have been developed to introduce the primary amine moiety into organic structures, given its importance as a fundamental functional group.[2,3] Classical synthesis of primary amines usually relies on manipulation of functional groups, such as reduction of azides or nitriles, reductive amination of carbonyl compounds,[4] and Buchwald-Hartwig amination of aryl halides.[5] Although C–H functionalization has emerged as a promising strategy that maximizes atom- and step-economy, state-of-the-art methods for amine synthesis typically provide *N*-protected or *N*-substituted products that require additional manipulation to access the corresponding primary amines.[6] Recently, a few methods have been developed for primary amination of C(sp^2)–H bonds, including photoredox catalysis,[7] electrochemical catalysis,[8] and the use of other novel amination reagents and metal catalysts.[9] Primary amination of C(sp^3)–H bonds, however, has remained elusive. Biocatalytic approaches have been exploited to selectively introduce primary amine functionalities, as exemplified by transaminases,[10] but these all require pre-oxidized substrates and do not introduce the amine at a C–H bond.

Natural evolution has created diverse enzymes for selective C–H functionalization, ranging from hydroxylation to halogenation, under ambient conditions.[11] Among these, the direct hydroxylation of C(sp^3)–H bonds catalyzed by cytochrome P450 enzymes (P450s)

exemplifies the remarkable ability of enzymes to mediate transformations that are challenging for small-molecule catalysts. A high-valent iron oxo species, compound I, has been identified as the key intermediate.[12] The biocatalytic repertoire of P450s has been expanded through directed evolution to include the ability to generate and transfer abiotic nitrene and carbene intermediates.[13] We and Fasan engineered P450s to catalyze C–H amination via sulfonyl nitrene intermediates leading to the formation of diverse sulfonamides.[14] In spite of the prevalence and fundamental importance of aliphatic primary amines in the biological world, however, C(sp^3)–H primary amination is unknown in biology.

Here we addressed this challenge by directed evolution of a P450 from *Bacillus megaterium* ($P450_{BM3}$), which natively catalyzes C(sp^3)–H hydroxylation of fatty acids. We hypothesized that a high-valent unprotected iron-nitrenoid intermediate would allow for direct C(sp^3)–H primary amination, a process analogous to C–H hydroxylation through compound I. This iron-nitrenoid intermediate was demonstrated recently by us to enable aminohydroxylation of styrenyl olefins in an engineered cytochrome *c*.[15] Considering the highly tunable activities of $P450_{BM3}$ for C(sp^3)–H functionalization conferred by the protein scaffold,[12a,13] we envisioned that directed evolution could divert and subsequently amplify the reactivity of the unprotected iron-nitrenoid species to primary amination of C(sp^3)–H bonds. For this investigation we focused on engineered $P450_{BM3}$ variants that have the iron-coordinating cysteine residue replaced by serine (cytochrome 'P411s'). Substitution of the axial ligand increased the FeIII/FeII reduction potential and enabled the resulting P411s to perform C−H functionalization under physiological conditions without additional reductants.[16]

2.2. Results and Discussion

We chose tetrahydronaphthalene **1a** as the model substrate to test in a reaction with the putative iron-nitrenoid generated from hydroxylamine ester **2** (**Figure 2-1A**). Although C(sp^2)–H primary amination of substrate **1a** and its analogues is well established in synthetic chemistry,[9] primary amination of benzylic C(sp^3)–H bonds remains unknown. Indeed, test reactions of **1a** and **2** with free heme only produced C(sp^2)–H primary amination product **4a** (**Table A-1**). The transformation was then evaluated with a panel of P411 variants derived from $P450_{BM3}$. The enzymes were expressed and screened in whole *Escherichia coli* cells. To suppress dioxygen inhibition, the reactions were performed under anaerobic conditions. A variant named P411-B2, previously engineered for benzylic C–H amination via a sulfonyl nitrene,[14d] was identified to produce the desired primary benzylic amine **3a** with 8 total turnover number (TTN). Side product **4a** resulting from C(sp^2)–H primary amination was also detected, in a negligible amount (< 1 TTN). These findings provided a starting point to test our hypothesis that protein engineering could divert and amplify the reactivity of the putative unprotected iron-nitrenoid to C(sp^3)–H primary amination.

Variant P411-B2 was chosen as the starting template for directed evolution of an efficient *benzylic C–H primary aminase* ($P411_{BPA}$). Sequential rounds of site-saturation mutagenesis (SSM) and screening were performed to improve catalytic activity and enantioselectivity for the synthesis of benzylic amine **3a**. We referred to the crystal structure of a related P411 variant,[14d] and mainly chose amino acids proximal to the heme and/or residing on flexible

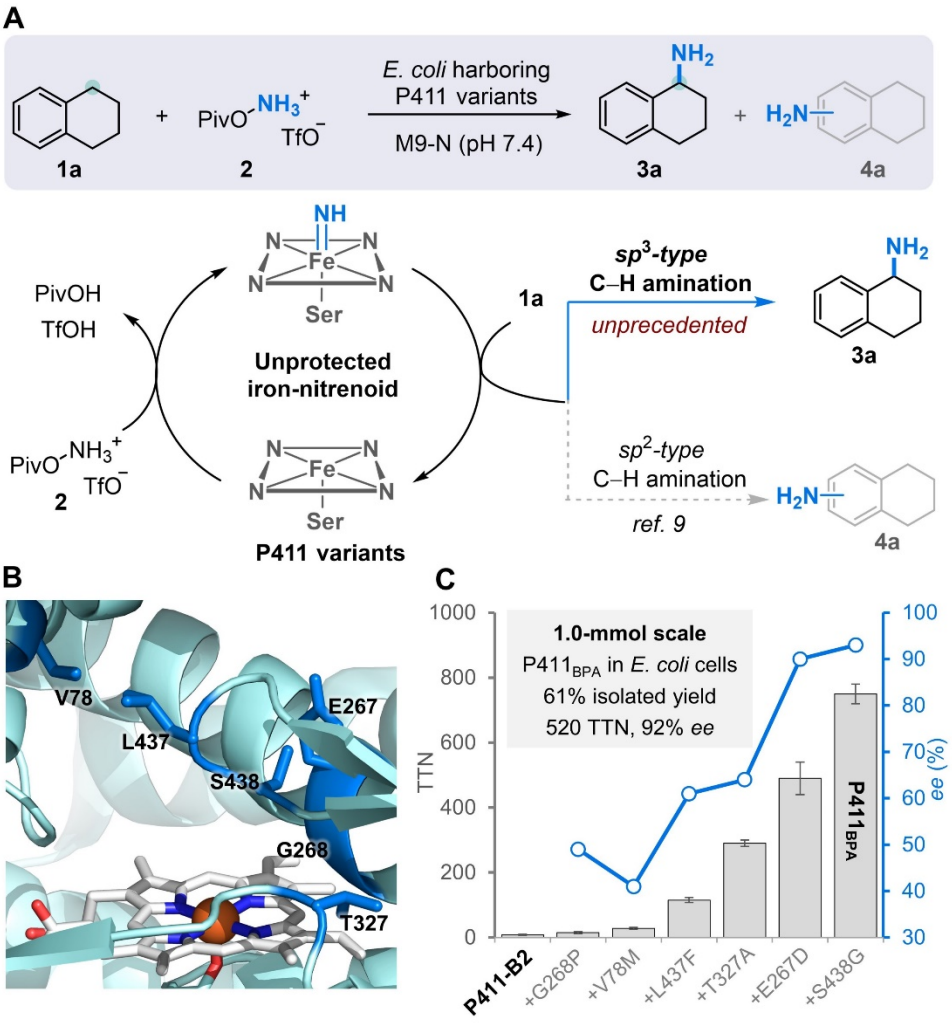


Figure 2-1. Enzymatic primary amination of benzylic C(sp^3)–H bonds. **(A)** Reaction scheme of benzylic C(sp^3)–H primary amination (Piv, pivaloyl, Tf, trifluoromethanesulfonyl). **(B)** Crystal structure of a variant (PDB ID: 5UCW) closely related to P411-B2 (Section A.8. in Supporting Information), with mutated residues marked in blue.[14d] **(C)** Evolutionary trajectory of *benzylic C–H primary aminase* ($P411_{BPA}$) for the synthesis of benzylic amine 3a and the scale-up reaction. $P411_{BPA}$ was evolved through six rounds of SSM and screening from P411-B2 (Table A-2). Indicated mutations are relative to P411-B2. Unless otherwise noted, reaction conditions were as follows: P411 variants in *E. coli* whole cells [optical density at 600 nm (OD_{600}) = 16], 2.0 mM substrate 1a, 4.0 mM substrate 2, 2.5 vol% EtOH in M9-N (pH 7.4) buffer, 800 µL reaction volume at 10 °C under anaerobic conditions for 12 h. The enantiomeric excess (*ee*) for P411-B2 was not determined due to the low yield. See the Supporting Information for further details of the scale-up reaction.

loops for mutagenesis (**Figure 2-1B**). In each round of SSM, enzyme libraries were generated and screened by HPLC for product formation in 96-well plates in the form of whole-cell catalysts. Four rounds of SSM and screening introduced mutations G268P, V78M, L437F, and T327A, leading to an increase of *ee* to 64% and a 36-fold improvement of TTN to 290 (**Figure 2-1C**). Subsequent rounds introduced mutations E267D and S438G, leading to final variant $P411_{BPA}$ and further boosting *ee* to 93% and TTN to 750 (**Table A-2**). The formation of side product **4** remained negligible during directed evolution. Notably, the HPLC yield of **3a** in reactions at analytical scale was as high as 85%, and primary amine **3a** could be isolated in 61% yield and with 92% *ee* by using simple acid-base extraction (1.0 mmol reaction scale). Although the biocatalytic process is scalable, the enzymes are sensitive to substrate concentration and higher concentration of hydroxylamine ester **2** usually leads to a decreased yield.

With $P411_{BPA}$ in hand, we evaluated the substrate scope of alkanes, systematically targeting primary, secondary, and tertiary benzylic C–H bonds for primary C–H amination (**Figure 2-2**). For cyclic substrates bearing secondary benzylic C–H bonds, ring sizes ranging from four to six members were well tolerated, affording primary amines **3a–3c**. Linear substrates with structurally and electronically diverse aryl groups proved to be compatible and delivered amines **3d–3l** with excellent enantioselectivity. In addition to secondary benzylic C–H bonds (**3d–3j**), primary benzylic C–H bonds (**3i–3l**) and tertiary benzylic C–H bonds (**3k**) were also suitable with varying levels of efficiency and regioselectivity, highlighting the generality of this enzymatic transformation. Reaction at secondary and tertiary benzylic C–H bonds was more favorable than at their primary counterparts, as demonstrated by

products **3i**, **3j**, and **3k**. Amination of larger substrates than those shown here or substrates with heteroatoms typically occurred with diminished efficiency, and further directed evolution based on P411$_{BPA}$ may be required for such substrates (**Figure A-1**).

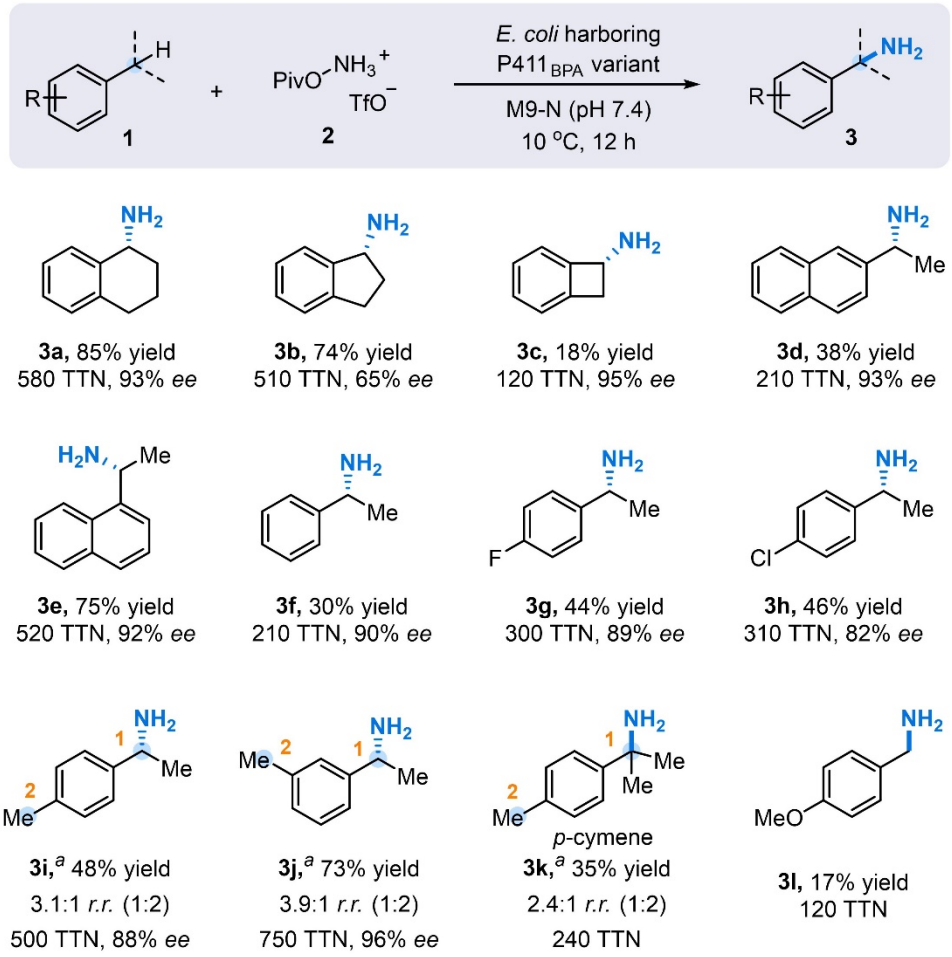


Figure 2-2. Scope of benzylic amine products. Experiments were performed at analytical scale using enzyme-expressing *E. coli* cells resuspended to OD_{600} = 16 or 24 in M9-N medium (pH 7.4), 2.0 mM substrate 1, 4.0 mM substrate 2, 2.5 vol% EtOH, and 800 µL reaction volume at 10 °C under anaerobic conditions for 12 hours(Table A-3). Regioisomeric ratio (*r.r.*) indicates the mole ratio of major product to combined minor regioisomers. Yields were quantified by HPLC based on the calibration curve of the corresponding reference products. [a] Yields and TTNs were calculated based on all regioisomers.

Having achieved primary amination of benzylic C–H bonds, we next asked whether the enzymes could aminate other types of C(sp^3)–H bonds, particularly in the presence of sensitive olefins. With hydroxylamine derivatives as nitrogen sources and small-molecule catalysts, alkene substrates were reported to undergo olefin functionalization.[17] As expected, we observed only amino alcohol **7a** in the reaction of model substrate **5a** and hydroxylamine **2** with free heme as the catalyst (**Table A-1**). Enzymes, however, often feature chemoselectivity in biochemical processes that can be tuned by protein engineering.[18] Therefore, we envisioned that appropriate engineering of $P411_{BPA}$-related enzymes would enable selective functionalization of allylic C–H bonds while minimizing the undesired olefin side reactions. Screening revealed that $P411_{BPA}$-related variants had the desired allylic C–H primary amination activity (**Figure 2-3A**). Byproduct **7a** was only formed in trace amounts (**Table A-4**), and anilines resulting from C(sp^2)–H primary amination were not observed. Starting from an intermediate variant $P411_{BPA'}$ having five mutations from P411-B2, we performed two rounds of SSM and screening to generate *allylic C–H primary aminase* ($P411_{APA}$) with two additional mutations (N395R, S438A). $P411_{APA}$ is exceptionally efficient and selective (3930 TTN, 93% HPLC yield, and 94% *ee*), and the enzymatic process is easily scalable, with 75% isolated yield (1.0 mmol reaction scale).

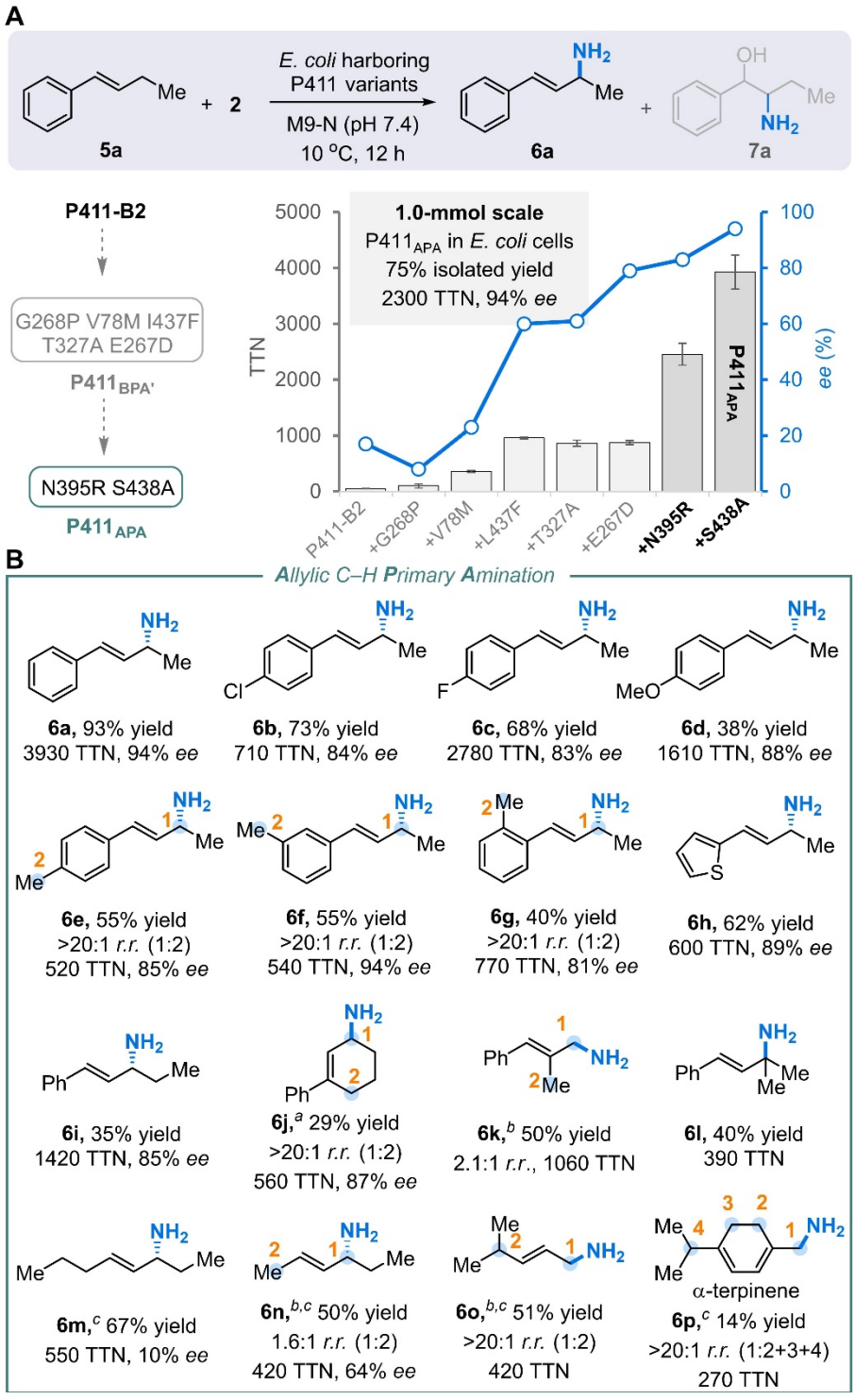


Figure 2-3. Engineering enzymes for allylic C–H primary amination and the substrate scope. **(A)** Evolutionary trajectory of *allylic C–H primary aminase* ($P411_{APA}$) for the

synthesis of allylic amine 6a and the scale-up reaction. P411$_{APA}$ was evolved through two more rounds of SSM and screening starting from an intermediate benzylic C–H aminase having five mutations with respect to P411-B2 (Table A-4). The indicated mutations are relative to P411-B2. **(B)** Substrate scope of allylic C–H primary amination (Table A-5). Yields were quantified by HPLC based on the calibration curve of the corresponding reference products. [a] Absolute configuration of 6j could not be determined. [b] Yields and TTNs were calculated based on all regioisomers. [c] Products were characterized and quantified by their benzoyl-protected amines.

P411$_{APA}$ is capable of functionalizing primary, secondary, and tertiary allylic C–H bonds with a broad substrate scope (**Figure 2-3B**). For styrene-type substrates bearing secondary allylic C–H bonds, phenyl groups with varying substituents were tolerated well to afford allylic amines **6a–6g** with excellent activity and enantioselectivity. In the presence of primary benzylic C–H bonds, allylic C–H bonds were functionalized with complete regioselectivity (**6e–6g**). Substrates with heterocyclic substituents such as thiophenyl (**6h**) and longer aliphatic substituents (**6i**) also proved to be suitable. Single regioisomer **6j** was exclusively afforded, indicating P411$_{APA}$ can discriminate between highly similar secondary allylic C–H bonds. Tertiary and primary allylic C–H bonds could also be aminated efficiently to give products **6k** and **6l**. Notably, aromatic groups on the alkene moiety are not necessary for allylic C–H amination activity, and products **6m–6p** were successfully synthesized. For substrates with different types of reactive C–H bonds, P411$_{APA}$ exhibits high levels of regioselectivity. While 2-hexene gave a mixture of regioisomers (**6n**), single regioisomers were observed for other alkenes with up to four sets of allylic C–H bonds (**3o**, **3p**), such as α-terpinene. Although the regioselectivity between secondary and primary C–H bonds is dependent on substrate patterns (**6n**, **6p**), P411$_{APA}$ exclusively functionalized primary C–H bonds over weaker tertiary C–H bonds (**6o**, **6p**), which is contrary to the regioselectivity of

P411$_{BPA}$ (**3k**). Similar to P411$_{BPA}$, further directed evolution of P411$_{APA}$ may be required for efficient functionalization of terminal or larger alkene substrates, especially complex natural products (**Figure A-1**).

In conclusion, we have engineered *C(sp^3)–H primary aminases* that constitute a general platform for selective primary amination of primary, secondary, and tertiary C–H bonds at benzylic and allylic positions. These laboratory-evolved enzymes provide a new approach to accessing aliphatic primary amines with high chemoselectivity, regioselectivity, and enantioselectivity. Considering the fundamental roles of primary amines in the biological world and the natural occurrence and biocompatibility of hydroxylamine derivatives, we anticipate that these new, fully genetically-encoded enzymes will provide a starting point to extend or even reformulate currently mapped nitrogen metabolism. Such activities may, in fact, already exist in nature, whose vast catalytic capabilities have only been partially explored. Further elucidation of the catalytic mechanism and characterization of the key intermediate would be valuable to generalize this process and inspire the design of corresponding small-molecule catalysts.

Chapter II Bibliography

1. (a) Hili, R.; Yudin, A. K. Making Carbon-Nitrogen Bonds in Biological and Chemical Synthesis. *Nat. Chem. Biol*. **2006**, *2*, 284– 287. (b) Vitaku, E.; Smith, D. T.; Njardarson, J. T. Analysis of the Structural Diversity, Substitution Patterns, and Frequency of Nitrogen Heterocycles among U.S. FDA Approved Pharmaceuticals. *J. Med. Chem.* **2014**, *57*, 10257–10274.
2. *Amino Group Chemistry: From Synthesis to the Life Sciences*; Ricci, A., Ed.; Wiley-VCH: Weinheim, Germany, **2008**.
3. Legnani, L.; Bhawal, B. N.; Morandi, B. Recent Developments in the Direct Synthesis of Unprotected Primary Amines. *Synthesis* **2017**, *49*, 776–789.
4. Gross, T.; Seayad, A. M.; Ahmad, M.; Beller, M. Synthesis of Primary Amines: First Homogeneously Catalyzed Reductive Amination with Ammonia. *Org. Lett*. **2002**, *4*, 2055–2058.
5. (a) Hartwig, J. F. Evolution of a Fourth Generation Catalyst for the Amination and Thioetherification of Aryl Halides. *Acc. Chem. Res.* **2008**, *41*, 1534–1544. (b) Surry, D. S.; Buchwald, S. L. Biaryl Phosphane Ligands in Palladium-Catalyzed Amination. *Angew. Chem, Int. Ed.* **2008**, *47*, 6338–6361.
6. Park, Y.; Kim, Y.; Chang, S. Transition Metal-Catalyzed C–H Amination: Scope, Mechanism, and Applications. *Chem. Rev*. **2017**, *117*, 9247–9301.
7. (a) Romero, N. A.; Margrey, K. A.; Tay, N. E.; Nicewicz, D. A. Site-Selective Arene C–H Amination via Photoredox Catalysis. *Science* **2015**, *349*, 1326–1330. (b) Zheng, Y.-W.; Chen, B.; Ye, P.; Feng, K.; Wang, W.; Meng, Q.-Y.; Wu, L.-Z.; Tung, C.-H. Photocatalytic Hydrogen-Evolution Cross-Couplings: Benzene C–H Amination and Hydroxylation. *J. Am. Chem. Soc.* **2016**, *138*, 10080–10083.
8. Morofuji, T.; Shimizu, A.; Yoshida, J.-I. Electrochemical C–H Amination: Synthesis of Aromatic Primary Amines via *N*-Arylpyridinium Ions. *J. Am. Chem. Soc*. **2013**, *135*, 5000–5003.
9. (a) Peng, J.; Chen, M.; Xie, Z.; Luo, S.; Zhu, Q. Copper-Mediated C(sp^2)–H Amination Using $TMSN_3$ as a Nitrogen Source: Redox-Neutral Access to Primary Anilines. *Org. Chem. Front*. **2014**, *1*, 777–781. (b) Tezuka, N.; Shimojo, K.; Hirano, K.; Komagawa, S.; Yoshida, K.; Wang, C.; Miyamoto, K.; Saito, T.; Takita, R.; Uchiyama, M. Direct Hydroxylation and Amination of Arenes via Deprotonative Cupration. *J. Am. Chem. Soc*. **2016**, *138*, 9166–9171. (c) Paudyal, M. P.; Adebesin, A. M.; Burt, S. R.; Ess, D. H.; Ma, Z.; Kürti, L.; Falck, J. R. Dirhodium-Catalyzed C–H Arene Amination Using Hydroxylamines. *Science* **2016**, *353*, 1144–1147. (d) Legnani, L.; Prina Cerai, G.; Morandi, B. Direct and Practical Synthesis of Primary Anilines through Iron-Catalyzed C–H Bond Amination. *ACS Catal*. **2016**, *6*, 8162–8165. (e) Liu, J.; Wu, K.; Shen, T.; Liang, Y.; Zou, M.; Zhu, Y.; Li, X.; Li, X.; Jiao, N. Fe-Catalyzed Amination of (Hetero)Arenes with a Redox-Active Aminating Reagent under Mild Conditions. *Chem. Eur. J*. **2017**, *23*, 563–567. (f) Kim, H.; Heo, J.; Kim, J.; Baik, M.-H.; Chang, S. Copper-Mediated Amination of Aryl C–H Bonds with the Direct Use of Aqueous Ammonia via a Disproportionation Pathway. *J. Am. Chem. Soc.* **2018**, *140*, 14350–14356. (*g*) D'Amato, E. M.; Börgel, J.; Ritter, T.

Aromatic C–H Amination in Hexafluoroisopropanol. *Chem. Sci*. **2019**, *10*, 2424–2428. (h) Anugu, R. R.; Munnuri, S.; Falck, J. R. Picolinate-Directed Arene *meta*-C–H Amination via $FeCl_3$ Catalysis. *J. Am. Chem. Soc*. **2020**, *142*, 5266–5271.

10. (a) Slabu, I.; Galman, J. L.; Lloyd, R. C.; Turner, N. J., Discovery, Engineering, and Synthetic Application of Transaminase Biocatalysts. *ACS Catal*. **2017**, *7*, 8263–8284. (b) Guo, F.; Berglund, P., Transaminase Biocatalysis: Optimization and Application. *Green Chem*. **2017**, *19*, 333–360. (c) Zhang, J.; Abidin, M. Z.; Saravanan, T.; Poelarends, G. J., Recent Applications of Carbon-Nitrogen Lyases in Asymmetric Synthesis of Noncanonical Amino Acids and Heterocyclic Compounds. *ChemBioChem* **2020**, DOI: 10.1002/cbic.202000214.
11. Lewis, J. C.; Coelho, P. S.; Arnold, F. H. Enzymatic Functionalization of Carbon–Hydrogen Bonds. *Chem. Soc. Rev.* **2011**, *40*, 2003–2021.
12. (a) Ortiz de Montellano, P. R. Hydrocarbon Hydroxylation by Cytochrome P450 Enzymes. *Chem. Rev*. **2010**, *110*, 932–948. (b) Rittle, J.; Green, M. T. Cytochrome P450 Compound I: Capture, Characterization, and C-H Bond Activation Kinetics. *Science* **2010**, *330*, 933–937.
13. (a) 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. (b) Fasan, R.; Kan, S. B.; Zhao, H. A Continuing Career in Biocatalysis: Frances H. Arnold. *ACS Catal*. **2019**, *9*, 9775–9788. (c) Chen, K.; Arnold, F. H. Engineering New Catalytic Activities in Enzymes. *Nat. Catal.* **2020**, *3*, 203–213. (d) Singh, R.; Kolev, J. N.; Sutera, P. A.; Fasan, R. Enzymatic C(sp^3)-H Amination: P450-Catalyzed Conversion of Carbonazidates into Oxazolidinones. *ACS Catalysis* **2015**, *5*, 1685–1691. (e) 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. (f) Dydio, P.; Key, H. M.; Hayashi, H.; Clark, D. S.; Hartwig, J. F. Chemoselective, Enzymatic C–H Bond Amination Catalyzed by a Cytochrome P450 Containing an Ir(Me)-PIX Cofactor. *J. Am. Chem. Soc*. **2017**, *139*, 1750–1753.
14. (a) McIntosh, J. A.; Coelho, P. S.; Farwell, C. C.; Wang, Z. J.; Lewis, J. C.; Brown, T. R.; Arnold, F. H. Enantioselective Intramolecular C–H Amination Catalyzed by Engineered Cytochrome P450 Enzymes in vitro and in vivo. *Angew. Chem, Int. Ed.* **2013**, *52*, 9309–9312. (b) Hyster, T. K.; Farwell, C. C.; Buller, A. R.; McIntosh, J. A.; Arnold, F. H. Enzyme-Controlled Nitrogen-Atom Transfer Enables Regiodivergent C–H Amination. *J. Am. Chem. Soc.* **2014**, *136*, 15505–15508. (c) Singh, R.; Bordeaux, M.; Fasan, R. P450-Catalyzed Intramolecular sp^3 C–H Amination with Arylsulfonyl Azide Substrates. *ACS Catalysis* **2014**, *4*, 546–552. (d) 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. (e) 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.

15. Cho, I.; Prier, C. K.; Jia, Z.-J.; Zhang, R. K.; Görbe, T.; Arnold, F. H. Enantioselective Aminohydroxylation of Styrenyl Olefins Catalyzed by an Engineered Hemoprotein. *Angew. Chem, Int. Ed.* **2019**, *58*, 3138–3142.
16. Coelho, P. S.; Wang, Z. J.; Ener, M. E.; Baril, S. A.; Kannan, A.; Arnold, F. H.; Brustad, E. M. A Serine-Substituted P450 Catalyzes Highly Efficient Carbene Transfer to Olefins i*n vivo*. *Nat. Chem. Biol.* **2013**, *9*, 485–487.
17. (a) Jat, J. L.; Paudyal, M. P.; Gao, H. Y.; Xu, Q. L.; Yousufuddin, M.; Devarajan, D.; Ess, D. H.; Kurti, L.; Falck, J. R.

Appendix A
SUPPLEMENTARY INFORMATION FOR CHAPTER II

A.1. General Procedures

A.1.1 General

Unless otherwise noted, all chemicals and reagents were obtained from commercial suppliers (Millipore Sigma, VWR, TCI America and Fischer Scientific) and were used without further purification. NMR spectra of chemicals in $CDCl_3$ were obtained using a Bruker Prodigy 400 MHz spectrometer, and were referenced to residual solvent signals. Data for ^1H NMR are reported in the conventional form: chemical shift (δ ppm), multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, hept = heptet, m = multiplet, br = broad), coupling constant (Hz), and integration. Sonication was performed using a Qsonica Q500 sonicator. High-resolution mass spectra HRMS were acquired from the Caltech Mass Spectral Facility using electrospray ionization (TOF ES+). Optical rotations were measured on a Jasco P-2000 polarimeter using a 100 mm path-length cell at 589 nm. Chemical reactions were monitored using thin layer chromatography (Merck 60 silica gel plates) and a UV lamp for visualization, if possible. Reverse-phase high-performance liquid chromatography (HPLC) and liquid chromatography-mass spectroscopy (LC-MS) for analysis were carried out using Agilent 1200 series instruments, with C8 (Agilent Poroshell 120, 4.6 × 50 mm, 4 μm) and C18 (Kromasil®, 4.6 × 50 mm, 5 μm) column. Water and acetonitrile containing 0.1% acetic acid were used as eluents. Normal-phase

chiral HPLC was performed using Daicel Chiralpak IB, IC, and Daicel Chiralcel OJ-H columns (all 4.6 × 250 mm, 5 μm) with hexane and isopropanol as the mobile phase.

A.1.2. Cloning, mutagenesis, and expression of enzymes

Expression vector pET22b(+) (Novagen) was used for cloning and expression of all variants described in this paper. Site-saturation mutagenesis was performed using a modified QuikChange™ mutagenesis protocol using the 22-codon trick.[1] The PCR products were digested with *Dpn*I, purified with New England Biolabs gel purification kit, and the gaps were repaired using Gibson Mix™.[2] Without further purification, 1 μL of the Gibson product was used to transform 50 μL of electrocompetent *Escherichia coli* BL21 E. cloni (Lucigen) cells. BL21 E. cloni cells transformed with pET22b(+) constructs encoding various $P411_{BM3}$ variants were grown overnight in 5-mL Luria-Bertani medium supplemented with ampicillin (LB_{amp}). Subsequently, 2 mL of this preculture were used to inoculate 50 mL of Hyperbroth medium supplemented with ampicillin (HB_{amp}) for expression. The expression culture was incubated at 37 °C and shaken at 220 rpm for 2.5 hours. Then, the expression culture was cooled on an ice bath for 20 minutes and was treated with 1 mM 5-aminolevulinic acid (ALA) and 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) (final concentrations). Cells were expressed at 22 °C and 130 rpm for 20–24 hours, and the shaking radius was 25 mm. When the shaking radius was 50 mm, the expression culture was shaken at 140 rpm before induction and 90 rpm after induction. Once expression was finished, the cultures were centrifuged (4,000 g, 3 minutes,

and 4 °C) and the pellets were resuspended to an optical density at 600 nm (OD_{600}) of 32 in M9-N minimal medium with pH adjusted to 7.4.

A.1.3. Determination of hemeprotein concentration

Protein concentration in the cell was determined by performing hemochrome assay on the cell lysate[3] Lysate was obtained by sonication (4 minutes, 1 second on, 1 second off, 30 % intensity, on ice). The cell debris was removed by centrifugation (14,000 g, 30 minutes, 4 °C). 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 (380–620 nm) of the oxidized Fe^{III} state was recorded immediately. Sodium dithionite (10 µL of 0.5 M solution in water) 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 following extinction coefficients: 34.7 mM^{-1} cm^{-1} at 557 nm.

A.1.4. Analytic reaction setup and product quantification

M9-N medium (pH 7.4) and D-glucose solution (500 mM in M9-N, pH 7.4) were placed in the anaerobic chamber with oxygen concentration below 20 ppm at least 24 hours. Harvested cells were resuspended with M9-N buffer (pH 7.4) in coy chamber to OD_{600} = 32, and a specific volume of resuspended cells were aliquoted to 2-mL screw cap vials. The screw cap vials with cells were precooled to 0 ℃ on an ice bath. Unless otherwise specified, 60 µL of the precooled glucose solution, 20 µL of a stock solution[4] containing

glucose oxidase (from *Aspergillus niger*, 1,000 U/mL) and catalase (from Bovine liver, 14,000 U/mL) in double-distilled water, and M9-N medium were added to make the total volume to 752 mL. The resulting mixtures stayed on the ice bath for another 10 minutes, and then the hydrocarbon substrate **1** or **5** (16 μL, 0.1 M stock in ethanol) and the nitrene precursor **2** (32 μL, 0.1 M stock in water) were added in a sequential manner. Unless otherwise noted, the reactions were then shaken at 10 ℃ for 12 hours at 250 rpm. Depending on UV sensitivity and yield of the products, two work-up procedures were performed after the reactions were completed.

For benzylic amines **3**, allylic products **6a**–**6l**, amino alcohol **7a**, 0.38 mL ethanol and 0.02 mL a specified internal standard in ethanol was added and the mixtures were subsequently transferred to 1.7-mL Eppendorf microcentrifuge tubes. After vortexing the mixtures, the microcentrifuge tubes were centrifuged at 14,000 g for 5 minutes to precipitate coagulated proteins and cell debris. To identify the products and determine yields, 400 μL of the supernatants were transferred to an HPLC vial with an insert and analyzed by reverse-phase HPLC-MS. Products were identified and quantified based on the corresponding reference compounds (**Section A.4., A-5 and A-6**). To further determine the enantiomeric excess (*ee*), the remaining supernatants were combined and transferred to a 4-mL vial and ethanol was evaporated. To the remaining supernatants were added 5 μL benzoyl chloride and 90 μL saturated $NaHCO_3$ solution. After shaking the vials overnight, the benzoyl-protected products were extracted to a solution of hexane and ethyl acetate mixture (1:1) and subjected to normal-phase HPLC to determine the *ee*.

For products **6m**–**6p**, 5 μL benzoyl chloride and 90 μL saturated $NaHCO_3$ solution were added directly when reactions were completed. The reaction mixtures were shaken overnight, and 0.38 mL ethanol and 0.02 mL of a specified internal standard in ethanol was added. After vortexing, coagulated proteins and cell debris were removed by centrifugation (14,000 g, 5 minutes). The benzoyl-protected products were analyzed by reverse-phase HPLC-MS to determine the yield and by normal-phase HPLC to determine the *ee*.

A.1.5. Reaction screening in 96-well plate in whole-cell format

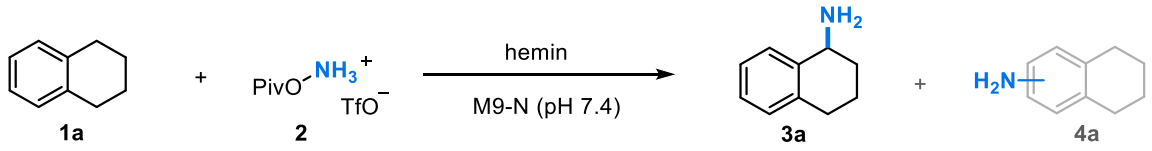
After a single site-saturation library was generated, 88 single colonies were randomly picked and cultured in 300 μL of LB medium with 0.1 mg/mL ampicillin (LB_{amp}) in a sterilized 96-well culture plate. The plate typically contained six wells inoculated with single colonies expressing the parent enzyme, and two sterile wells. The cells in LB medium were cultured at 37 °C, 250 rpm, and 80% relative humidity for 8–12 hours. A separate sterilized 96-well culture plate was filled with 1,000 μL of Hyperbroth medium containing 0.1 mg/mL ampicillin (HB_{amp}) in each well. The plate with HB_{amp} was inoculated with the LB preculture (50 μL/well), and incubated at 37 °C, 250 rpm, and 80% relative humidity for 2.5 hours. The plate was cooled on an ice bath for 1 hour, induced with 0.5 mM IPTG and 1 mM 5-aminolevulinic acid (final concentrations), and then expressed at 22 °C and 220 rpm for 20–22 hours. The cells were pelleted (4,000 g, 3 minutes), and 300 μL of M9-N (pH 7.4) together with 60 μL of D-glucose solution (500 mM in M9-N, pH 7.4) were added to each well. After cells were fully resuspended by

shaking at 500 rpm, the 96-well plate was transferred to an anaerobic chamber. Subsequently, the hydrocarbon substrate **1** or **5** (20 μL, 0.1 M stock in ethanol) and the nitrene precursor **2** (20 μL, 0.1 M stock in water) were added. The plate was sealed with an aluminum foil and shaken at 600 rpm in the anaerobic chamber at room temperature overnight.

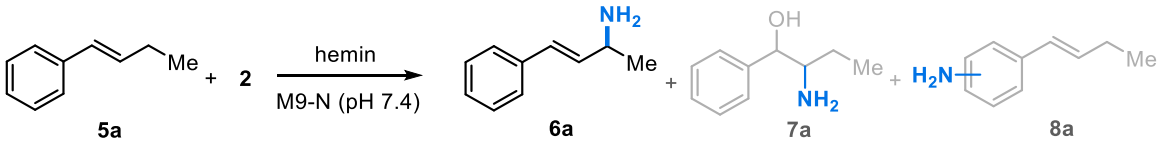
Once the plate was taken out of the anaerobic chamber and the seal was removed, ethanol (800 μL/well) was added. The resulting suspension in the wells was mixed by pipetting, and the plate was left to shake at 600 rpm at room temperature for an hour for additional mixing. The plate was then centrifuged (4,500 g, 5 minutes) 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-MS analysis to determine the yield.

A.2. Supporting Tables and Figures

Table A-1. Hemin catalyzed reaction of **1a** or **5a** with **2**]



Catalysts	Yield [%]	
	3a	**4a**
Hemin[(a)]	N. D.	< 0.1% yield
Hemin + BSA[(b)]	N. D.	< 0.1% yield



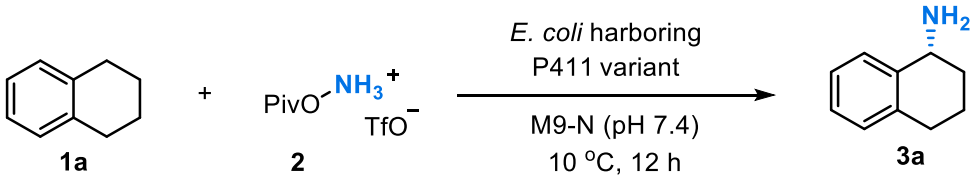
Catalysts	Yield [%]		
	6a	**7a**	**8a**
Hemin[(a)]	N. D.	< 0.1% yield	N. D.
Hemin + BSA[(b)]	N. D.	< 0.1% yield	N. D.

Reactions were performed in an anaerobic chamber with oxygen concentration below 20 ppm using indicated small-molecule catalysts, 4 mM substrate **1a**, 8 mM substrate **2**, 400 µL final reaction volume at room temperature for 12 h. Products were identified by HPLC-MS based on the corresponding reference compounds (**Section A.5.**). N. D. = none detected.

(a) To 332 µL M9-N medium (pH 7.4) were added 10 µL hemin (1 mM in DMSO), 10 µL sodium hydrosulfite solution (10 mg/mL), 16 µL substrate **1a** or **5a** (100 mM in EtOH), and 32 µL substrate **2** (100 mM in water) in a sequential manner.

(b) To 322 µL M9-N medium (pH 7.4) were added 10 µL hemin solution (1 mM in DMSO), 10 µL bovine serum albumin (BSA, 1 mM in water), 10 µL sodium hydrosulfite solution (10 mg/mL), 16 µL substrate **1a** or **5a** (100 mM in EtOH), and 32 µL substrate **2** (100 mM in water) in a sequential manner.

Table A-2. Summary of directed evolution for benzylic C–H primary aminase (**Fig. 3A**).



Variant[(a)]	Yield [%]	TTN	*ee* [%]
P411-B2	1.0 ± 0.2	8 ± 2	N. D.
+ G268P	1.7 ± 0.4	15 ± 3	49
+ V78M	3.0 ± 0.3	28 ± 3	– 41[(b)]
+ L437F	9.3 ± 1.0	115 ± 8	61
+ T327A	29.0 ± 0.3	290 ± 8	64
+ E267D	56.1 ± 4.0	490 ± 50	90
+ S438G (**P411$_{BPA}$**)	73.1 ± 2.0	750 ± 30	93

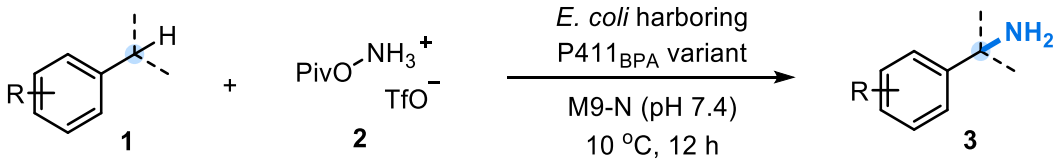
Reactions for site-saturation mutagenesis and screening were performed in the 96-well format (**Section A.1.**). Identified variants with beneficial mutations were then tested in reaction conditions as follows: P411s variants in *E. coli* whole cells (OD_{600} = 16), 2.0 mM

substrate **1a**, 4.0 mM substrate **2**, 2.5 vol % EtOH in M9-N (pH 7.4) buffer. Reaction volume was 800 µL in 2-mL screw cap vials, and the reactions proceeded at 10 °C under anaerobic conditions for 12 h. Product formation was quantified by HPLC-MS based on the calibration curve of the corresponding racemic reference compound (**Section A.6.**), and *ee* was measured with normal phase HPLC. Results are the average of four reactions (performed from two independent cell cultures, each used for duplicate reactions). N. D. = none determined due to a very low yield.

(a) The indicated mutations are relative to the closet parent. See the detailed mutations relative to wild-type $P450_{BM3}$ in **Table A-6** and the sequence of final variant $P411_{BPA}$ in **Section A.8.**.

(b) The *S*-enantiomer was obtained instead.

Table A-3. Substrate scope of benzylic primary amines.



Product	OD_{600}	Yield [%]	*r.r.*	TTN	*ee* [%]
3a	24	85.0 ± 4.0	-	580 ± 30	93
3b	24	74.0 ± 6.0	-	510 ± 40	65
3c	24	18.0 ± 2.0	-	120 ± 10	95
3d	24	38.0 ± 2.0	-	210 ± 80	93
3e	24	75.0 ± 6.0	-	520 ± 40	92
3f	24	30.0 ± 2.0	-	210 ± 10	90

3g	24	44.0 ± 4.0	-	300 ± 30	89
3h	24	46.0 ± 5.0	-	310 ± 30	82
3i	16	48.0 ± 2.0	3.1:1	500 ± 10	88
3j	16	73.0 ± 8.0	3.9:1	750 ± 70	96
3k	24	35.0 ± 2.0	2.4:1	240 ± 10	-
3l	24	17.0 ± 1.0	-	120 ± 10	-

Reaction condition: $P411_{BPA}$ in *E. coli* whole cells (OD_{600} = 16 or 24), 2.0 mM substrate **1**, 4.0 mM substrate **2**, 2.5 vol % EtOH in M9-N (pH 7.4) buffer. Reaction volume was 800 µL in 2-mL screw cap vials, and the reactions proceeded at 10 ℃ under anaerobic conditions for 12 h. Product formation was quantified by HPLC-MS based on the calibration curve of the corresponding racemic reference compound (**Section A.6.**), and *ee* was determined by normal-phase chiral HPLC. Results are the average of four reactions (performed from two independent cell cultures, each used for duplicate reactions). The absolute configurations of **3a** and **3i** were determined by comparison to chiral standards (**Section A.7.**).

Figure A-1. Less reactive (< 1% yield) and unreactive substrates .

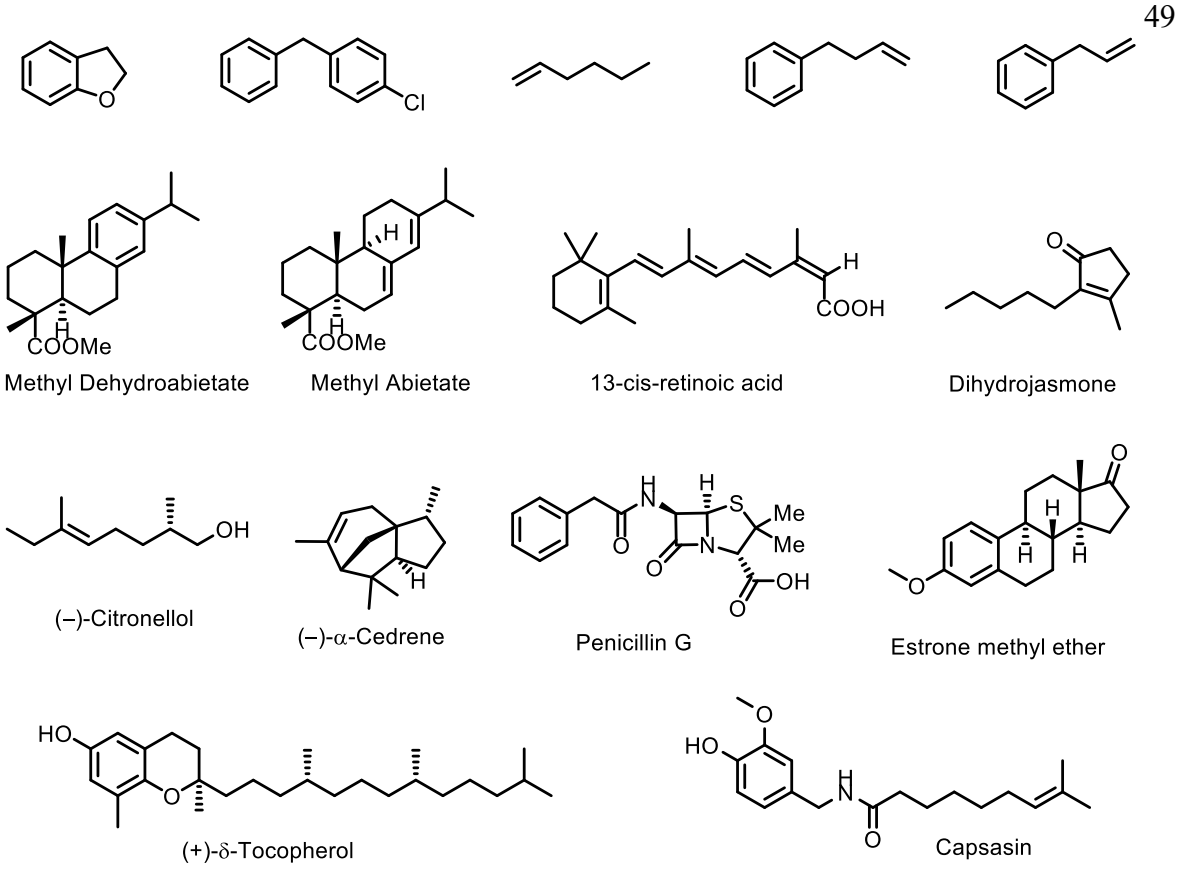
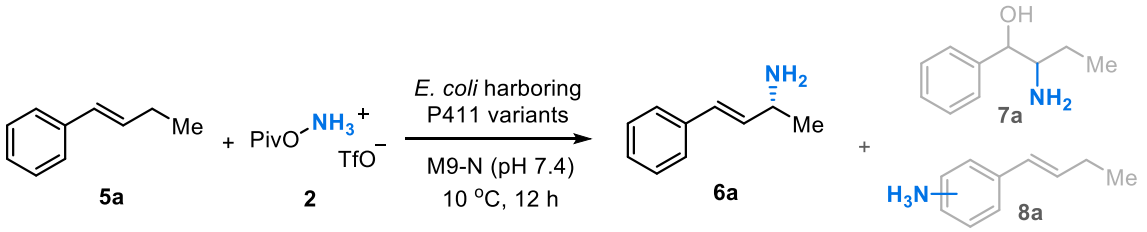


Table A-4. Summary of directed evolution for P411$_{APA}$.



Variant[a]	6a			7a	8a
	Yield [%]	TTN	*ee* [%]	Yield [%]	Yield [%]
P411-B2	1.5 ± 0.1	51 ± 2	–17[b]	0.8 ± 0.1	N. D.
+ G268P	3.0 ± 1.0	100 ± 40	–18[b]	0.9 ± 0.1	N. D.
+ V78M	9.5 ± 0.5	360 ± 20	–23[b]	0.7 ± 0.1	N. D.
+ L437F	19.0 ± 1.0	960 ± 10	60	0.9 ± 0.2	N. D.

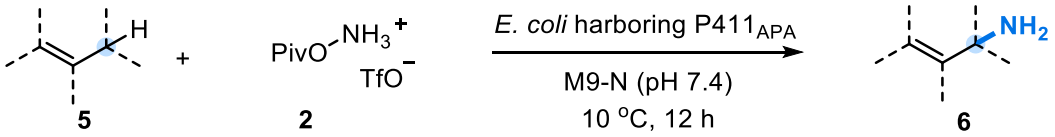
+ T327A	21.6 ± 0.9	870 ± 50	61	1.6 ± 0.1	N. D.
+ E267D	25.1 ± 0.5	870 ± 40	82	2.3 ± 0.2	N. D.
+ N395R	54.0 ± 3.0	2450 ± 200	82	3.1 ± 0.3	N. D.
+ S438A (**P411$_{APA}$**)	93.0 ± 2.0	3930 ± 310	94	3.7 ± 0.5	N. D.

Reactions for site-saturation mutagenesis and screening were performed in the 96-well format (**Section A.1.**). Identified variants with beneficial mutations were then tested in reaction conditions as follows: P450s variants in *E. coli* whole cells (OD_{600} = 4.0), 2.0 mM substrate **5a**, 4 mM substrate **2**, 2.5 vol % EtOH in M9-N (pH 7.4) buffer. Reaction volume was 800 μL in 2-mL vials with screw cap, and the reactions proceeded at 10 °C under anaerobic conditions for 12 h. The formation of product **6a** and side product **7a** was quantified by HPLC-MS based on the calibration curve of the corresponding racemic reference compound (**Section A.6.**), and *ee* was measured with normal phase HPLC. Results are the average of four reactions (performed from two independent cell cultures, each used for duplicate reactions). N. D. = not detected.

(a) The indicated mutations are relative to the closet parent. See the detailed mutations relative to wild-type P450$_{BM3}$ in **Table A-6** and the sequence of final variant P411$_{APA}$ in **Section A.8.**.

(b) The *S*-enantiomer was obtained.

Table A-5. Substrate scope of allylic amines.



Product	OD_{600}	Yield [%]	*r.r.*	TTN	*ee* [%]
6a[(a)]	4	93.0 ± 2.0	-	3930 ± 310	94
6b	16	73.0 ± 3.0	-	710 ± 30	84
6c	4	68.0 ± 5.0	-	2780 ± 300	83
6d	4	38.0 ± 4.0	-	1610 ± 100	88
6e	16	55.0 ± 2.0	> 20:1	520 ± 30	85
6f	16	55.0 ± 12.0	> 20:1	540 ± 100	94
6g	8	40.0 ± 4.0	> 20:1	770 ± 60	81
6h	16	62.0 ± 11.0	-	600 ± 120	89
6i	4	35.0 ± 2.0	-	1420 ± 120	85
6j[(a)]	8	29.0 ± 1.0	> 20:1	560 ± 30	87
6k	8	50.0 ± 2.0	2.1:1	1060 ± 110	-
6l	16	40.0 ± 2.0	-	390 ± 10	-
6m[(b)]	16	67.0 ± 7.0	-	550 ± 70	10
6n[(b)]	16	50.0 ± 5.0	1.6:1	420 ± 60	64
6o[(b)]	16	51.0 ± 3.0	> 20:1	420 ± 40	-
6p[(b)]	8	14.0 ± 3.0	> 20:1	270 ± 60	-

Reaction conditions are as follows: $P411_{APA}$ in *E. coli* whole cells at indicated OD_{600}, 2.0 mM substrate **5**, 4.0 mM substrate **2**, 2.5 vol % EtOH in M9-N (pH 7.4) buffer. For substrates **5a–5l** and **5p**, the reaction volume was 800 µL in 2-mL vials with screw cap. For substrates **5m–5o**, the reaction volume was 1600 µL in 2-mL vials with screw caps due to the volatility of substrates. The reactions proceeded at 10 °C under anaerobic conditions

for 12 h. Product formation was quantified by HPLC-MS based on the calibration curve of the corresponding racemic reference compound (**Section A.6.**), and *ee* was measured with normal phase HPLC. Results are the average of four reactions (performed from two independent cell cultures, each used for duplicate reactions).

(a) The absolute configuration of **6a** and **6j** was determined by comparison with chiral standards (see **Section A.5.** and **Section A.7.**).

(b) Products **6m–6p** were characterized and quantified by their benzoyl-protected amines.

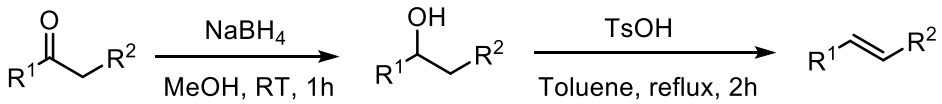
Table A-6. Summary of mutations in each variant relative to wild-type $P450_{BM3}$.[(a)]

Variant	Mutations relative to P411-B2
P411-B2	A82L, F87A, P142S, T175I, A184V, S226R, H236Q, E252G, I263M, T268G, A290V, A328V, L353V, I366V, C400S, T438S, E442K
$P411_{BPA}$	**V78M**, A82L, F87A, P142S, T175I, A184V, S226R, H236Q, E252G, I263M, **E267D**, **T268P**, A290V, **T327A**, A328V, L353V, I366V, C400S, **I437F**, **T438G**, E442K
$P411_{APA}$	**V78M**, A82L, F87A, P142S, T175I, A184V, S226R, H236Q, E252G, I263M, **E267D**, **T268P**, A290V, **T327A**, A328V, L353V, I366V, **N395R**, C400S, **I437F**, **T438A**, E442K

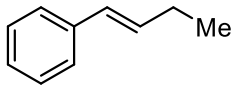
(a) See the sequences of $P411_{BPA}$ and $P411_{APA}$ in **Section A.8.**.

A.3. Synthesis and Characterization of Substrates

Benzylic substrates **1a–1l**, and allylic substrates **5j, 5k,** and **5m–5p** were purchased from commercial vendors (Millipore Sigma, VWR, TCI America, and Fischer Scientific). Other allylic substrates were synthesized according to the literature.[5]

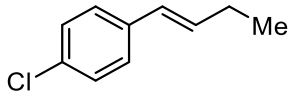


To the solution of a ketone (5.0 mmol, 1.0 eq) in 20 mL MeOH was added $NaBH_4$ solid (7.5 mmol, 1.5 eq) in batches over 10 minutes. The reaction mixture was stirred at room temperature for 1 hour. Then the reaction mixture was quenched with saturated NH_4Cl aqueous solution, extracted with ethyl acetate for three times and washed with saturated brine. The combined organic layer was dried over anhydrous Na_2SO_4 and concentrated by rotary evaporation. The crude product was used directly in the next step without further purification.The crude product and 4-methylbenzenesulfonic acid (0.5 mmol, 0.1 eq) were solved in toluene (20 mL). The mixture was heated to reflux for 2 hours. After cooling to room temperature, the mixture was extracted with ethyl acetate for three times and washed with saturated brine. The combined organic layer was dried over anhydrous Na_2SO_4 and concentrated by rotary evaporation. The crude product was purified by flash chromatography on silica gel to afford the corresponding alkene **5a**–**5i** and **5l** in 56–89% yield.



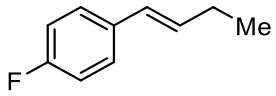
(*E*)-But-1-en-1-ylbenzene

5a, *E*:*Z* > 20:1; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.38 – 7.33 (m, 2H), 7.32 – 7.26 (m, 2H), 7.25 – 7.15 (m, 1H), 6.44 – 6.22 (m, 2H), 2.29 – 2.18 (m, 2H), 1.10 (t, *J* = 7.5 Hz, 3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 138.0, 132.7, 128.9, 128.6, 126.8, 126.0, 26.2, 13.8.



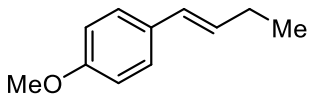
(*E*)-1-(But-1-en-1-yl)-4-chlorobenzene

5b, *E*:*Z* > 20:1; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.29 – 7.23 (m, 4H), 6.37 – 6.20 (m, 2H), 2.30 – 2.16 (m, 2H), 1.09 (t, *J* = 7.4 Hz, 3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 136.5, 133.5, 132.3, 128.7, 127.7, 127.2, 26.1, 13.6.



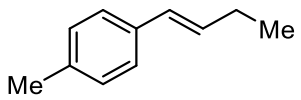
(*E*)-1-(But-1-en-1-yl)-4-fluorobenzene

5c, *E*:*Z* > 20:1; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.33 – 7.27 (m, 2H), 7.02 – 6.94 (m, 2H), 6.34 (dt, *J* = 15.8, 1.6 Hz, 1H), 6.24 – 6.11 (m, 1H), 2.27 – 2.18 (m, 2H), 1.09 (t, *J* = 7.5 Hz, 3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 161.9 (d, *J* = 245.3 Hz), 134.2 (d, *J* = 3.2 Hz), 132.5 (d, *J* = 2.2 Hz), 127.7, 127.4 (d, *J* = 7.9 Hz), 115.4 (d, *J* = 21.4 Hz), 26.1, 13.7.



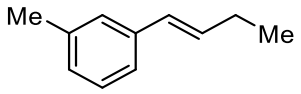
(*E*)-1-(But-1-en-1-yl)-4-methoxybenzene

5d, *E*:*Z* > 20:1; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.30 – 7.26 (m, 2H), 6.86 – 6.82 (m, 2H), 6.33 (dt, *J* = 15.7, 1.6 Hz, 1H), 6.13 (dt, *J* = 15.8, 6.5 Hz, 1H), 3.80 (s, 3H), 2.25 – 2.17 (m, 2H), 1.08 (t, *J* = 7.5 Hz, 3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 158.7, 130.9, 130.6, 128.2, 127.0, 114.0, 55.4, 26.1, 13.9.



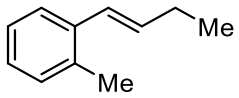
(*E*)-1-(But-1-en-1-yl)-4-methylbenzene

5e, *E*:*Z* = 8:1; for the major *E* isomer: ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.26 – 7.22 (m, 2H), 7.10 (d, *J* = 7.8 Hz, 2H), 6.35 (dd, *J* = 15.8, 1.5 Hz, 1H), 6.22 (ddd, *J* = 15.8, 6.9, 5.8 Hz, 1H), 2.33 (s, 3H), 2.27 – 2.17 (m, 2H), 1.09 (t, *J* = 7.5 Hz, 3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 136.5, 135.2, 131.7, 129.3, 128.7, 125.9, 26.2, 21.2, 13.8.



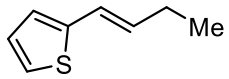
(*E*)-1-(But-1-en-1-yl)-3-methylbenzene

5f, *E*:*Z* > 20:1; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.22 – 7.13 (m, 3H), 7.04 – 6.98 (m, 1H), 6.36 (dt, *J* = 15.7, 1.3 Hz, 1H), 6.26 (dt, *J* = 15.8, 6.2 Hz, 1H), 2.34 (d, *J* = 0.8 Hz, 3H), 2.27 – 2.18 (m, 2H), 1.09 (t, *J* = 7.4 Hz, 3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 138.1, 138.0, 132.5, 128.9, 128.5, 127.6, 126.7, 123.1, 26.2, 21.5, 13.8.



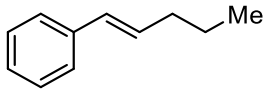
(*E*)-1-(But-1-en-1-yl)-2-methylbenzene

5g, *E*:*Z* > 20:1; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.43 – 7.40 (m, 1H), 7.17 – 7.11 (m, 4H), 6.57 (dt, *J* = 15.7, 1.6 Hz, 1H), 6.14 (dt, *J* = 15.7, 6.5 Hz, 1H), 2.34 (s, 3H), 2.30 – 2.22 (m, 2H), 1.11 (t, *J* = 7.5 Hz, 3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 137.1, 135.0, 134.1, 130.2, 126.8, 126.7, 126.1, 125.5, 26.5, 19.9, 13.9.



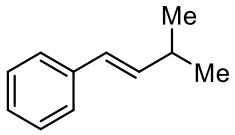
(*E*)-2-(But-1-en-1-yl)thiophene

5h, *E*:*Z* = 6:1; for the major *E* isomer: ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.11 – 7.05 (m, 1H), 6.93 (dd, *J* = 5.1, 3.5 Hz, 1H), 6.89 – 6.85 (m, 1H), 6.51 (ddt, *J* = 15.7, 1.7, 0.9 Hz, 1H), 6.12 (dt, *J* = 15.7, 6.5 Hz, 1H), 2.26 – 2.16 (m, 2H), 1.08 (t, *J* = 7.5 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 143.3, 132.7, 127.3, 124.2, 123.1, 122.2, 26.0, 13.6.



(*E*)-Pent-1-en-1-ylbenzene

5i, *E*:*Z* > 20:1; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.37 – 7.32 (m, 2H), 7.32 – 7.27 (m, 2H), 7.22 – 7.16 (m, 1H), 6.38 (dt, *J* = 15.8, 1.4 Hz, 1H), 6.23 (dt, *J* = 15.8, 6.8 Hz, 1H), 2.25 – 2.14 (m, 2H), 1.56 – 1.46 (m, 2H), 0.96 (t, *J* = 7.4 Hz, 3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 138.0, 131.1, 129.9, 128.6, 126.8, 126.0, 35.2, 22.7, 13.9.



(*E*)-(3-Methylbut-1-en-1-yl)benzene

5l, *E*:*Z* > 20:1; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.38 – 7.34 (m, 2H), 7.33 – 7.27 (m, 2H), 7.22 – 7.17 (m, 1H), 6.35 (dd, *J* = 15.9, 1.2 Hz, 1H), 6.21 (dd, *J* = 15.9, 6.7 Hz, 1H), 2.54 – 2.42 (m, 1H), 1.10 (d, *J* = 6.8 Hz, 6H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 138.1, 138.0, 128.6, 126.9, 126.8, 126.0, 31.6, 22.6.

A.4. 1.0 mmol-Scale Biosynthesis of Selected Amines

Protein expression:

BL21 E. cloni cells transformed with pET22b(+) plasmid encoding BPA or APA variant were grown in 50 mL LB_{amp} (37 °C, 230 rpm) in 125 mL Erlenmeyer flask for 10–12 hours. HB_{amp} medium (1,000 mL) with in a 2,800 mL flask was inoculated with 20 mL of the overnight culture. The HB culture was shaken at 37 °C and 120 rpm for about 2.5 hours in an incubator with a 25 mm shaking radius, until the OD_{600} was between 0.8 and 0.9. The culture was cooled on an ice bath for 1 hour and then treated with 0.5 mM IPTG and 1 mM 5-aminolevulinic acid (final concentrations). The culture was allowed to shake for 22 hours in the same incubator at 90 rpm and with the temperature reduced to 22 °C.

Reaction setup:

While harvesting cells, M9-N buffer (pH 7.4), D-glucose solution (0.5 M), substrates (0.1 M in ethanol or water) were precooled on an ice bath. Then, cells were pelleted by centrifugation (4,500 g, 5 minutes, 4 °C), resuspended in about 300 mL precooled M9-N buffer, and adjusted to OD_{600} = 32. Depending on substrates, a specific volume of cell suspension, precooled M9-N buffer, and precooled D-glucose solution (0.5 M, 37.5 mL) was then poured into a 500 mL or 1000 mL Erlenmeyer flask with screw cap on an ice bath to make a final volume of 457.5 mL (see below for more details). Subsequently, the flask with cells, substrate **1** or **5** (100 mM in ethanol, 10 mL, 1.0 mmol), substrate **2** (100 mM

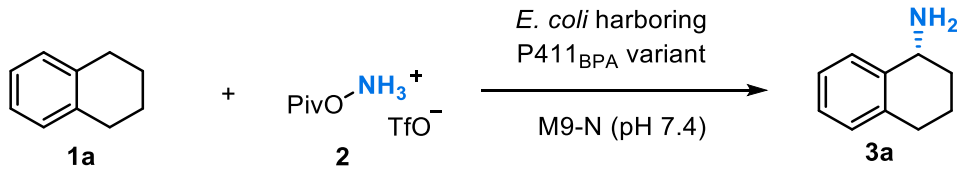
in water, 20 mL, 2.0 mmol), and a precooled stock solution for oxygen depletion (12.5 mL, 14,000 U/mL catalase and 1,000 U/mL glucose oxidase in M9-N buffer) were quickly transferred into an anaerobic chamber without the ice bath. In the anaerobic chamber, to the flask with cells was added the solution for oxygen depletion and the mixture was gently shaken by hands around 5 minutes. Then, substrate **1** or **5** and substrate **2** were added sequentially. The final reaction volume was 500 mL, and the final OD_{600} ranges from 4 to 16. The flask was sealed tightly, and the reaction proceeded outside the anaerobic chamber at 10 ℃ for 12 hours with shaking at 250 rpm with a 25 mm shaking radius. The leftover cell suspension was lysed to determine the hemoprotein concentration using hemochrome assay (**Section A.1.**).

Reaction work-up:

For products **3a**, **6a**, **6j**, **6k**, and **6l**: To the resulting mixture was added 250 mL ethanol and mixed vigorously by shaking. The cell debris was removed by centrifugation (5500 g, 10 minutes, 4 °C). Then ethanol and the majority of water in the supernatant were removed by rotary evaporation until a final volume of about 100 mL was achieved, and the remaining aqueous layer was treated with HCl solution (1M) to pH < 2 in an ice bath. After washing two times with diethyl ether (20 mL each time), the aqueous phase was treated with NaOH solution (1M) to PH > 12 on an ice bath and extracted with ethyl acetate for six times (30 mL each time). During the extraction, centrifugation was used for phase separation (5,000 g, 5 minutes, 4 °C). The combined organic layer was washed with brine

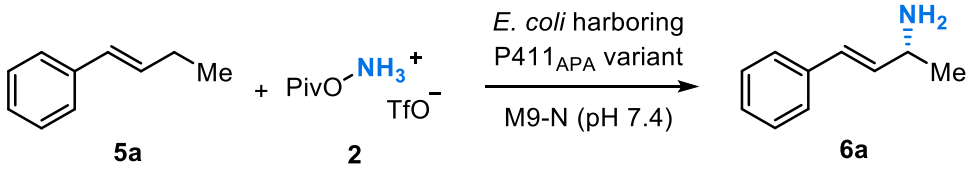
(30 mL), dried by Na_2SO_4, and concentrated *in vacuo* to give the final products without further purification.

For products **6o** and **6p**: To the resulting reaction mixture was directly added 50 mL NaHCO3 (sat) and 5 mL benzoyl chloride. The mixture was shaken for overnight at room temperature at 250 rpm with a 25 mm shaking radius. Then 250 mL EtOH was added, and the cell debris was removed by centrifugation (5500 g, 10 minutes, 4 °C). Ethanol and the majority of water in the supernatant were removed by rotary evaporation until a final volume of about 100 mL, and the aqueous phase was extracted by ethyl acetate for 5 times. The solvent was removed *in vacuo* and the resulting crude products were purified by silica gel with hexane/ethyl acetate to give Bz-**6o** and Bz-**6p**.



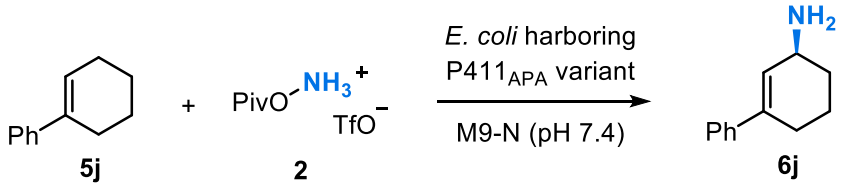
(*R*)-1,2,3,4-tetrahydronaphthalen-1-amine

3a: 89 mg, 61% yield, 520 TTN, 92% *ee*. The scale-up reaction was performed with 250 mL cells and 170 mL M9-N buffer, and the final OD_{600} was 16. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.41 (dd, *J* = 7.2, 1.9 Hz, 1H), 7.21 – 7.13 (m, 2H), 7.11 – 7.04 (m, 1H), 4.03 (t, *J* = 5.5 Hz, 1H), 2.91 – 2.68 (m, 2H), 2.62 (br s, 2H), 2.08 – 2.00 (m, 1H), 1.99 – 1.88 (m, 1H), 1.88 – 1.65 (m, 2H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 136.9, 129.2, 128.2, 126.9, 126.2, 49.5, 33.0, 29.6, 19.6.



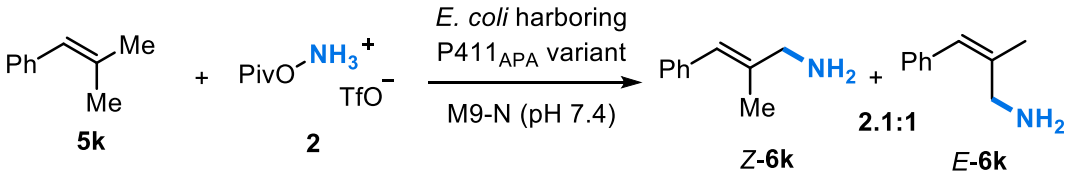
(*R*, *E*)-4-Phenylbut-3-en-2-amine

6a: 89 mg, 75% yield, 2300 TTN, 94% *ee*. The scale-up reaction was performed with 62.5 mL cells and 357.5 mL M9-N buffer, and the final OD_{600} was 4. The absolute configuration was determined by comparing with the chiral standard synthesized by chemical method (see chiral HPLC traces in **Section A.7.**)[6] The ^{1}H NMR and ^{13}C NMR of **6a** are identical to synthesized racemic *rac*-**6a** (see **Section A.9.**). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.39 – 7.33 (m, 2H), 7.33 – 7.27 (m, 2H), 7.25 – 7.19 (m, 1H), 6.48 (dd, J = 15.9, 1.0 Hz, 1H), 6.21 (dd, J = 15.9, 6.8 Hz, 1H), 3.76 – 3.65 (m, 1H), 2.07 (br s, 3H), 1.27 (d, J = 6.5 Hz, 3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 137.1, 135.3, 128.6, 128.6, 127.5, 126.4, 49.6, 23.7; $[\alpha]_D^{RT}$ = +27.8 (MeOH, c = 1.0).



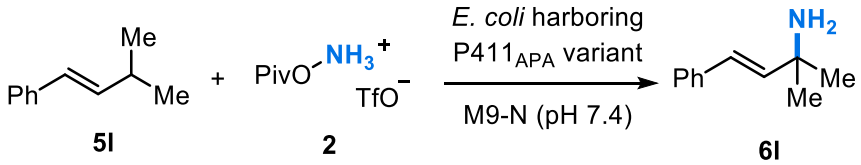
(*S*)-3,4,5,6-Tetrahydro-[1,1'-biphenyl]-3-amine

6j, 28 mg, 16% yield, 490 TTN, 87% *ee*. The scale-up reaction was performed with 125. mL cells and 295 mL M9-N buffer, and the final OD_{600} was 8. After acid-base extraction, the resulting product was further purified by flash chromatography on silica gel with ethyl acetate and methanol (0.1 % triethylamine) as eluents. The absolute configuration, determined by comparing the optical rotation of its hydrochloride salt with the literature, $[\alpha]_D^{RT}$ = +2.5 (hydrochloride salt of **6j** in MeOH, c = 0.5);[7] *r.r.* > 95:5, determined by ^{1}H NMR of the crude product. The ^{1}H NMR and ^{13}C NMR of **6a** are identical to synthesized racemic *rac*-**6a** (see **Section A.9.**). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.38 – 7.29 (m, 2H), 7.29 – 7.20 (m, 2H), 7.20 – 7.16 (m, 1H), 6.01 – 5.93 (m, 1H), 3.53 – 3.46 (m, 1H), 2.59 (br s, 2H), 2.44 – 2.22 (m, 2H), 2.00 – 1.79 (m, 2H), 1.70 – 1.57 (m, 1H), 1.44 – 1.30 (m, 1H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 141.7, 138.7, 128.4, 128.2, 127.3, 125.4, 47.9, 32.4, 27.4, 20.6.



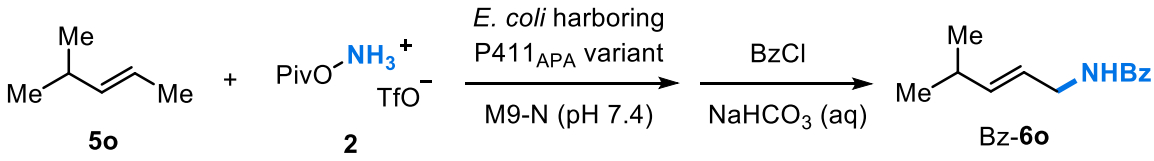
(*E*) and (Z)-2-Methyl-3-phenylprop-2-en-1-amine

6k, 47 mg, 32% yield, 980 TTN, 2.1:1 *r.r.* determined by ^{1}H NMR of the crude product (see **Section A.9.**). The scale-up reaction was performed with 125 mL cells and 295 mL M9-N buffer, and the final OD_{600} was 8. For *E* isomer: ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.37 – 7.17 (m, 5H), 6.44 (s, 1H), 3.38 (s, 2H), 1.89 (s, 3H), 1.87 (br s, 2H); ^{13}C NMR (101 MHz, $CDCl_3$) δ ^{13}C NMR (101 MHz, $CDCl_3$) δ 139.7, 138.1, 128.9, 128.2, 126.2, 123.9, 50.5, 16.5; HRMS (ESI) m/z: 131.0853 (M – NH_2); calc. for $C_{10}H_{11}$: 131.0861.



(*E*)-2-Methyl-4-phenylbut-3-en-2-amine

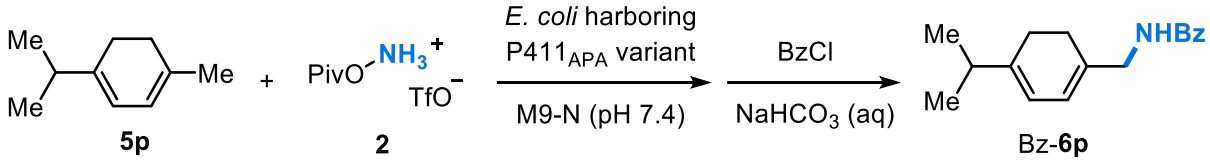
6l, 45 mg, 28% yield, 300 TTN. The scale-up reaction was performed with 250 mL cells and 170 mL M9-N buffer, and the final OD_{600} was 16. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.39 – 7.35 (m, 2H), 7.32 – 7.27 (m, 2H), 7.24 – 7.19 (m, 1H), 6.50 (d, J = 16.1 Hz, 1H), 6.31 (d, J = 16.2 Hz, 1H), 3.01 (br s, 2H), 1.37 (s, 6H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 137.5, 136.9, 128.7, 127.6, 126.7, 126.5, 52.1, 29.6; HRMS (ESI) m/z: 145.1005 (M – NH_2); calc. for $C_{11}H_{13}$: 145.1017.



(*E*)-*N*-(4-methylpent-2-en-1-yl)benzamide

Bz-**6o**, 83 mg, 41 % yield, 870 TTN, *r.r.* > 20:1, determined by ^{1}H NMR of the crude product. The scale-up reaction was performed in 500 mL Erlenmeyer flask with 250

mL cells and 170 mL M9-N buffer, and the final OD_{600} was 16. ^1H NMR (400 MHz, $CDCl_3$) δ 7.82 – 7.73 (m, 2H), 7.53 – 7.47 (m, 1H), 7.46 – 7.39 (m, 2H), 6.10 (s, 1H), 5.72 – 5.63 (m, 1H), 5.55 – 5.45 (m, 1H), 4.05 – 3.99 (m, 2H), 2.37 – 2.24 (m, 1H), 1.01 (s, 3H), 0.99 (s, 3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 167.3, 141.3, 134.8, 131.5, 128.6, 127.0, 122.7, 42.1, 30.9, 22.3; HRMS (ESI) m/z: 204.1380 ($M + H^+$); calc. for $C_{13}H_{18}NO$: 204.1383.



_N_-((4-isopropylcyclohexa-1,3-dien-1-yl)methyl)benzamide

Bz-**6p**, 25 mg, 10 % yield, 230 TTN, *r.r.* > 20:1, determined by ^1H NMR of the crude product. The scale-up reaction was performed with 125 mL cells and 295 mL M9-N buffer, and the final OD_{600} was 8. ^1H NMR (400 MHz, $CDCl_3$) δ 7.81 – 7.75 (m, 2H), 7.53 – 7.47 (m, 1H), 7.46 – 7.40 (m, 2H), 6.17 (br s, 1H), 5.84 (dt, J = 5.5, 1.4 Hz, 1H), 5.64 (dd, J = 5.5, 1.2 Hz, 1H), 4.10 (dd, J = 5.9, 1.4 Hz, 2H), 2.38 – 2.25 (m, 1H), 2.24 – 2.09 (m, 4H), 1.03 (d, J = 6.8 Hz, 6H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 167.5, 145.6, 134.7, 132.6, 131.5, 128.7, 127.0, 127.0, 127.0, 121.2, 115.8, 45.0, 34.7, 25.5, 25.3, 21.2; HRMS (ESI) m/z: 256.1690 ($M + H^+$); calc. for $C_{17}H_{22}NO$: 256.1696.

A.5. Synthesis and Characterization of Reference Products

Achiral and racemic amines *rac*-**3a**–*rac*-**3l** and optical pure amines *S*-**3a** and *S*-**3i** were purchased from commercial vendors. Racemic amines *rac*-**6a**–*rac*-**6j**, benzoyl-protected amines Bz-*rac*-**6m** and Bz-*rac*-**6n**, and *S*-enantiomer of **6a** were obtained through chemical synthesis. Enantioenriched amines (**3a**, **6a**, **6j)**, achiral amines (**6k**, **6l)**, and benzoyl-protected amines (Bz-**6o**, Bz-**6p)** were obtained by scale-up enzymatic reactions as described in **Section A.4.**.

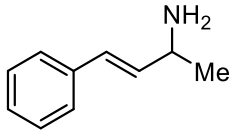
Procedure:[8]

A mixture of ketone or aldehyde (2 mmol, 1.0 eq), titanium(IV) isopropoxide (4 mmol, 2.0 eq), and ammonia in ethyl alcohol (2 M, 5 mL, 10 mmol, 5.0 eq) was stirred under argon in a capped flask at ambient temperature for 8 hours. $NaBH_4$ solid (3 mmol, 1.5 eq) was then added and the resulting mixture was stirred at room temperature for 12 hours. A second batch of $NaBH_4$ (3 mmol, 1.5 eq) was added, and the reaction was stirred for an additional 12 hours.

The reaction was then quenched by pouring into ammonium hydroxide (2 M, 25 mL), the resulting inorganic precipitate was filtered off, and washed with ethyl acetate for two times. The organic layer was separated, and the remaining aqueous layer was extracted with ethyl acetate for two times. To the aqueous phase was added HCl solution (1M) until pH < 2. After washing three times with ethyl acetate, the aqueous phase was treated with NaOH

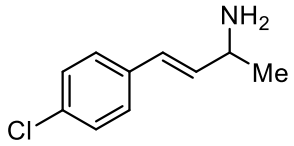
solution (1M) to PH > 12 and extracted three times with ethyl acetate. The combined organic layer was washed with brine, dried by Na_2SO_4, and concentrated *in vacuo* to afford racemic amines *rac*-**6a**–*rac*-**6j** with 15–75% yields.

For Bz-*rac*-**6m** and Bz-*rac*-**6n** (including major isomer and minor isomer), protection of amines was performed. To a solution of the crude products in 2 mL DCM was added benzoyl chloride (2 mmol, 1.0 eq) in an ice bath. Then the resulting mixtures were allowed to stir at room temperature for 2 hours and then were directly subjected to flash chromatography on silica gel.



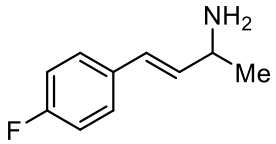
(*E*)-4-Phenylbut-3-en-2-amine

rac-**6a**, ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.39 – 7.34 (m, 2H), 7.34 – 7.27 (m, 2H), 7.24 – 7.18 (m, 1H), 6.47 (dt, *J* = 15.9, 0.7 Hz, 1H), 6.21 (dd, *J* = 15.9, 6.7 Hz, 1H), 3.72 – 3.63 (m, 1H), 1.48 (br s, 2H), 1.26 (d, *J* = 6.5 Hz, 3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 137.3, 136.3, 128.6, 127.9, 127.3, 126.3, 49.4, 24.0; HRMS (ESI) m/z: 131.0862 ($M - NH_2$); calc. for $C_{10}H_{11}$: 131.0861.



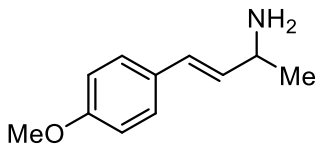
(*E*)-4-(*p*-Chlorophenyl)but-3-en-2-amine

rac-**6b**, ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.30 – 7.24 (m, 5H), 6.43 (dd, *J* = 15.9, 1.1 Hz, 1H), 6.18 (dd, *J* = 15.9, 6.6 Hz, 1H), 3.73 – 3.63 (m, 1H), 1.73 (br s, 2H), 1.26 (d, *J* = 6.5 Hz, 3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 136.6, 135.7, 133.0, 128.8, 127.6, 127.0, 49.4, 23.8; HRMS (ESI) m/z: 165.0471 ($M - NH_2$); calc. for $C_{10}H_{10}Cl$: 165.0471.



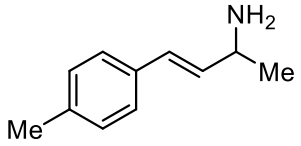
(*E*)-4-(*p*-Fluorophenyl)but-3-en-2-amine

rac-**6c**, ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.35 – 7.28 (m, 2H), 7.02 – 6.95 (m, 2H), 6.46 – 6.39 (m, 1H), 6.12 (ddd, *J* = 15.9, 6.7, 0.5 Hz, 1H), 3.71 – 3.62 (m, 1H), 1.70 (br s, 2H), 1.26 (d, *J* = 6.6 Hz, 3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 162.2 (d, *J* = 246.2 Hz), 135.7 (d, *J* = 2.2 Hz), 133.3 (d, *J* = 3.3 Hz), 127.8 (d, *J* = 8.0 Hz), 127.0, 115.5 (d, *J* = 21.6 Hz), 49.4, 23.9; HRMS (ESI) m/z: 149.0766 ($M - NH_2$); calc. for $C_{10}H_{10}F$: 149.0767.



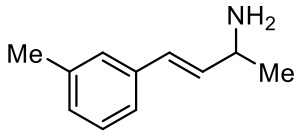
(*E*)-4-(4-Methoxyphenyl)but-3-en-2-amine

rac-**6d**, ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.34 – 7.27 (m, 2H), 6.88 – 6.82 (m, 2H), 6.46 – 6.37 (m, 1H), 6.07 (dd, *J* = 15.9, 6.8 Hz, 1H), 3.80 (s, 3H), 3.71 – 3.61 (m, 1H), 1.79 (br s, 2H), 1.25 (d, *J* = 6.5 Hz, 3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 159.1, 133.8, 130.0, 127.6, 127.5, 114.0, 55.4, 49.5, 24.0; HRMS (ESI) m/z: 161.0958 (M – NH_2); calc. for $C_{11}H_{13}O$: 161.0966.



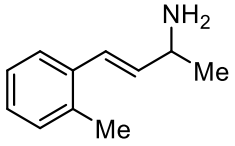
(*E*)-4-(*p*-Tolyl)but-3-en-2-amine

rac-**6e**, ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.30 – 7.25 (m, 2H), 7.15 – 7.08 (m, 2H), 6.47 – 6.38 (m, 1H), 6.15 (dd, *J* = 15.8, 6.7 Hz, 1H), 3.71 – 3.61 (m, 1H), 2.33 (s, 3H), 1.43 (br s, 2H), 1.24 (d, *J* = 6.5 Hz, 3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 137.1, 135.4, 134.5, 129.3, 127.7, 126.2, 49.5, 24.1, 21.3; HRMS (ESI) m/z: 145.1015 (M – NH_2); calc. for $C_{10}H_{13}$: 145.1017.



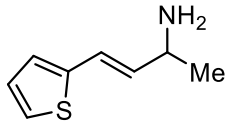
(*E*)-4-(*m*-Tolyl)but-3-en-2-amine

rac-**6f**, ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.22 – 7.14 (m, 3H), 7.06 – 7.01 (m, 1H), 6.43 (dd, *J* = 15.9, 1.1 Hz, 1H), 6.20 (dd, *J* = 15.9, 6.7 Hz, 1H), 3.70 – 3.61 (m, 1H), 2.34 (s, 3H), 1.45 (br s, 2H), 1.25 (d, *J* = 6.5 Hz, 3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 138.2, 137.2, 136.2, 128.5, 128.1, 127.9, 127.0, 123.5, 49.5, 24.1, 21.5; HRMS (ESI) m/z: 145.1010 (M – NH_2); calc. for $C_{10}H_{13}$: 145.1017.



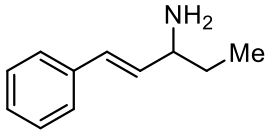
(*E*)-4-(*o*-Tolyl)but-3-en-2-amine

rac-**6g**, ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.47 – 7.38 (m, 1H), 7.20 – 7.07 (m, 4H), 6.66 (dd, *J* = 15.8, 1.2 Hz, 1H), 6.09 (dd, *J* = 15.7, 6.7 Hz, 1H), 3.74 – 3.64 (m, 1H), 2.34 (s, 3H), 1.45 (br s, 2H), 1.26 (d, *J* = 6.5 Hz, 3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 137.8, 136.3, 135.4, 130.3, 127.3, 126.2, 125.6, 125.6, 49.7, 24.2, 19.9; HRMS (ESI) m/z: 145.1036 (M – NH_2); calc. for $C_{10}H_{13}$: 145.1017.



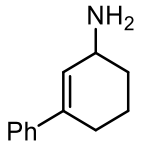
(*E*)-4-(Thiophen-2-yl)but-3-en-2-amine

rac-**6h**, ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.12 (dt, *J* = 5.0, 0.9 Hz, 1H), 6.98 – 6.88 (m, 2H), 6.65 – 6.55 (m, 1H), 6.06 (dd, *J* = 15.7, 6.6 Hz, 1H), 3.68 – 3.57 (m, 1H), 1.40 (br s, 2H), 1.24 (d, *J* = 6.5 Hz, 3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 142.5, 136.1, 127.4, 125.2, 123.8, 121.2, 49.2, 23.9; HRMS (ESI) m/z: 137.0427 (M – NH_2); calc. for C_8H_9S: 137.0425.



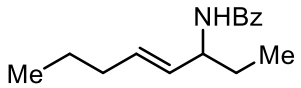
(*E*)-1-Phenylpent-1-en-3-amine

rac-**6i**, ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.42 – 7.35 (m, 2H), 7.35 – 7.28 (m, 2H), 7.25 – 7.19 (m, 1H), 6.49 (dd, *J* = 15.9, 1.0 Hz, 1H), 6.14 (dd, *J* = 15.9, 7.3 Hz, 1H), 3.45 – 3.37 (m, 1H), 1.88 (br s, 2H), 1.62 – 1.50 (m, 2H), 0.94 (t, *J* = 7.4 Hz, 3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 137.2, 134.4, 129.6, 128.6, 127.4, 126.4, 55.8, 30.6, 10.6; HRMS (ESI) m/z: 145.0997 (M – NH_2); calc. for $C_{10}H_{13}$: 145.1017.



3,4,5,6-Tetrahydro-[1,1'-biphenyl]-3-amine

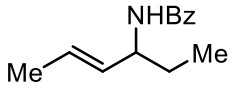
rac-**6j**, ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.44 – 7.36 (m, 2H), 7.36 – 7.27 (m, 2H), 7.29 – 7.19 (m, 1H), 6.06 – 6.00 (m, 1H), 3.56 – 3.45 (m, 1H), 2.47 – 2.32 (m, 2H), 2.05 – 1.86 (m, 2H), 1.77 – 1.65 (m, 1H), 1.46 – 1.35 (m, 3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 141.9, 137.7, 129.7, 128.3, 127.1, 125.3, 47.9, 33.3, 27.5, 20.8; HRMS (ESI) m/z: 157.1018 (M – NH_2); calc. for $C_{12}H_{13}$: 157.1017.



(*E*)-*N*-(Oct-4-en-3-yl)benzamide

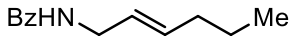
Bz-*rac*-**6m**, ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.81 – 7.72 (m, 2H), 7.53 – 7.38 (m, 3H), 5.97 (d, *J* = 8.3 Hz, 1H), 5.66 (dtd, *J* = 15.1, 6.8, 1.3 Hz, 1H), 5.42 (ddt, *J* = 15.4, 6.4, 1.5

Hz, 1H), 4.61 – 4.50 (m, 1H), 2.08 – 1.96 (m, 2H), 1.72 – 1.59 (m, 2H), 1.46 – 1.34 (m, 2H), 0.96 (t, *J* = 7.4 Hz, 3H), 0.89 (t, *J* = 7.4 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 166.8, 135.1, 132.2, 131.4, 129.9, 128.6, 126.9, 52.8, 34.5, 28.5, 22.4, 13.8, 10.2. HRMS (ESI) m/z: 232.1673 ($M + H^+$); calc. for $C_{15}H_{22}NO$: 232.1701.



(*E*)-*N*-(Hex-4-en-3-yl)benzamide

Bz-*rac*-**6n**-major, ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.81 – 7.72 (m, 2H), 7.53 – 7.39 (m, 3H), 5.95 (d, *J* = 8.1 Hz, 1H), 5.76 – 5.61 (m, 1H), 5.51 – 5.40 (m, 1H), 4.59 – 4.47 (m, 1H), 1.70 (ddd, *J* = 6.4, 1.7, 1.1 Hz, 3H), 1.68 – 1.61 (m, 2H), 0.96 (t, *J* = 7.4 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 166.8, 135.0, 131.4, 131.1, 128.6, 126.9, 52.8, 28.4, 17.9, 10.4. HRMS (ESI) m/z: 204.1414 ($M + H^+$); calc. for $C_{13}H_{18}NO$: 204.1388.



(*E*)-*N*-(Hex-4-en-3-yl)benzamide

Bz-**6n**-minor, ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.80 – 7.70 (m, 2H), 7.52 – 7.45 (m, 1H), 7.45 – 7.35 (m, 2H), 6.20 (s, 1H), 5.75 – 5.62 (m, 1H), 5.59 – 5.48 (m, 1H), 4.02 (td, *J* = 5.8, 1.2 Hz, 2H), 2.06 – 1.96 (m, 2H), 1.45 – 1.35 (m, 2H), 0.90 (t, *J* = 7.4 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 167.3, 134.7, 134.1, 131.5, 128.6, 127.0, 125.7, 42.1, 34.4, 22.3, 13.8. HRMS (ESI) m/z: 204.1402 ($M + H^+$); calc. for $C_{13}H_{18}NO$: 204.1388.

A.6. Analytic Scale Enzymatic Reactions and Calibration Curves for Products

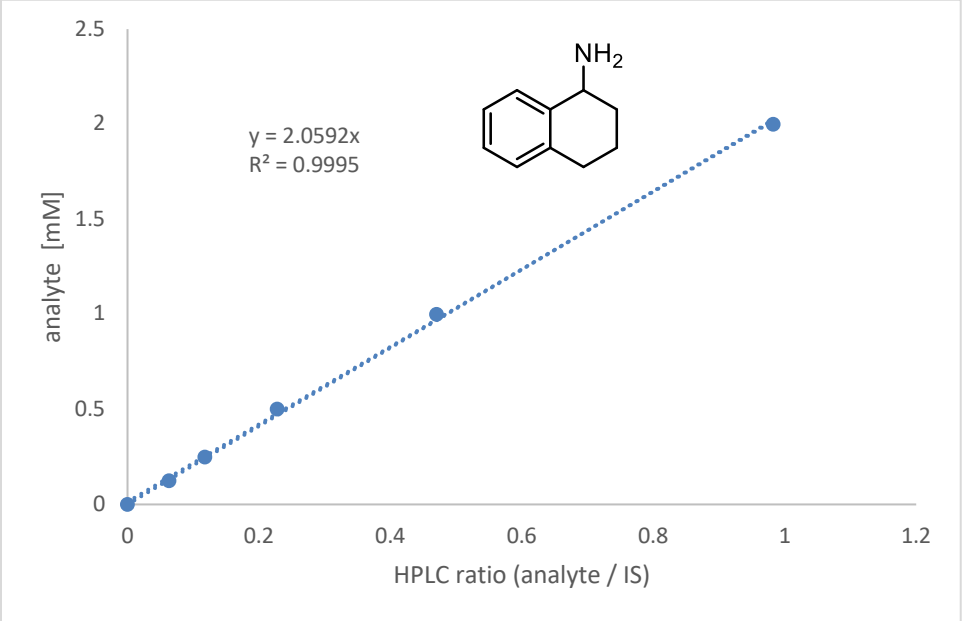
Enzymatic reactions on analytic scale were performed following the general procedure described in **Section A- (D)**. Product formation was quantified by HPLC based on the calibration curve of the corresponding racemic reference compound (**Section A.5.**). TTN is defined as the concentration of product divided by the concentration of heme protein measured by the hemochrome assay (**Section A.1. (C)**). Analysis data presented in this section are for results shown in **Fig. 2** and **Fig. 3** of the main text, as well as in **Table A-3** and **Table A-5**.

Calibration curves of commercially available or synthesized reference compounds were created for the determination of yield and TTN. To a 2-mL vial were added 1160 µL water:EtOH (2:1), and 20 µL product stock solution in EtOH, 20 µL internal standard (20 mM 1,3,5-trimethoxybenzene in EtOH for benzylic amines **3**, 300 mM ethyl benzoate in EtOH for allylic amines **6**, and 200 mM 2-amino-1,2,3,4-tetrahydronaphthalen-1-ol for byproduct **7a**). The mixture was vortexed and then analyzed by HPLC. The concentration of UV-active analytes was plotted based on the UV absorbance ratio of analytes (products **3** and **6**) or the ion count at selected ion monitoring mode of byproduct **7a** over the indicated internal standard. Cell impurities shared the same retention time. Methods were adjusted so that none of the products elute at the same time with any of the cell impurities. For most analyses, water and acetonitrile containing 0.1% acetic acid were used as eluents for a Kromasil® C18 column. For analysis of byproduct **7a**, a C8 column (Agilent Poroshell 120,

4.6 × 50 mm, 4 µm) was used. The methods used 10–95% acetonitrile (0–3 minutes) and 5-µL sample injections. The flow rate was 1.5 mL/minute, and the column was maintained at 37 °C.

Benzylic amines 3a–3l:

3a: based on UV absorbance at 210 nm and using 1,3,5-trimethoxybenzene as internal standard.

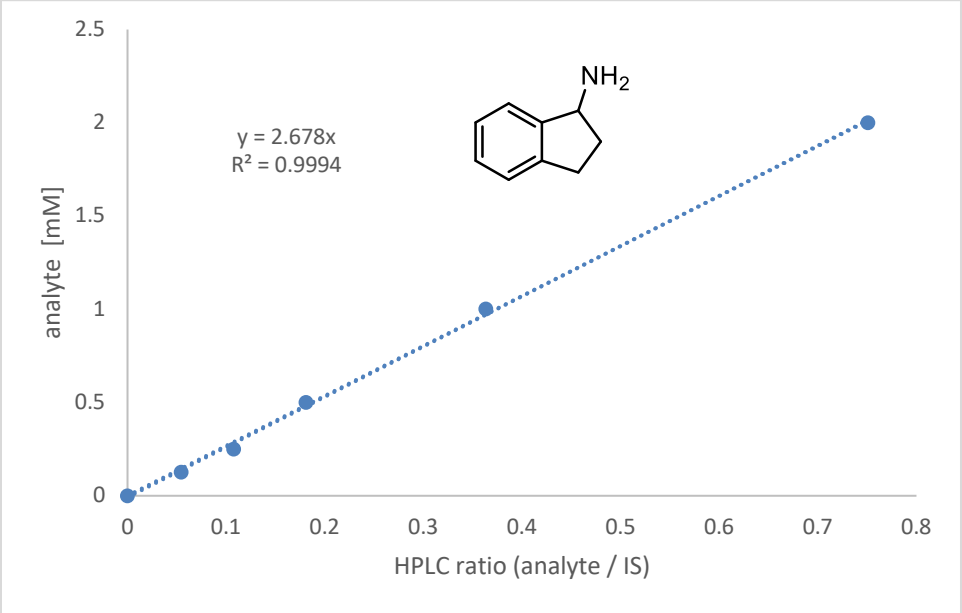


Analysis data for P411$_{BPA}$ catalyzed primary amination of alkane **1a** (2 mM, OD_{600} = 24):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/μM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	1707.2	1934.3	0.883	1.82	2.94	580	30	85	4
2	1637.4	2055.2	0.797	1.64	2.94				
3	1582.9	1950.5	0.812	1.67	2.9				
4	1578.3	1968.2	0.802	1.65	2.9				

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.

3b: based on UV absorbance at 210 nm and using 1,3,5-trimethoxybenzene as internal standard.

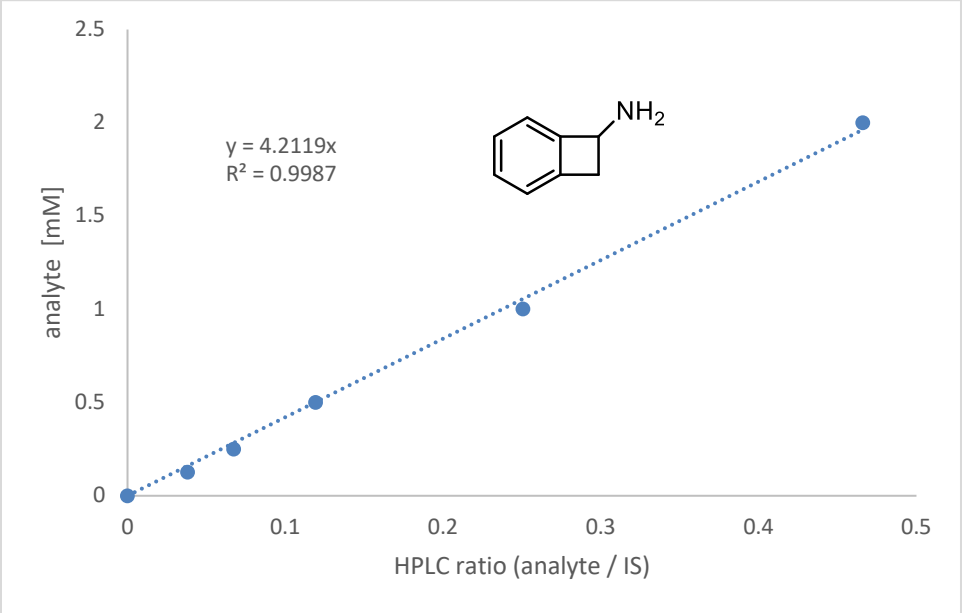


Analysis data for P411$_{BPA}$ catalyzed primary amination of alkane **1b** (2 mM, OD_{600} = 24):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/μM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	1311.8	2153.9	0.609	1.63	2.94	510	40	74	6
2	1116.3	2167.4	0.515	1.38	2.94				
3	1137.9	2173.1	0.524	1.40	2.9				
4	1245.8	2217.3	0.562	1.50	2.9				

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.

3c: based on UV absorbance at 210 nm and using 1,3,5-trimethoxybenzene as internal standard.

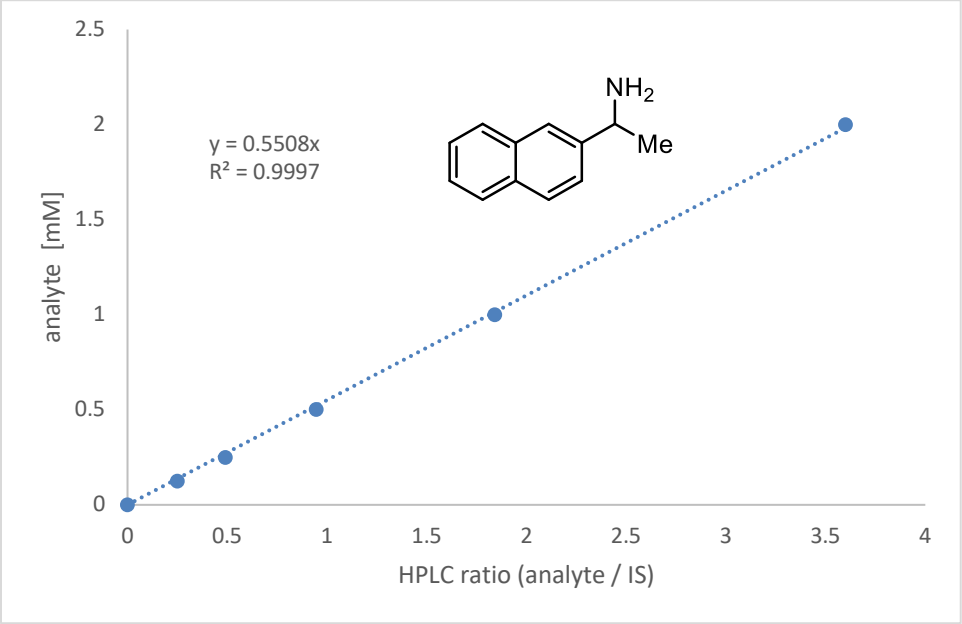


Analysis data for P411$_{BPA}$ catalyzed primary amination of alkane **1c** (2 mM, OD_{600} = 24):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/μM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	145.4	1942.6	0.075	0.32	2.94	120	10	18	2
2	156.3	2004.4	0.078	0.33	2.94				
3	179.8	1963.9	0.092	0.39	2.9				
4	188.6	2036.8	0.093	0.39	2.9				

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.

3d: based on UV absorbance at 210 nm and using 1,3,5-trimethoxybenzene as internal standard.

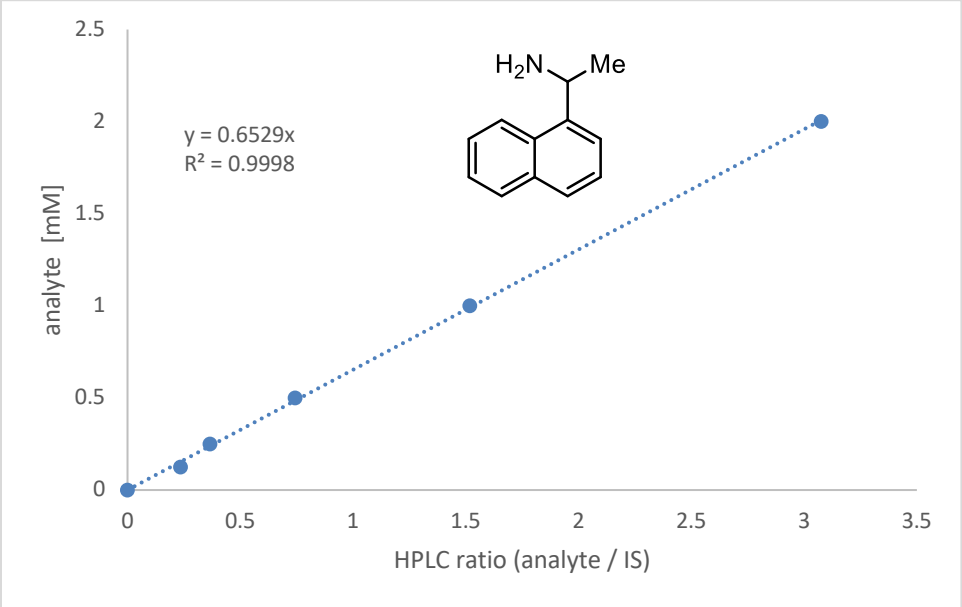


Analysis data for P411$_{BPA}$ catalyzed primary amination of alkane **1d** (2 mM, OD_{600} = 24):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/μM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	3026	2029	1.491	0.82	2.94	210	80	38	2
2	2856.4	2128.8	1.342	0.74	2.94				
3	2880.4	2074.4	1.389	0.76	2.9				
4	2828.7	2092.9	1.352	0.74	2.9				

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.

3e: based on UV absorbance at 210 nm and using 1,3,5-trimethoxybenzene as internal standard.

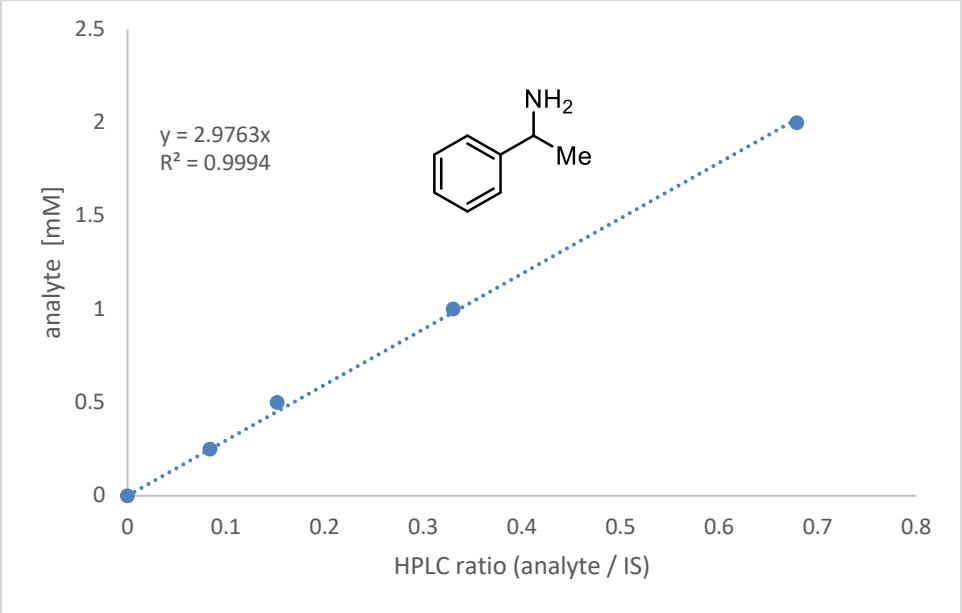


Analysis data for P411$_{BPA}$ catalyzed primary amination of alkane **1e** (2 mM, OD_{600} = 24):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/μM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	5132.5	2009.5	2.554	1.67	2.94	520	40	75	6
2	4968.2	2146.6	2.314	1.51	2.94				
3	4572	2068.9	2.210	1.44	2.9				
4	4669.7	2168.5	2.153	1.41	2.9				

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.

3f: based on UV absorbance at 210 nm and using 1,3,5-trimethoxybenzene as internal standard.

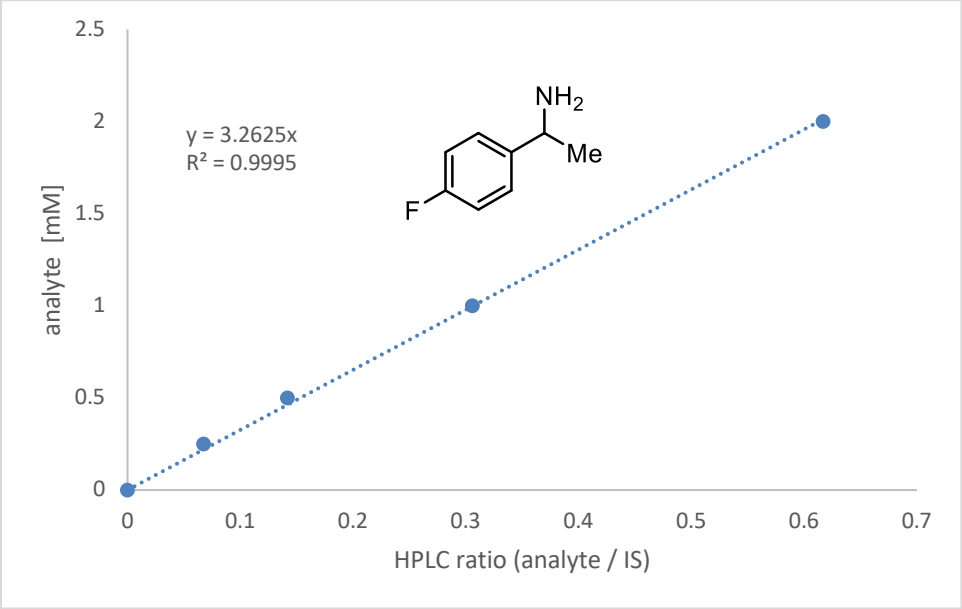


Analysis data for P411$_{BPA}$ catalyzed primary amination of alkane **1f** (2 mM, OD_{600} = 24):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/μM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	437.6	1963.1	0.223	0.66	2.94	210	10	30	2
2	404.8	2057.6	0.197	0.59	2.94				
3	396.7	2054.7	0.193	0.57	2.9				
4	422.4	2082.5	0.203	0.60	2.9				

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.

3g: based on UV absorbance at 210 nm and using 1,3,5-trimethoxybenzene as internal standard.

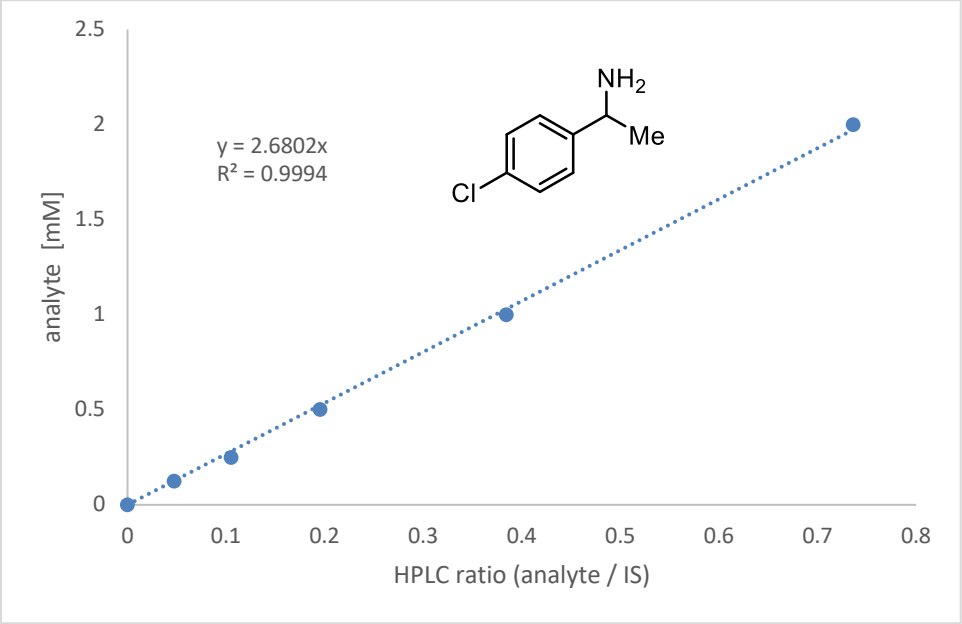


Analysis data for P411$_{BPA}$ catalyzed primary amination of alkane **1g** (2 mM, OD_{600} = 24):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/µM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	552.7	1996.8	0.277	0.90	2.94	300	30	44	4
2	489.6	2030.4	0.241	0.79	2.94				
3	519.1	2008.6	0.258	0.84	2.9				
4	586.1	2017.4	0.291	0.95	2.9				

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.

3h: based on UV absorbance at 210 nm and using 1,3,5-trimethoxybenzene as internal standard.

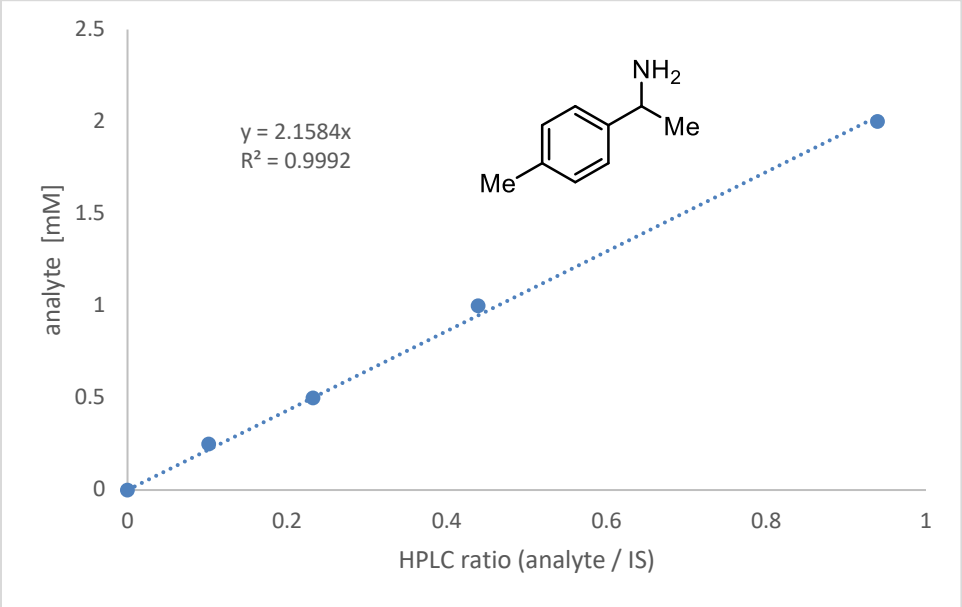


Analysis data for P411$_{BPA}$ catalyzed primary amination of alkane **1h** (2 mM, OD_{600} = 24):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/µM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	760	1928.8	0.394	1.06	2.94	310	30	46	5
2	627.4	1951.7	0.321	0.86	2.94				
3	652.7	1931.8	0.338	0.91	2.9				
4	596.4	1944.3	0.307	0.82	2.9				

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.

3i-major: based on UV absorbance at 210 nm and using 1,3,5-trimethoxybenzene as internal standard.

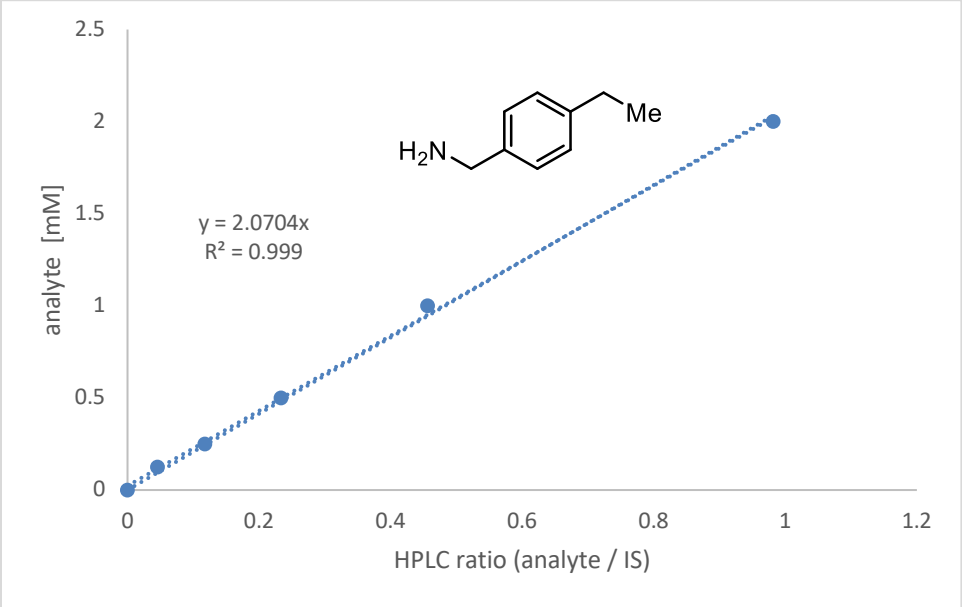


Analysis data for P411$_{BPA}$ catalyzed primary amination of alkane **1i** (2 mM, OD_{600} = 16):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/µM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	695.1	1943.1	0.358	0.77	1.98	380	10	37	3
2	644.3	1954.1	0.330	0.71	1.98				
3	648.9	1924.3	0.337	0.73	1.9				
4	646.8	1933.5	0.335	0.72	1.9				

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.

3i-minor: based on UV absorbance at 210 nm and using 1,3,5-trimethoxybenzene as internal standard.

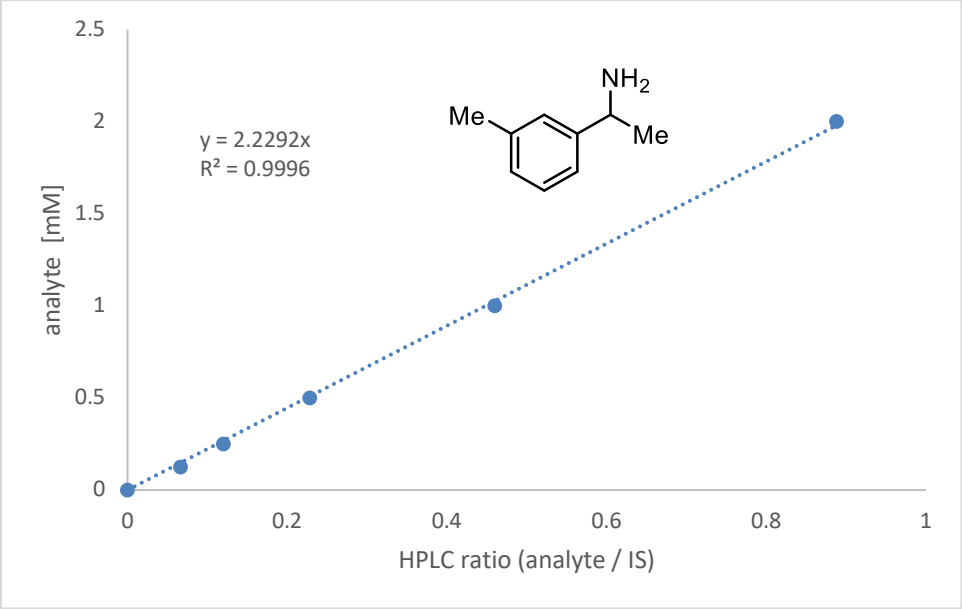


Analysis data for P411$_{BPA}$ catalyzed primary amination of alkane **1i** (2 mM, OD_{600} = 16):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/μM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	233.7	1943.1	0.120	0.25	1.98	120	10	12	1
2	238.8	1954.1	0.122	0.25	1.98				
3	184.9	1924.3	0.096	0.20	1.9				
4	219.1	1933.5	0.113	0.23	1.9				

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.

3j-major: based on UV absorbance at 210 nm and using 1,3,5-trimethoxybenzene as internal standard.

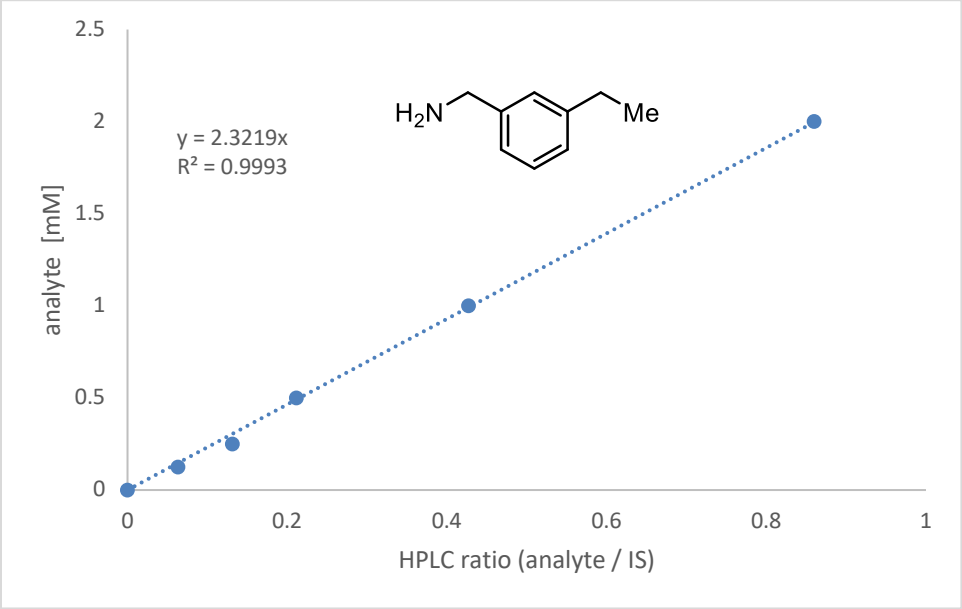


Analysis data for P411$_{BPA}$ catalyzed primary amination of alkane **1j** (2 mM, OD_{600} = 16):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/μM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	1071	1885.8	0.568	1.27	1.98	600	60	58	7
2	1060.4	1831.3	0.579	1.29	1.98				
3	952	1926.3	0.494	1.10	1.9				
4	832.2	1859.4	0.448	1.00	1.9				

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.

3j-minor: based on UV absorbance at 210 nm and using 1,3,5-trimethoxybenzene as internal standard.

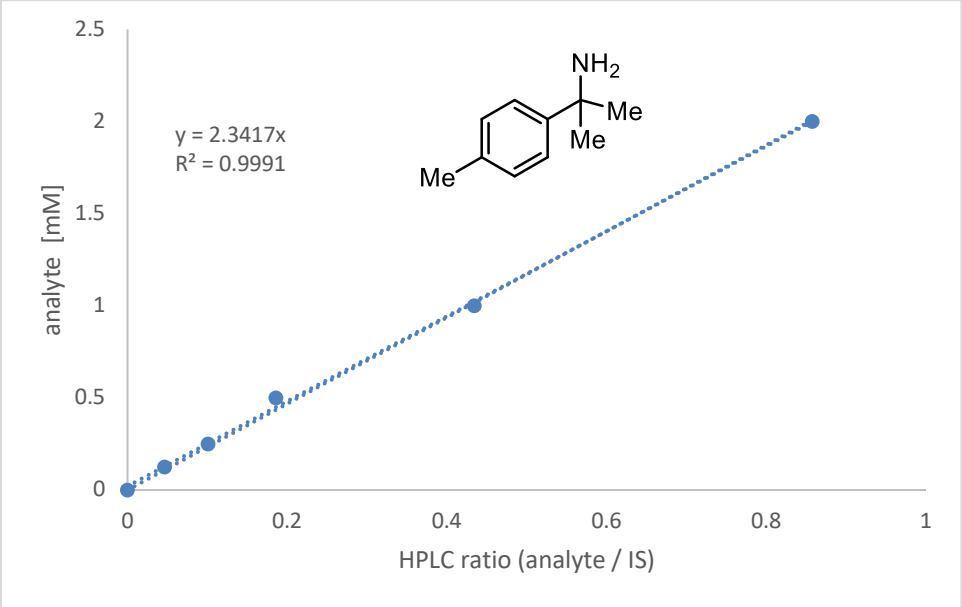


Analysis data for P411$_{BPA}$ catalyzed primary amination of alkane **1j** (2 mM, OD_{600} = 16):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/µM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	262.9	1885.8	0.139	0.32	1.98	150	10	15	2
2	259.6	1831.3	0.142	0.33	1.98				
3	214.9	1926.3	0.112	0.26	1.9				
4	227.3	1859.4	0.122	0.28	1.9				

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.

3k-major: based on UV absorbance at 210 nm and using 1,3,5-trimethoxybenzene as internal standard.

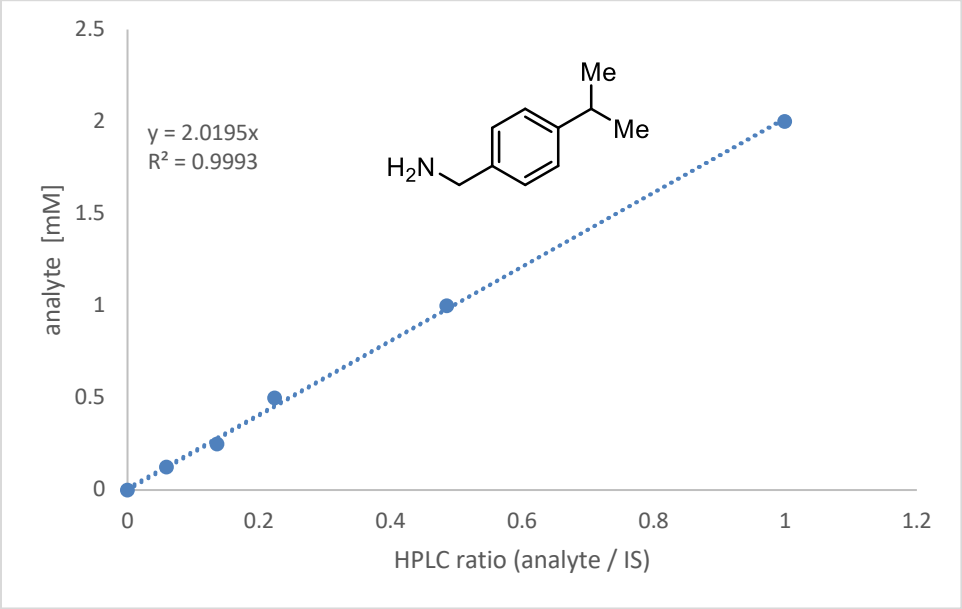


Analysis data for P411$_{BPA}$ catalyzed primary amination of alkane **1k** (2 mM, OD_{600} = 24):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/μM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	456.5	2042.1	0.224	0.52	2.94	170	10	25	2
2	427.9	2114.7	0.202	0.47	2.94				
3	460.1	2048.6	0.225	0.53	2.9				
4	432.7	2135.9	0.203	0.47	2.9				

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.

3k-minor: based on UV absorbance at 210 nm and using 1,3,5-trimethoxybenzene as internal standard.

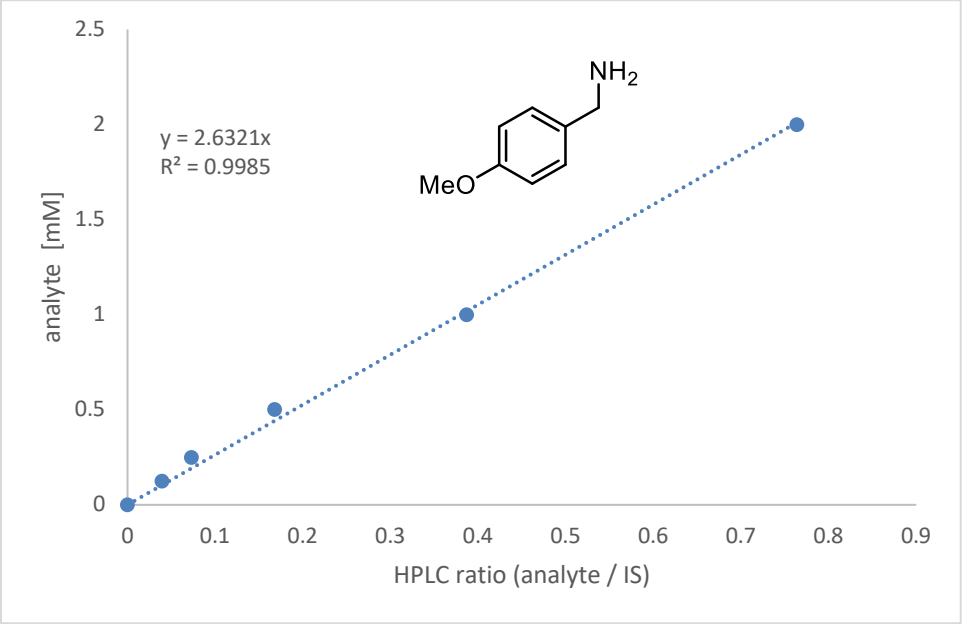


Analysis data for P411$_{BPA}$ catalyzed primary amination of alkane **1k** (2 mM, OD_{600} = 24):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/μM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	210.6	1984.8	0.106	0.21	2.94	70	5	10	1
2	215.6	2057.8	0.105	0.21	2.94				
3	207.6	1972.2	0.105	0.21	2.9				
4	188.3	2076	0.091	0.18	2.9				

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.

3l: based on UV absorbance at 210 nm and using 1,3,5-trimethoxybenzene as internal standard.



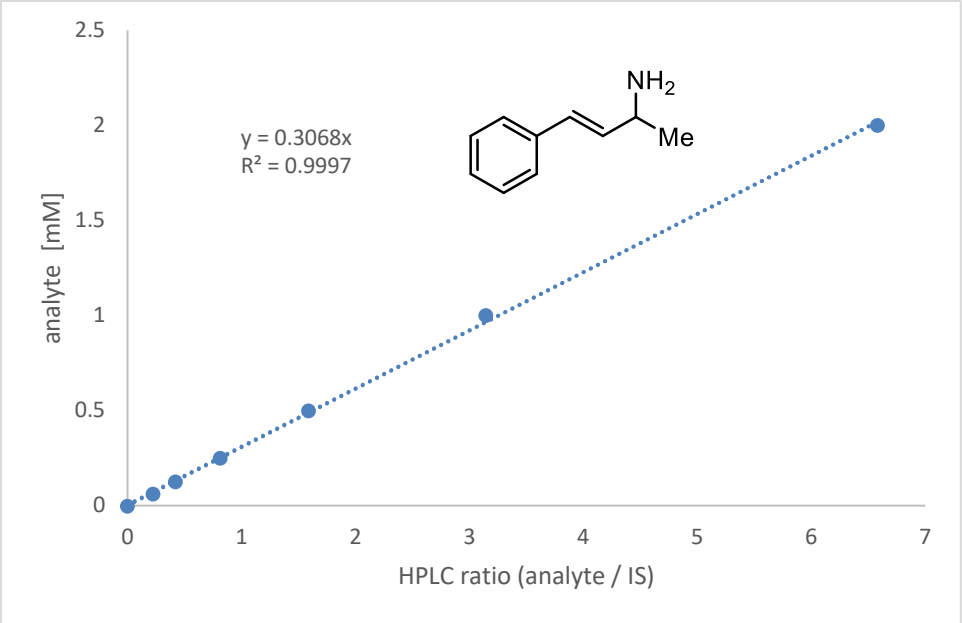
Analysis data for P411$_{BPA}$ catalyzed primary amination of alkane **1l** (2 mM, OD_{600} = 24):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/μM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	234.6	1942.7	0.121	0.32	2.94	120	10	17	1
2	265.8	1942.5	0.137	0.36	2.94				
3	273.1	1924.6	0.142	0.37	2.9				
4	237.5	1937.9	0.123	0.32	2.9				

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.

Allylic amines 6a-6p and byproduct 7a:

6a: based on UV absorbance at 254 nm and using ethyl benzoate as internal standard.

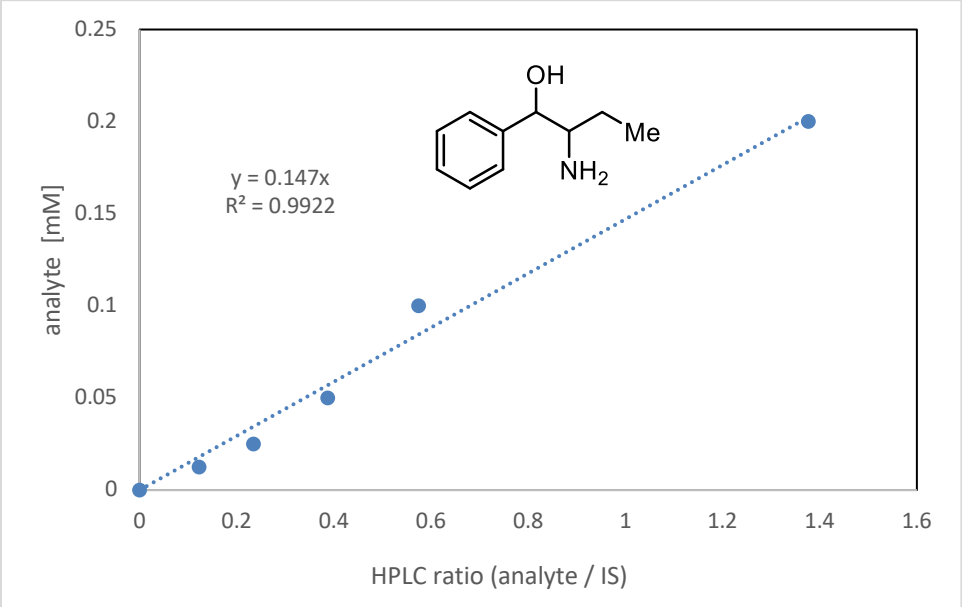


Analysis data for P411$_{APA}$ catalyzed primary amination of alkane **5a** (2 mM, OD_{600} = 4):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/μM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	3787.8	635.4	5.961	1.83	0.51	3930	310	93	2
2	3881.7	622.3	6.238	1.91	0.51				
3	3852.8	634.5	6.072	1.86	0.44				
4	3704	624.9	5.927	1.82	0.44				

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.

7a: based on ion count at single ion monitoring mode and using 2-amino-1,2,3,4-tetrahydronaphthalen-1-ol as internal standard.

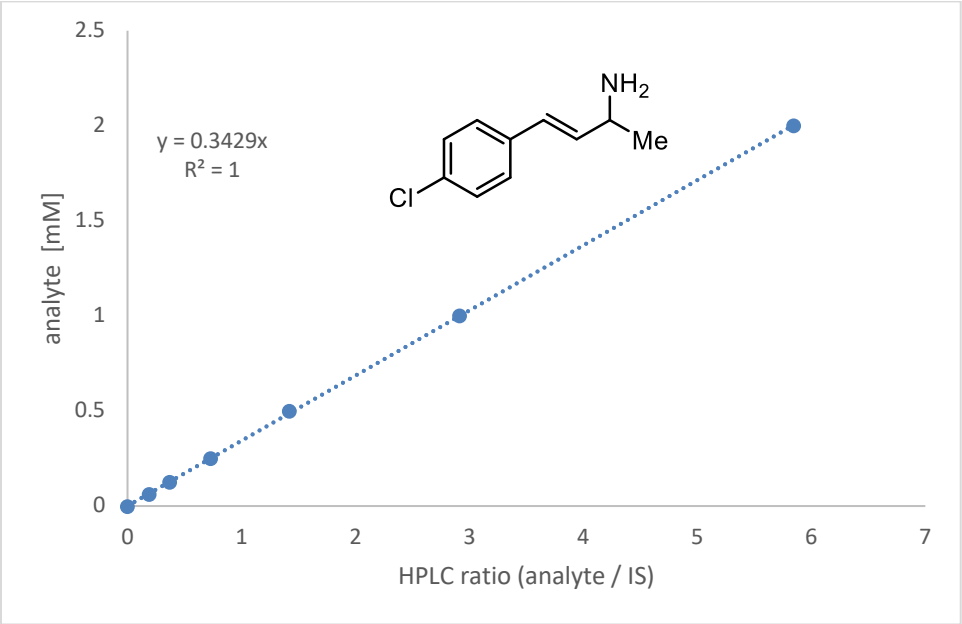


Analysis data for P411$_{APA}$ catalyzed primary amination of alkane **5a** (2 mM, OD_{600} = 4):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	Avg. Yield [%]	SD Yield [%]
1	2677411	6121475	0.437	0.06	3.7	0.5
2	3260312	5441517	0.599	0.09		
3	2916000	5792005	0.503	0.07		
4	2832826	5926976	0.478	0.07		

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.

6b: based on UV absorbance at 254 nm and using ethyl benzoate as internal standard.

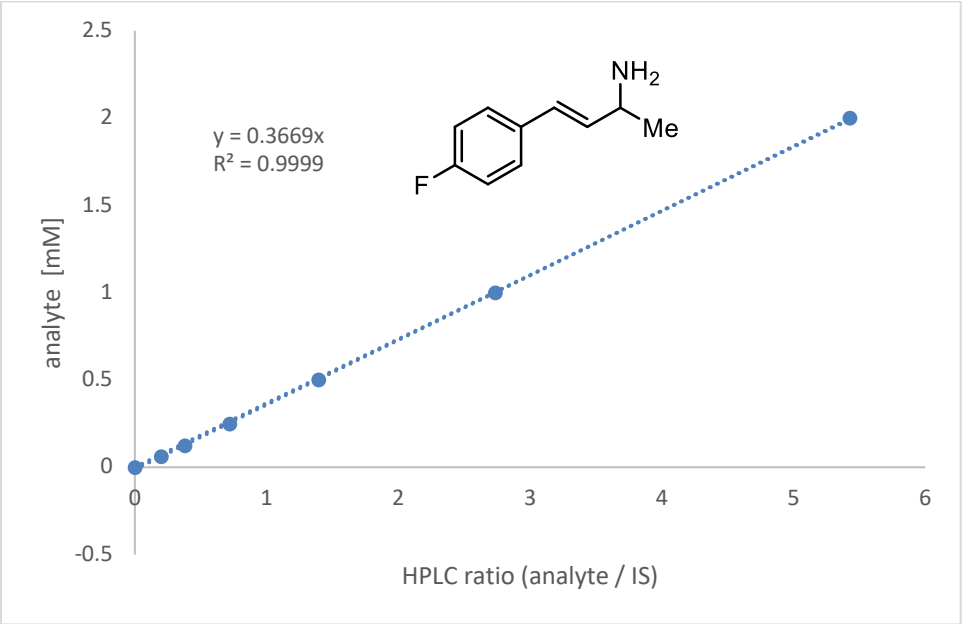


Analysis data for P411$_{APA}$ catalyzed primary amination of alkane **5b** (2 mM, OD_{600} = 16):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/µM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	2518	617.1	4.080	1.40	2.03	710	30	73	3
2	2910.5	663.5	4.387	1.50	2.03				
3	3017.3	678.5	4.447	1.52	2.13				
4	2792.2	666.7	4.188	1.44	2.13				

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.

6c: based on UV absorbance at 254 nm and using ethyl benzoate as internal standard.

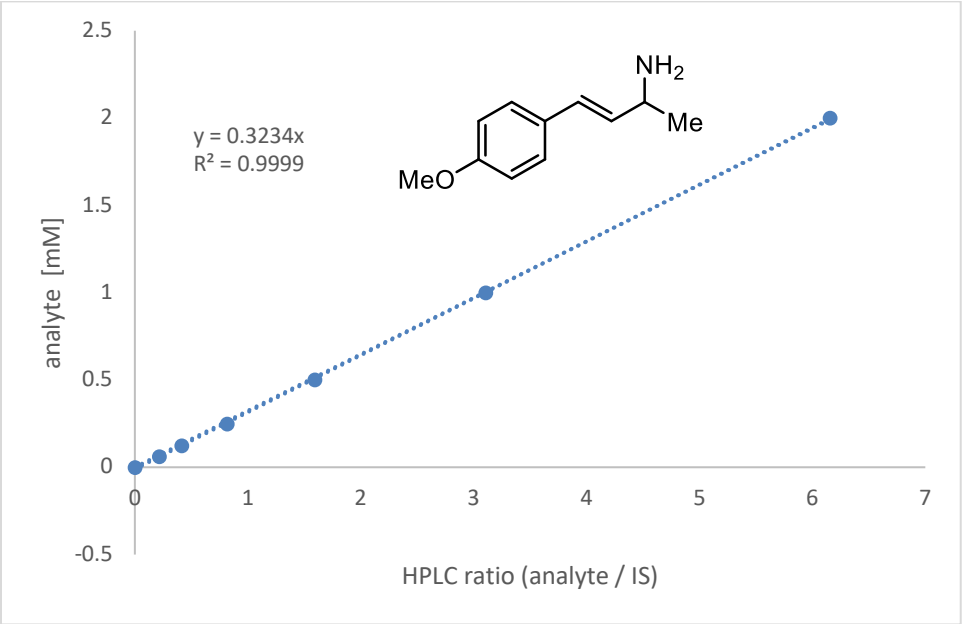


Analysis data for P411$_{APA}$ catalyzed primary amination of alkane **5c** (2 mM, OD_{600} = 4):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/μM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	2292.2	586.2	3.910	1.43	0.51	2780	300	68	5
2	2296.8	595.8	3.855	1.41	0.51				
3	2274.4	605.5	3.756	1.38	0.44				
4	2044.6	609.7	3.353	1.23	0.44				

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.

6d: based on UV absorbance at 254 nm and using ethyl benzoate as internal standard.

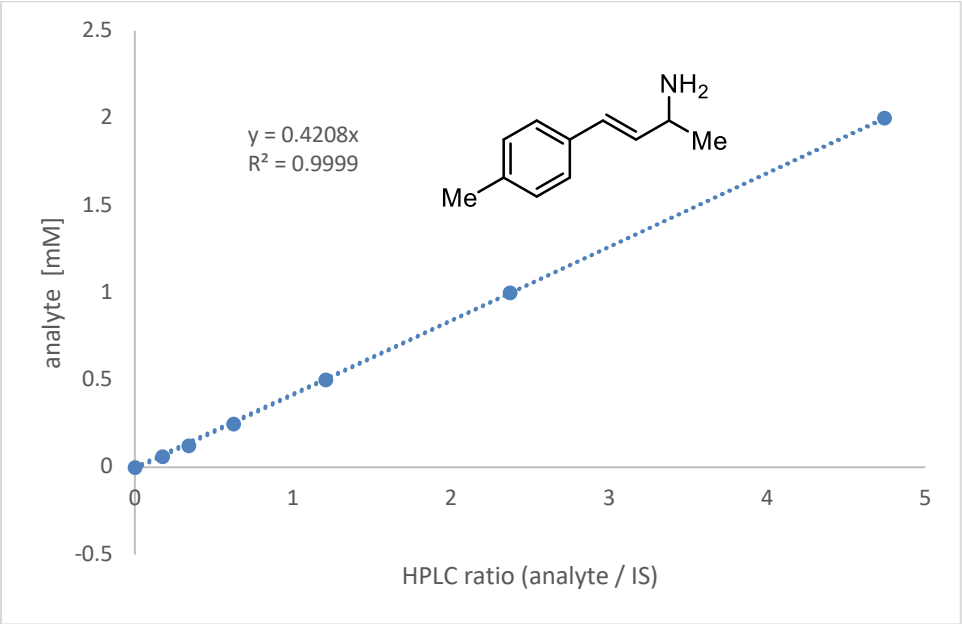


Analysis data for P411$_{APA}$ catalyzed primary amination of alkane **5d** (2 mM, OD_{600} = 4):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/μM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	1573	594.8	2.645	0.86	0.51	1610	100	38	4
2	1417.9	596.3	2.378	0.77	0.51				
3	1394.8	597.9	2.333	0.75	0.44				
4	1236.9	594.7	2.080	0.67	0.44				

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.

6e: based on UV absorbance at 254 nm and using ethyl benzoate as internal standard.

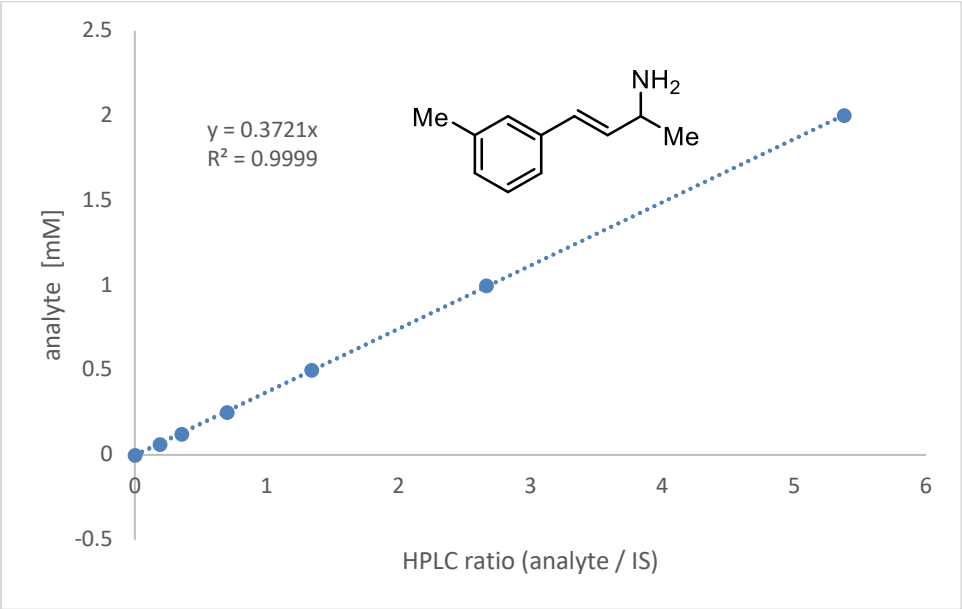


Analysis data for P411$_{APA}$ catalyzed primary amination of alkane **5e** (2 mM, OD_{600} = 16):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/μM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	1724.1	649.1	2.656	1.12	2.03	520	30	55	2
2	1624	650.6	2.496	1.05	2.03				
3	1668.3	662.3	2.519	1.06	2.13				
4	1699.3	621.7	2.733	1.15	2.13				

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.

6f: based on UV absorbance at 254 nm and using ethyl benzoate as internal standard.

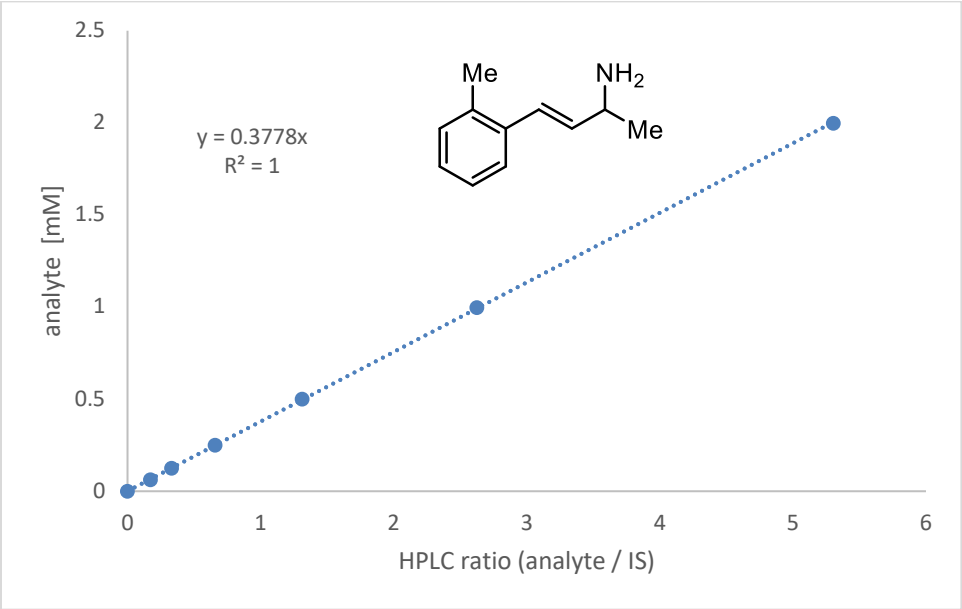


Analysis data for P411$_{APA}$ catalyzed primary amination of alkane **5f** (2 mM, OD_{600} = 16):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/μM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	2293.1	658.4	3.483	1.30	2.03	540	100	55	12
2	1600.4	668.4	2.394	0.89	2.03				
3	2312.7	650	3.558	1.32	2.13				
4	1749.8	658.3	2.658	0.99	2.13				

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.

6g: based on UV absorbance at 254 nm and using ethyl benzoate as internal standard.

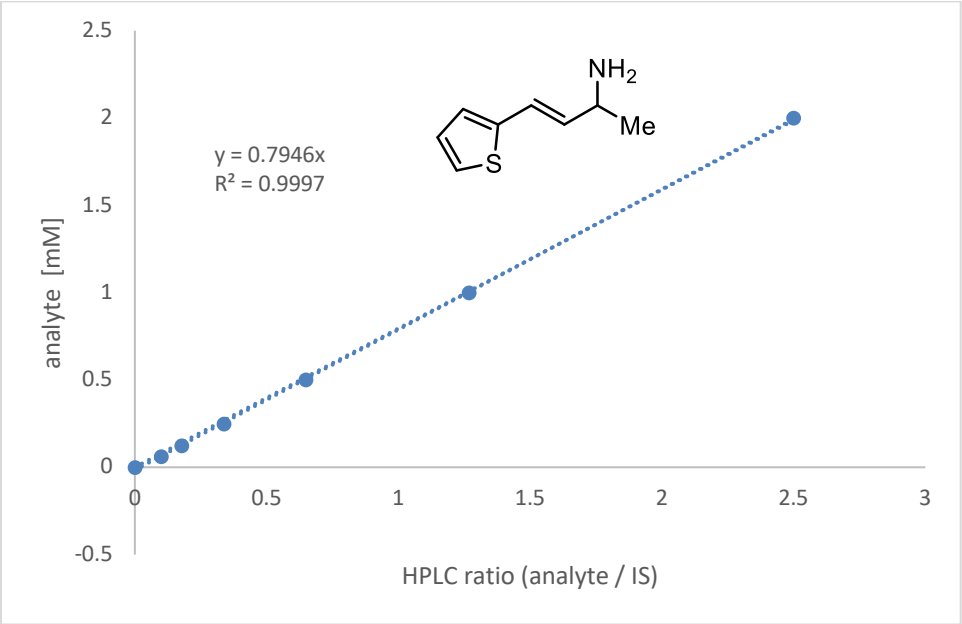


Analysis data for P411$_{APA}$ catalyzed primary amination of alkane **5g** (2 mM, OD_{600} = 8):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/μM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	1218.6	662.5	1.839	0.69	1.01	770	60	40	4
2	1341	649.7	2.064	0.78	1.01				
3	1480.5	667.7	2.217	0.84	1.07				
4	1542.1	650.9	2.369	0.90	1.07				

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.

6h: based on UV absorbance at 254 nm and using ethyl benzoate as internal standard.

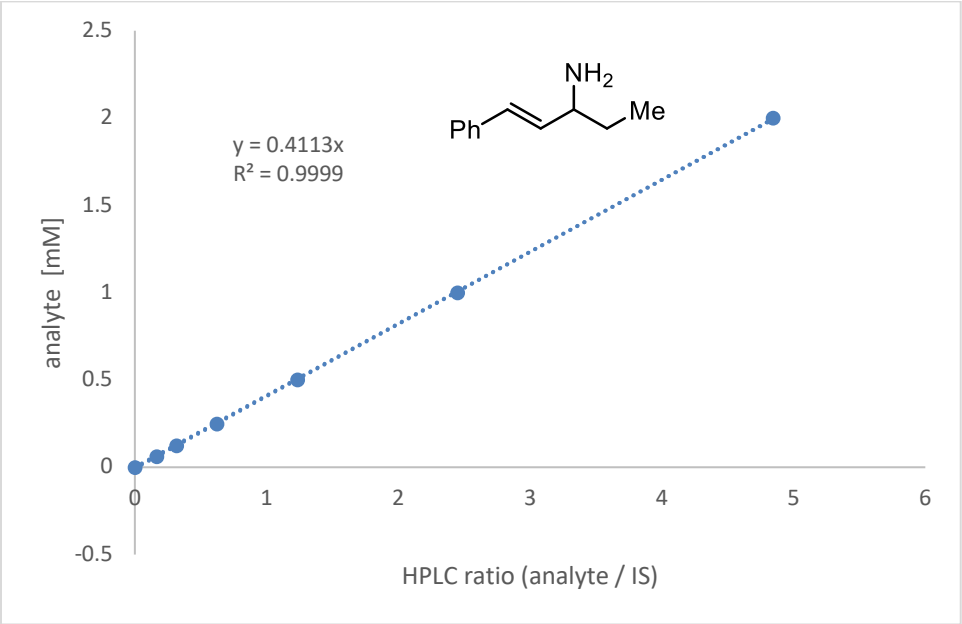


Analysis data for P411$_{APA}$ catalyzed primary amination of alkane **5h** (2 mM, OD_{600} = 16):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/μM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	1171.7	676.7	1.731	1.38	2.03	600	120	62	11
2	1181	672.7	1.756	1.40	2.03				
3	1055.8	658.9	1.602	1.27	2.13				
4	773.9	673.4	1.149	0.91	2.13				

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.

6i: based on UV absorbance at 254 nm and using ethyl benzoate as internal standard.

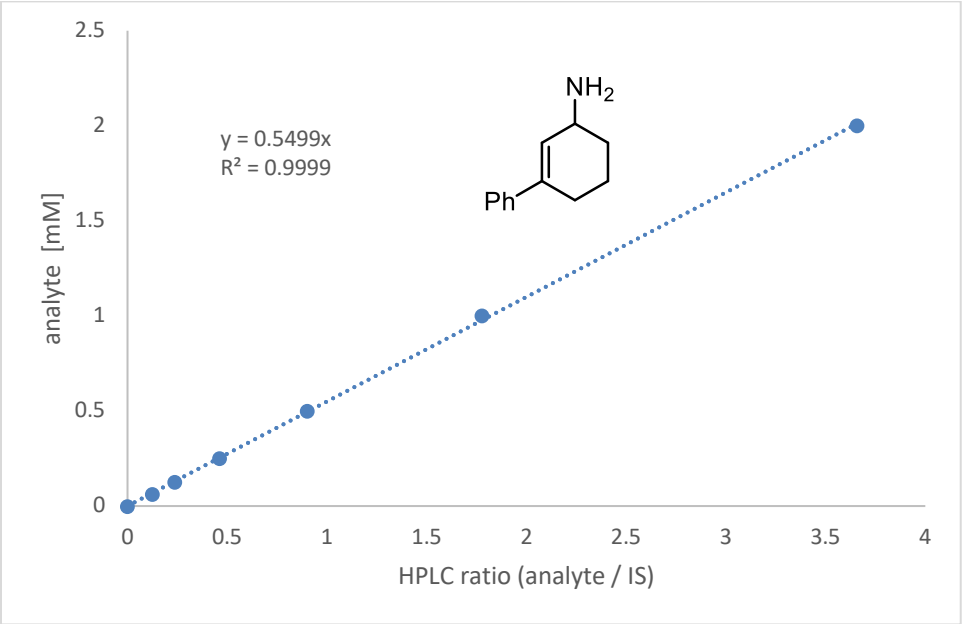


Analysis data for P411$_{APA}$ catalyzed primary amination of alkane **5i** (2 mM, OD_{600} = 4):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/μM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	1089.3	597.5	1.823	0.75	0.51	1420	120	35	2
2	1040.7	595.2	1.748	0.72	0.51				
3	984.2	602	1.635	0.67	0.44				
4	967.3	599.2	1.614	0.66	0.44				

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.

6j: based on UV absorbance at 254 nm and using ethyl benzoate as internal standard.

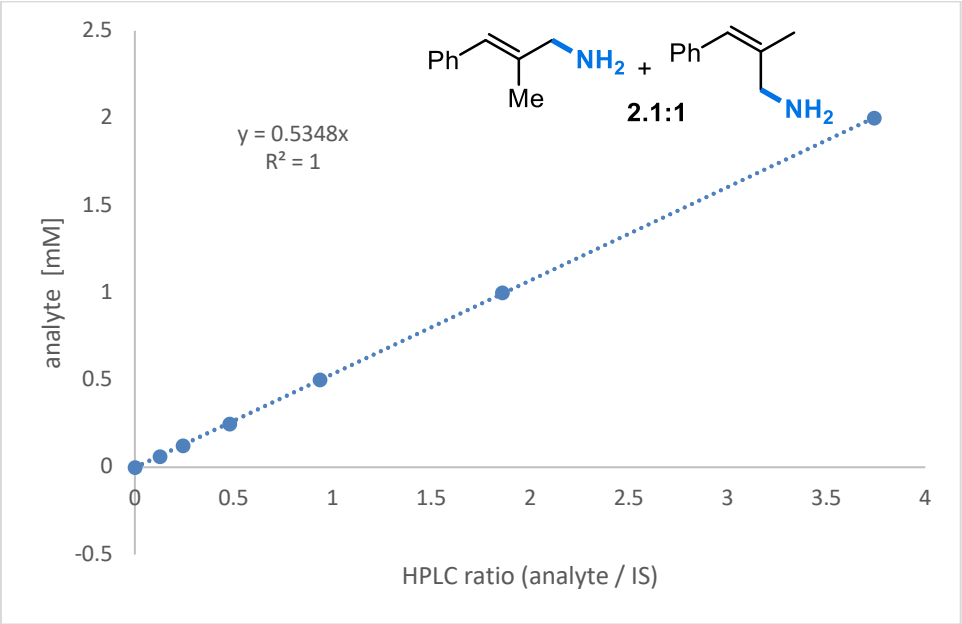


Analysis data for P411$_{APA}$ catalyzed primary amination of alkane **5j** (2 mM, OD_{600} = 8):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/μM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	703.9	637.9	1.103	0.61	1.01	560	30	29	1
2	675.9	663.5	1.019	0.56	1.01				
3	694.3	662.8	1.048	0.58	1.07				
4	713.5	666	1.071	0.59	1.07				

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.

6k: based on UV absorbance at 254 nm and using ethyl benzoate as internal standard.

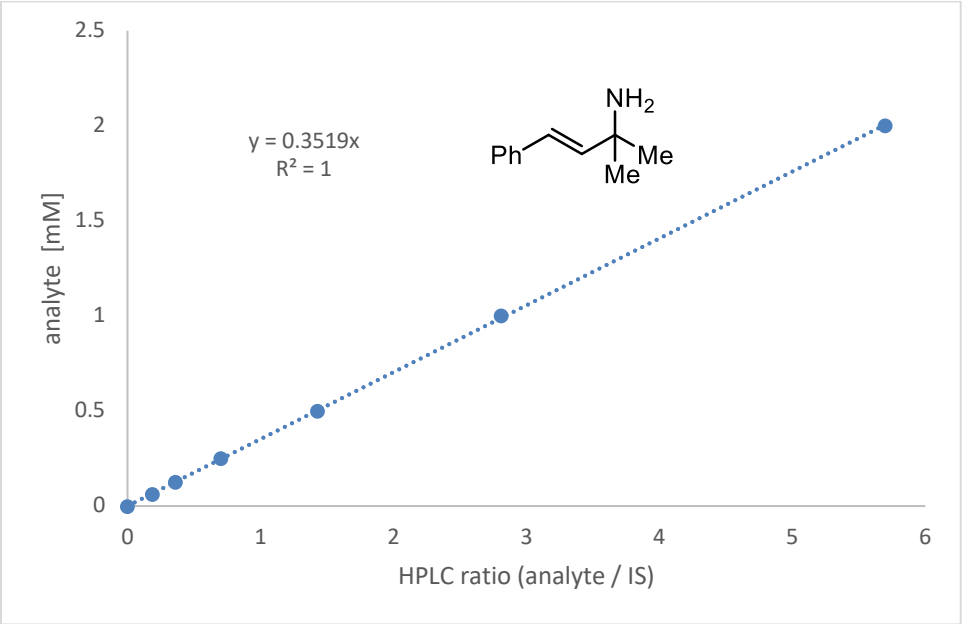


Analysis data for P411$_{APA}$ catalyzed primary amination of alkane **5k** (2 mM, OD_{600} = 8):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/μM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	1071.2	605.8	1.768	0.95	1.01	1060	110	50	2
2	1145.8	587.1	1.952	1.04	1.01				
3	1155.1	619	1.866	1.00	0.88				
4	1160.8	602.2	1.928	1.03	0.88				

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.

6l: based on UV absorbance at 254 nm and using ethyl benzoate as internal standard.

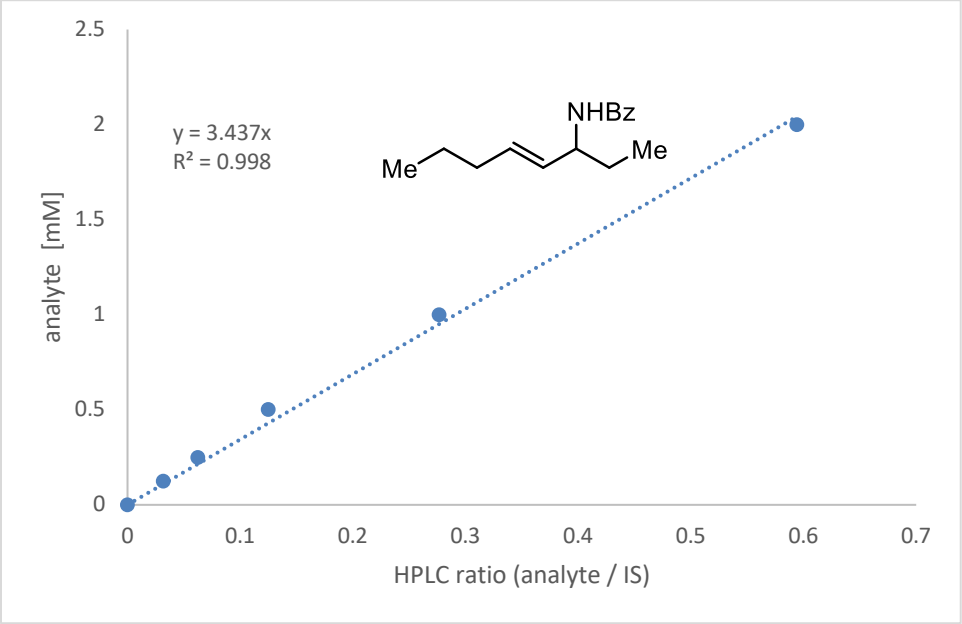


Analysis data for P411$_{APA}$ catalyzed primary amination of alkane **5l** (2 mM, OD_{600} = 16):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/μM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	1535.6	676.7	2.269	0.80	2.03	390	7	40	2
2	1465.5	673.8	2.175	0.77	2.03				
3	1579	667.2	2.367	0.83	2.13				
4	1567.3	673.8	2.326	0.82	2.13				

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.

Bz-**6m**: based on UV absorbance at 254 nm and using ethyl benzoate as internal standard.

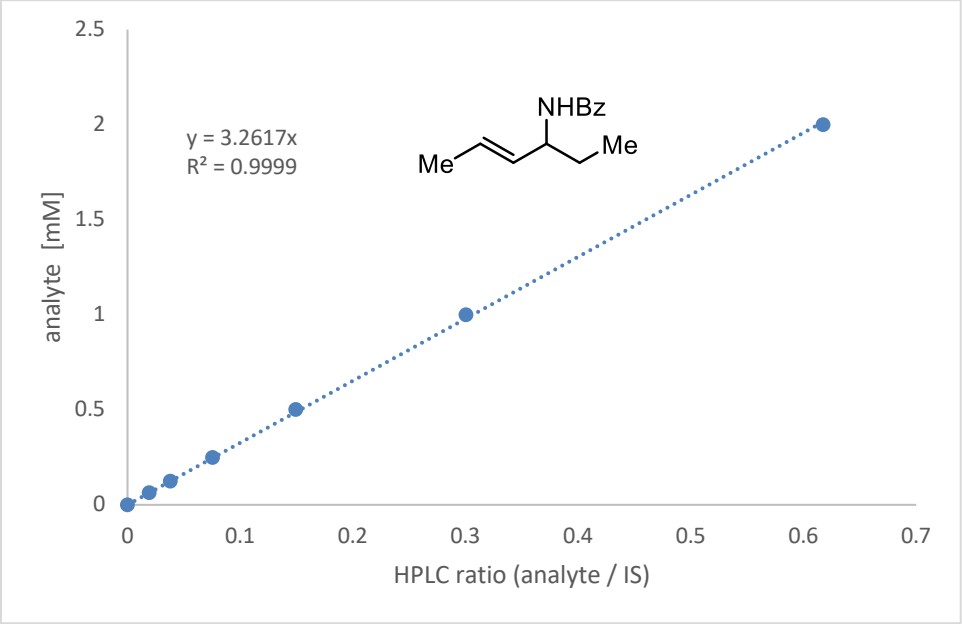


Analysis data for P411$_{APA}$ catalyzed primary amination of alkane **5m** (2 mM, OD_{600} = 16):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/μM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	260.7	654.6	0.398	1.37	2.36	550	70	67	7
2	277.9	633.5	0.439	1.51	2.36				
3	244.7	670.3	0.365	1.25	2.46				
4	230.2	662.4	0.348	1.19	2.46				

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.

Bz-**6n**-major: based on UV absorbance at 254 nm and using ethyl benzoate as internal standard.

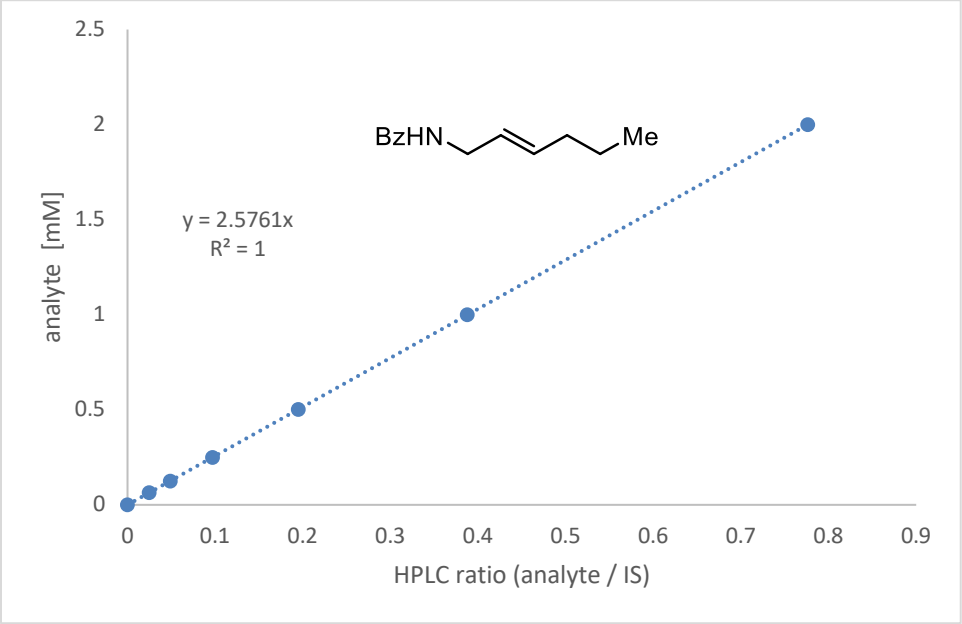


Analysis data for P411$_{APA}$ catalyzed primary amination of alkane **5n** (2 mM, OD_{600} = 16):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/μM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	147.7	672	0.220	0.72	2.36	260	40	31	4
2	131.9	668	0.197	0.64	2.36				
3	102.2	659.9	0.155	0.51	2.46				
4	130.4	671	0.194	0.63	2.46				

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.

Bz-**6n**-minor: based on UV absorbance at 254 nm and using ethyl benzoate as internal standard.

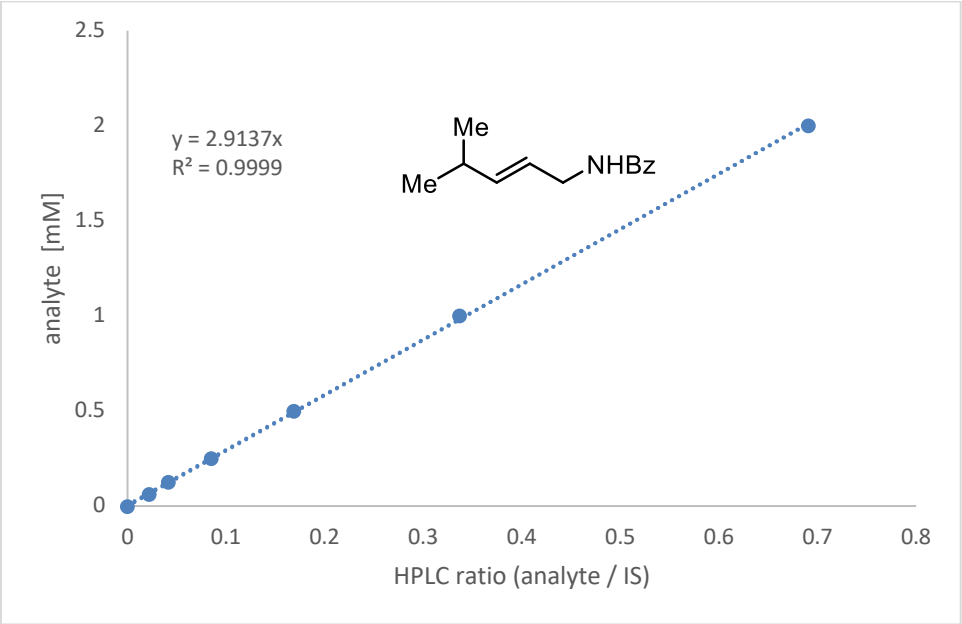


Analysis data for P411$_{APA}$ catalyzed primary amination of alkane **5n** (2 mM, OD_{600} = 16):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/μM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	86	672	0.128	0.33	2.36	160	20	19	3
2	113	668	0.169	0.44	2.36				
3	87.4	659.9	0.132	0.34	2.46				
4	107.2	671	0.160	0.41	2.46				

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.

6o: based on UV absorbance at 254 nm and using ethyl benzoate as internal standard.

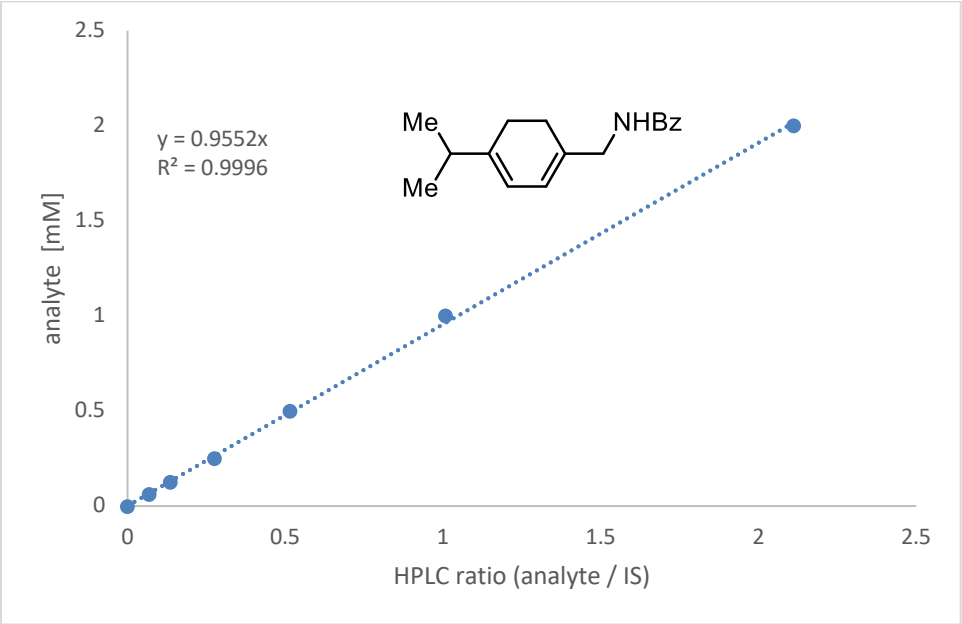


Analysis data for P411$_{APA}$ catalyzed primary amination of alkane **5o** (2 mM, OD_{600} = 16):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/μM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	233.1	632.8	0.368	1.07	2.36	420	40	51	3
2	238.3	653.4	0.365	1.06	2.36				
3	222.5	685.6	0.325	0.95	2.46				
4	218.8	665.5	0.329	0.96	2.46				

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.

6p: based on UV absorbance at 254 nm and using ethyl benzoate as internal standard.



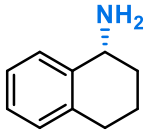
Analysis data for P411$_{APA}$ catalyzed primary amination of alkane **5p** (2 mM, OD_{600} = 8):

Entry	Pdt	IS	Pdt/IS	[Pdt]/mM	[PC]/μM	Avg. TTN	SD TTN	Avg. Yield [%]	SD Yield [%]
1	251	779.7	0.322	0.31	1.01	270	60	14	3
2	184.9	770.1	0.240	0.23	1.01				
3	190.6	771.5	0.247	0.24	1.07				
4	301.4	790.8	0.381	0.36	1.07				

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.

A.7. Assignment of Absolute Configuration and Chiral HPLC Traces

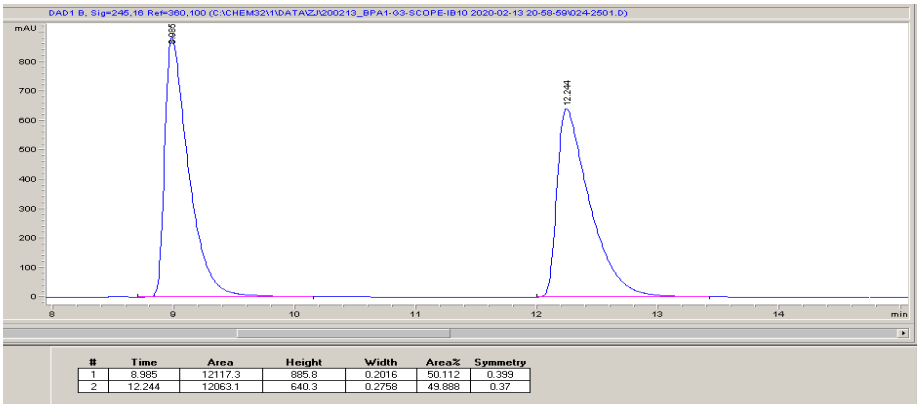
Before analysis of *ee*, the amine products were protected by benzoyl group (**Section A.1.**). The absolute stereochemistry for benzylic amine **3a** and **3i**-major was assigned to be *R* based on their commercially available (*S*)-enantiomers. The other benzylic amines **3b–3h** and **3j** were assigned by analogy. The absolute stereochemistry for allylic amine **6a** was assigned to be *R* based on its (*S*)-enantiomer synthesized by a reported method,[6] while the other linear allylic amines **6b–6i, 6m,** and **6n** were assigned by analogy. The absolute stereochemistry for cyclic amine **6j** was assigned to be *S* configuration by comparing the optical rotation of its hydrochloride salt with the literature (see **Section A.4.**).[7]



(*R*)-1,2,3,4-tetrahydronaphthalen-1-amine (3a)

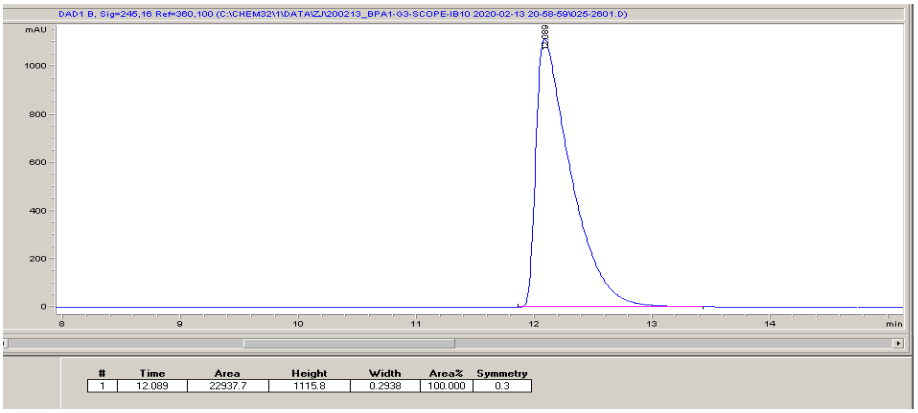
Conditions: Chiralpak IB, 10% *i*-PrOH in hexane, 1.0 mL/min, 254 nm

Racemic 3a:



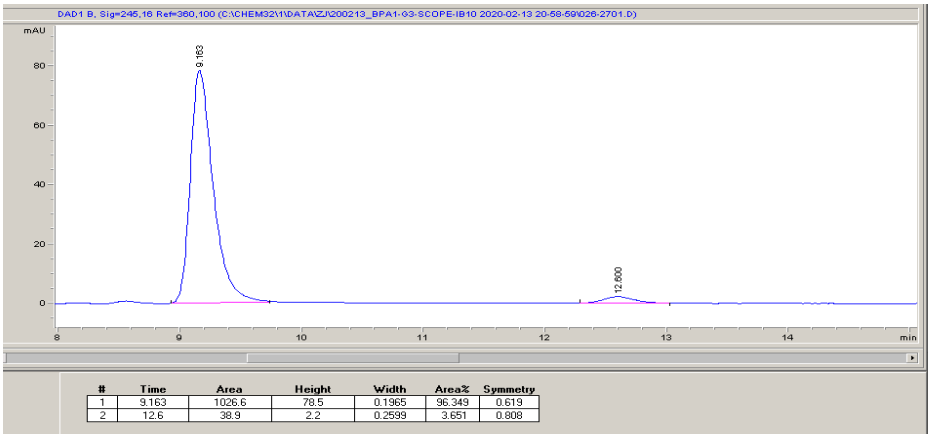
#	Time	Area	Height	Width	Area%	Symmetry
1	8.985	12117.3	885.8	0.2016	50.112	0.399
2	12.244	12063.1	640.3	0.2758	49.888	0.37

***S*-enantiomer of 3a**:



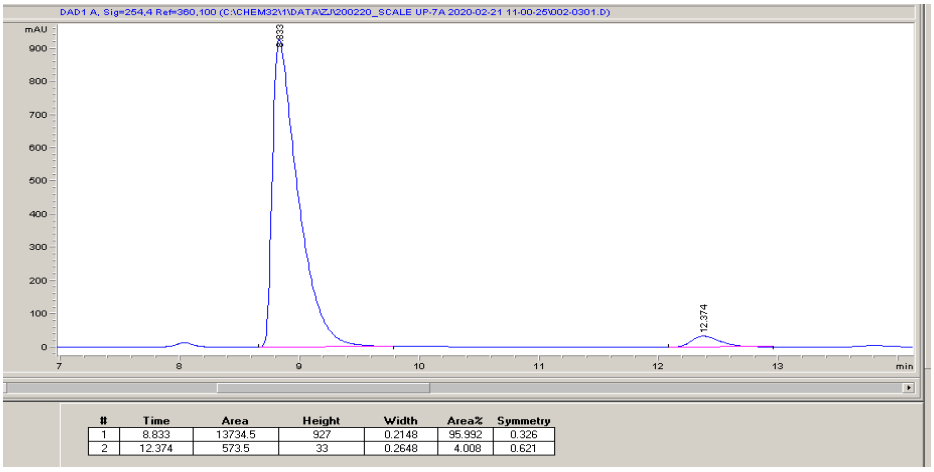
#	Time	Area	Height	Width	Area%	Symmetry
1	12.089	22937.7	1115.8	0.2938	100.000	0.3

Enzymatic product 3a: 93% *ee*

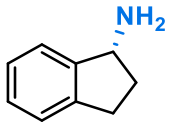


#	Time	Area	Height	Width	Area%	Symmetry
1	9.163	1026.6	78.5	0.1965	96.349	0.619
2	12.6	38.9	2.2	0.2599	3.651	0.808

Enzymatic product 3a from the scale-up reaction: 92% *ee*



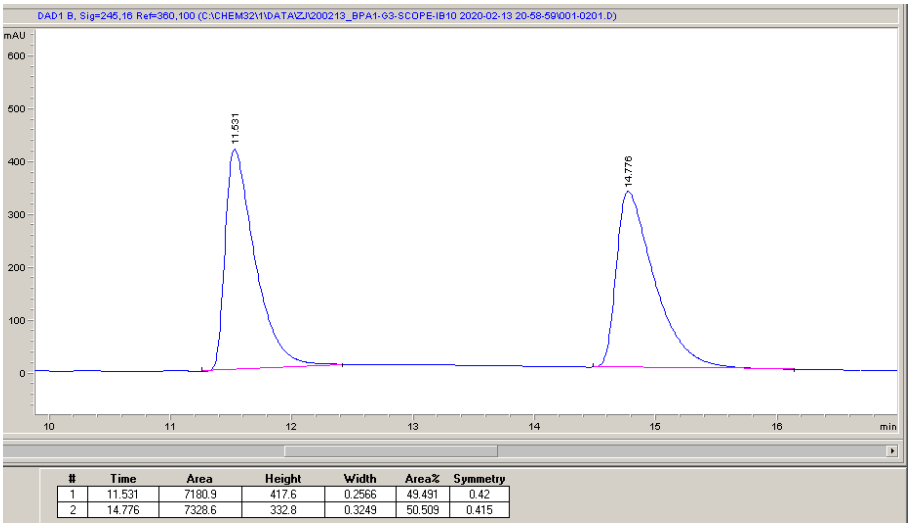
#	Time	Area	Height	Width	Area%	Symmetry
1	8.833	13734.5	927	0.2148	95.992	0.326
2	12.374	573.5	33	0.2648	4.008	0.621



(*R*)-2,3-dihydro-1H-inden-1-amine (3b)

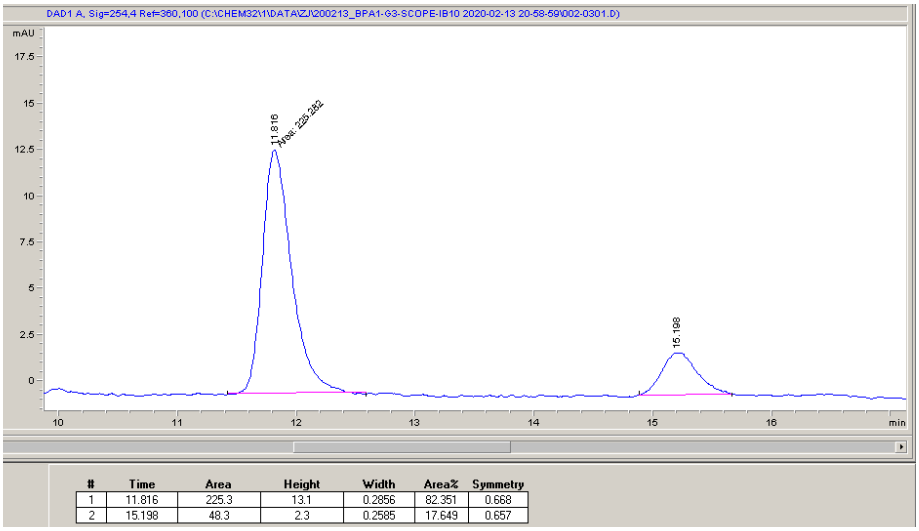
Conditions: Chiralpak IB, 10% *i*-PrOH in hexane, 1.0 mL/min, 254 nm

Racemic 3b:

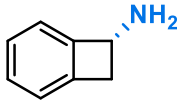


#	Time	Area	Height	Width	Area%	Symmetry
1	11.531	7180.9	417.6	0.2566	49.491	0.42
2	14.776	7328.6	332.8	0.3249	50.509	0.415

Enzymatic product 3b: 65% *ee*



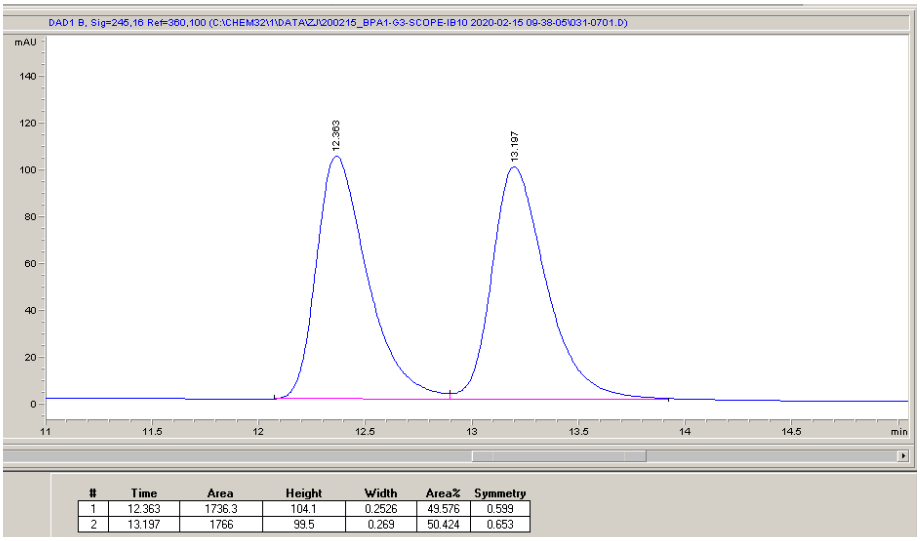
#	Time	Area	Height	Width	Area%	Symmetry
1	11.816	225.3	13.1	0.2856	82.351	0.668
2	15.198	48.3	2.3	0.2585	17.649	0.657



(*R*)-bicyclo[4.2.0]octa-1,3,5-trien-7-amine (3c)

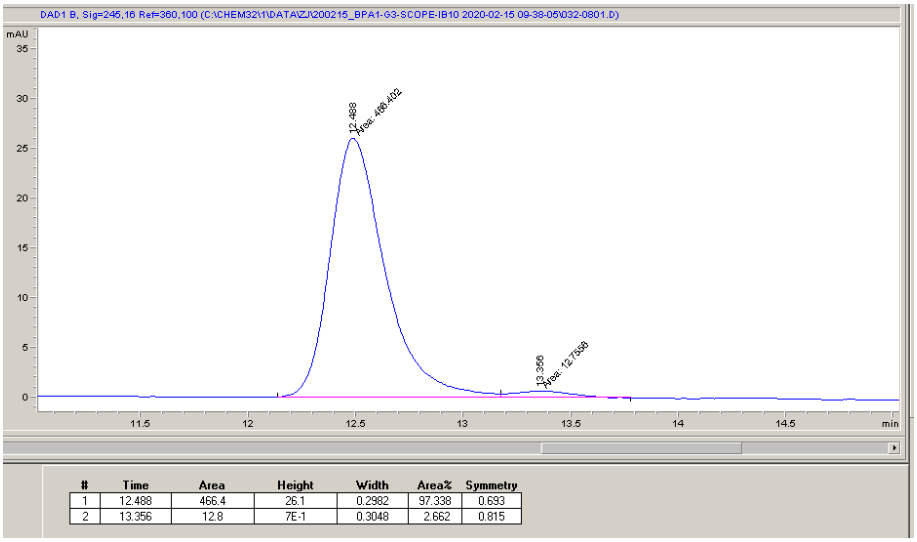
Conditions: Chiralpak IB, 10% *i*-PrOH in hexane, 1.0 mL/min, 245 nm

Racemic 3c:

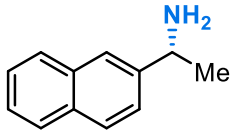


#	Time	Area	Height	Width	Area%	Symmetry
1	12.363	1736.3	104.1	0.2526	49.576	0.599
2	13.197	1766	99.5	0.269	50.424	0.653

Enzymatic product 3c: 95% *ee*



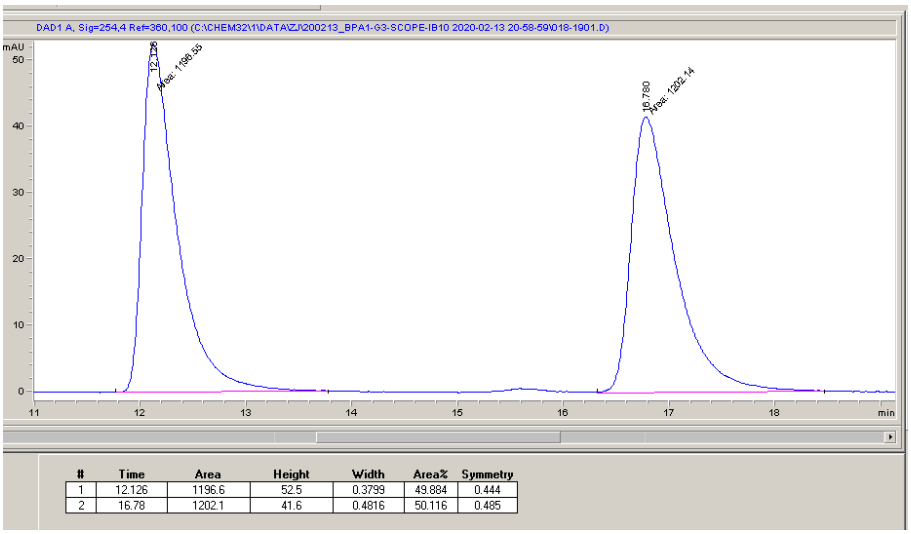
#	Time	Area	Height	Width	Area%	Symmetry
1	12.488	466.4	26.1	0.2982	97.338	0.693
2	13.356	12.8	7E-1	0.3048	2.662	0.815



(*R*)-1-(naphthalen-2-yl)ethan-1-amine (3d)

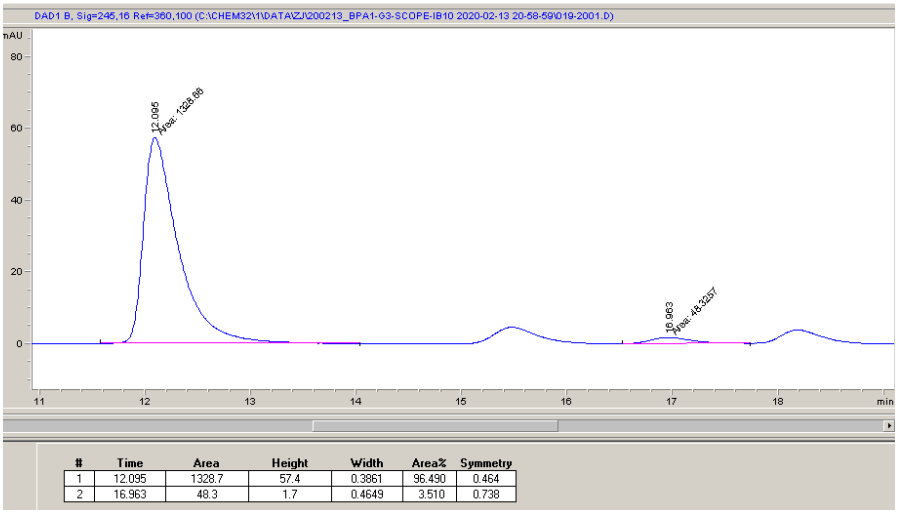
Conditions: Chiralpak IB, 10% *i*-PrOH in hexane, 1.0 mL/min, 245 nm

Racemic 3d:

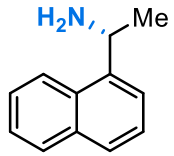


#	Time	Area	Height	Width	Area%	Symmetry
1	12.126	1196.6	52.5	0.3799	49.884	0.444
2	16.78	1202.1	41.6	0.4816	50.116	0.485

Enzymatic product 3d: 93% *ee*



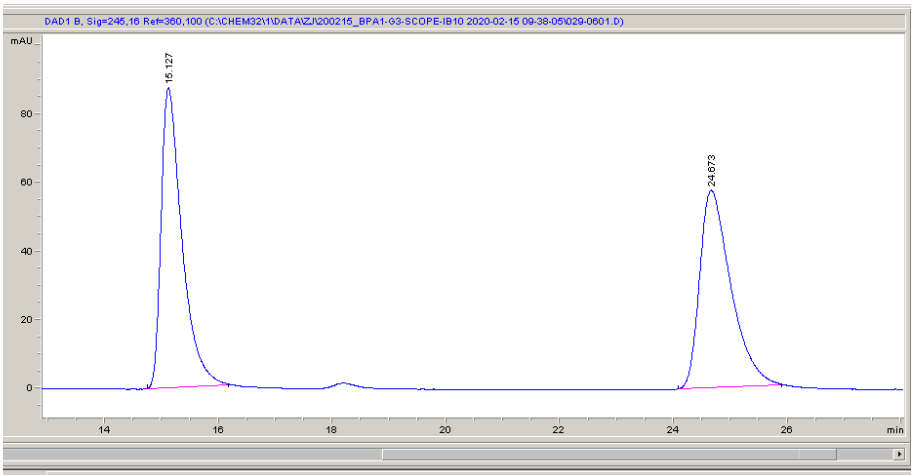
#	Time	Area	Height	Width	Area%	Symmetry
1	12.095	1328.7	57.4	0.3861	96.490	0.464
2	16.963	48.3	1.7	0.4649	3.510	0.738



(*R*)-1-(naphthalen-1-yl)ethan-1-amine (3e)

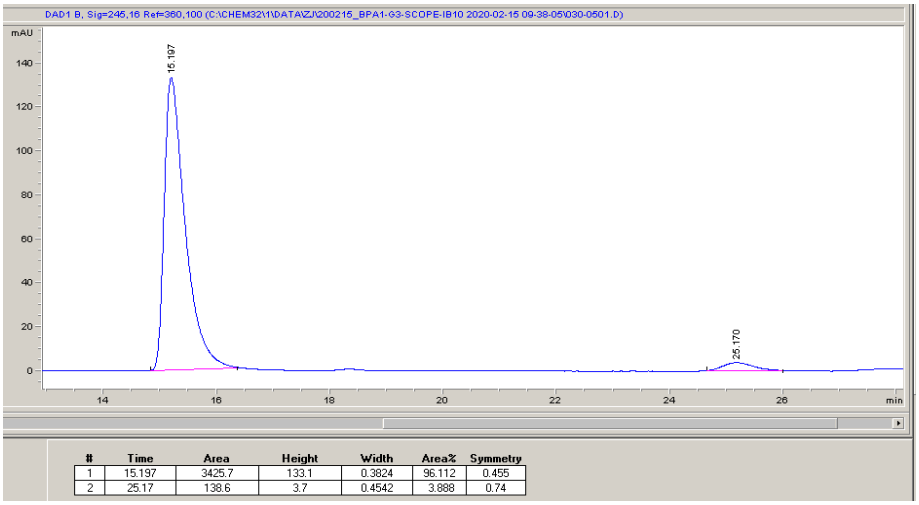
Conditions: Chiralpak IB, 10% *i*-PrOH in hexane, 1.0 mL/min, 245 nm

Racemic 3e:

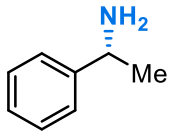


#	Time	Area	Height	Width	Area%	Symmetry
1	15.127	2233.5	87.4	0.378	50.207	0.495
2	24.673	2215	57.7	0.5742	49.793	0.547

Enzymatic product 3e: 92% *ee*



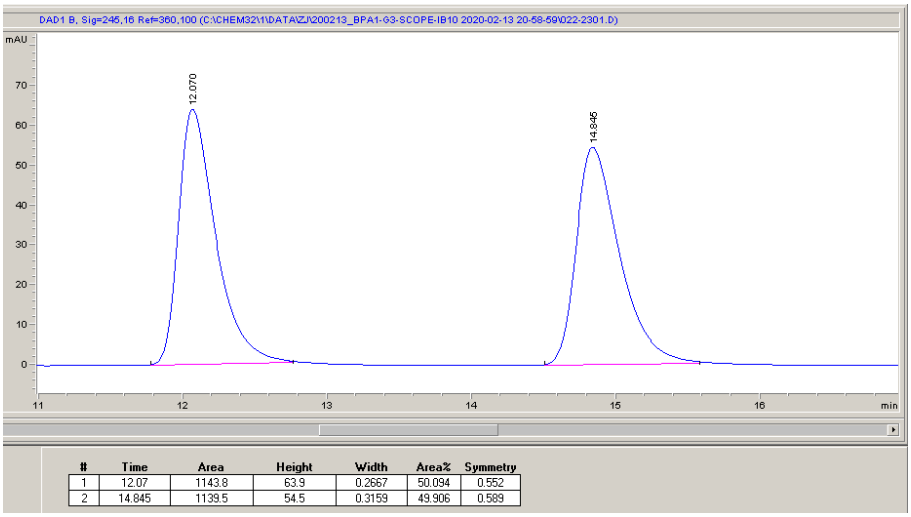
#	Time	Area	Height	Width	Area%	Symmetry
1	15.197	3425.7	133.1	0.3824	96.112	0.455
2	25.17	138.6	3.7	0.4542	3.888	0.74



(*R*)-1-phenylethan-1-amine (3f)

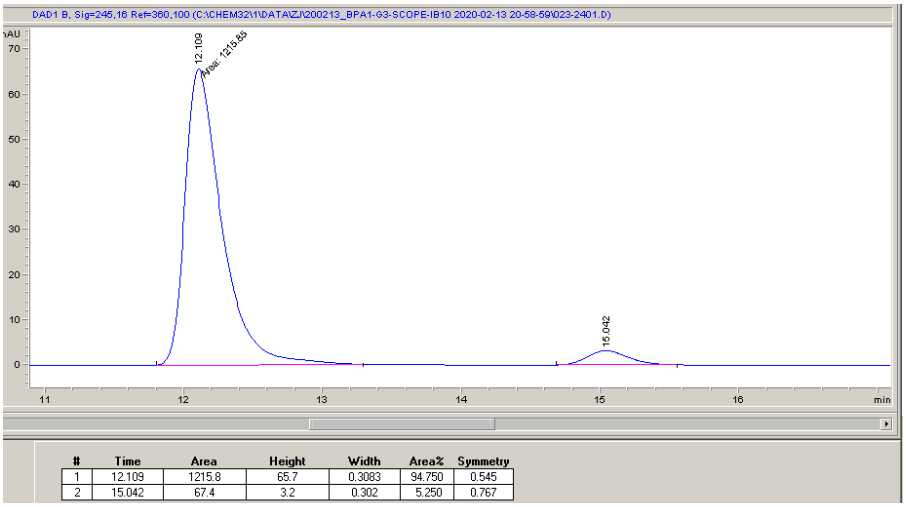
Conditions: Chiralpak IB, 10% *i*-PrOH in hexane, 1.0 mL/min, 245 nm

Racemic 3f:

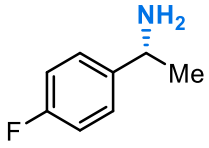


#	Time	Area	Height	Width	Area%	Symmetry
1	12.07	1143.8	63.9	0.2667	50.094	0.552
2	14.845	1139.5	54.5	0.3159	49.906	0.589

Enzymatic product 3f: 90% *ee*



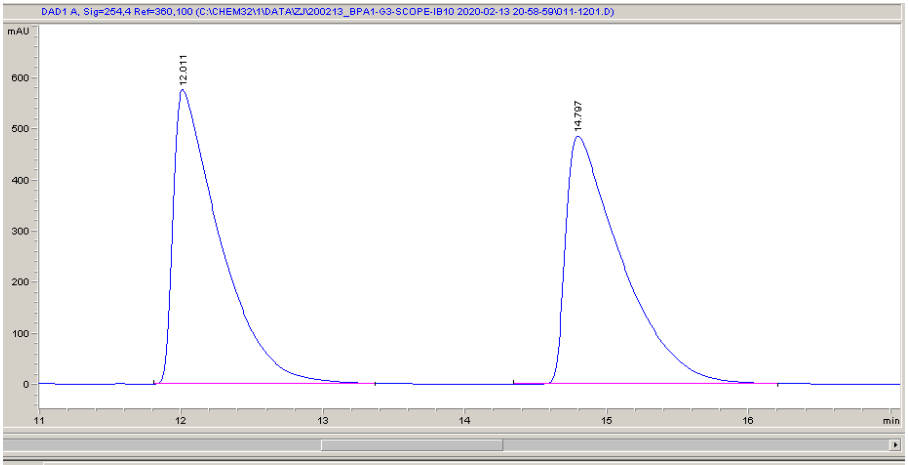
#	Time	Area	Height	Width	Area%	Symmetry
1	12.109	1215.8	65.7	0.3083	94.750	0.545
2	15.042	67.4	3.2	0.302	5.250	0.767



(*R*)-1-(4-fluorophenyl)ethan-1-amine (3g)

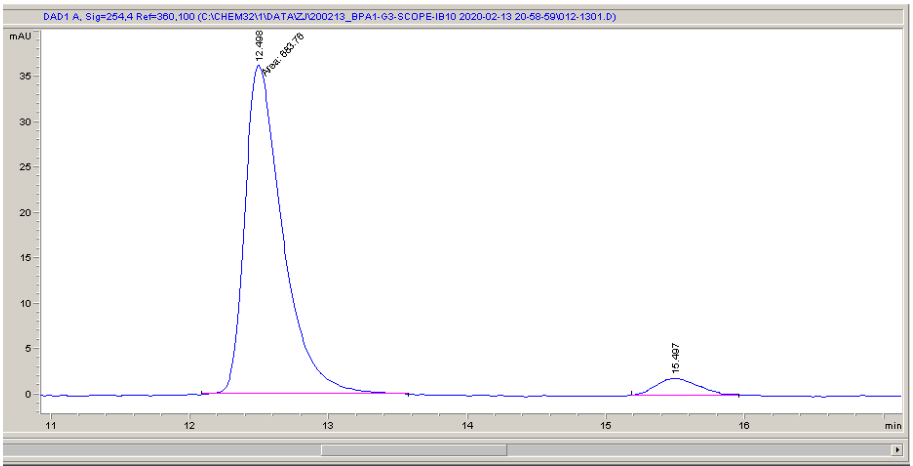
Conditions: Chiralpak IB, 10% *i*-PrOH in hexane, 1.0 mL/min, 254 nm

Racemic 3g:

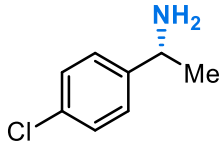


#	Time	Area	Height	Width	Area%	Symmetry
1	12.011	13202.1	575.3	0.3237	50.005	0.241
2	14.797	13199.6	484.9	0.3816	49.995	0.257

Enzymatic product 3g: 89% *ee*



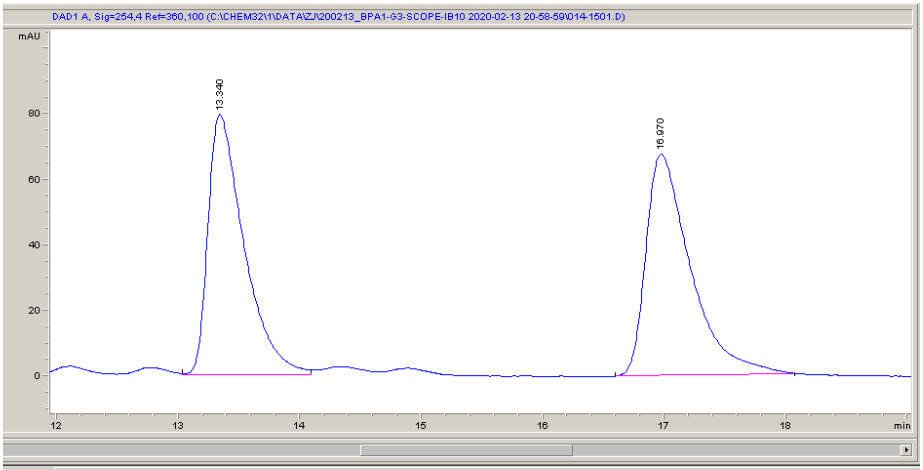
#	Time	Area	Height	Width	Area%	Symmetry
1	12.498	683.8	36.2	0.3151	94.430	0.552
2	15.497	40.3	1.9	0.2662	5.570	0.773



(*R*)-1-(4-chlorophenyl)ethan-1-amine (3h)

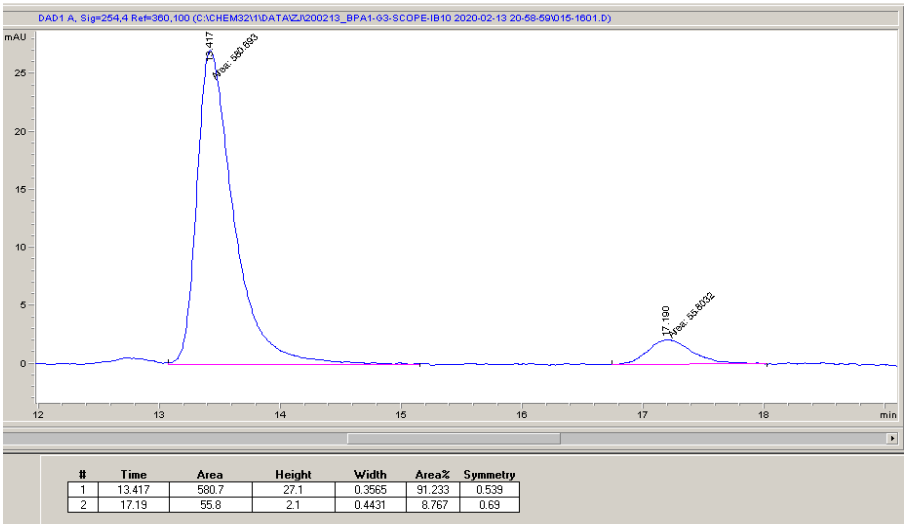
Conditions: Chiralpak IB, 10% *i*-PrOH in hexane, 1.0 mL/min, 254 nm

Racemic 3h:

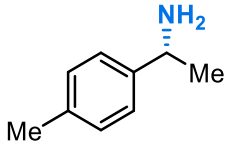


#	Time	Area	Height	Width	Area%	Symmetry
1	13.34	1646	79.5	0.3055	48.588	0.454
2	16.97	1741.6	67.7	0.3783	51.412	0.469

Enzymatic product 3h: 82% *ee*



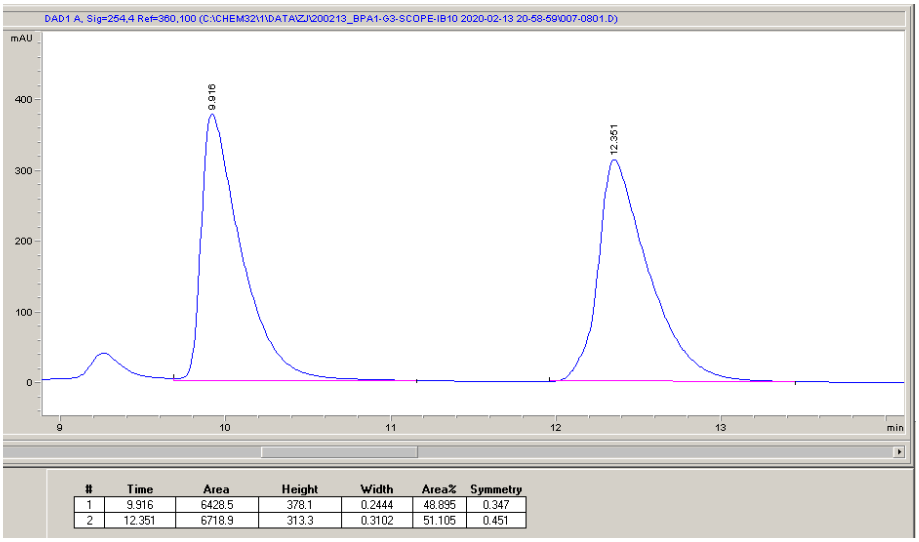
#	Time	Area	Height	Width	Area%	Symmetry
1	13.417	580.7	27.1	0.3565	91.233	0.539
2	17.19	55.8	2.1	0.4431	8.767	0.69



(*R*)-1-(4-methylphenyl)ethan-1-amine (3i-major)

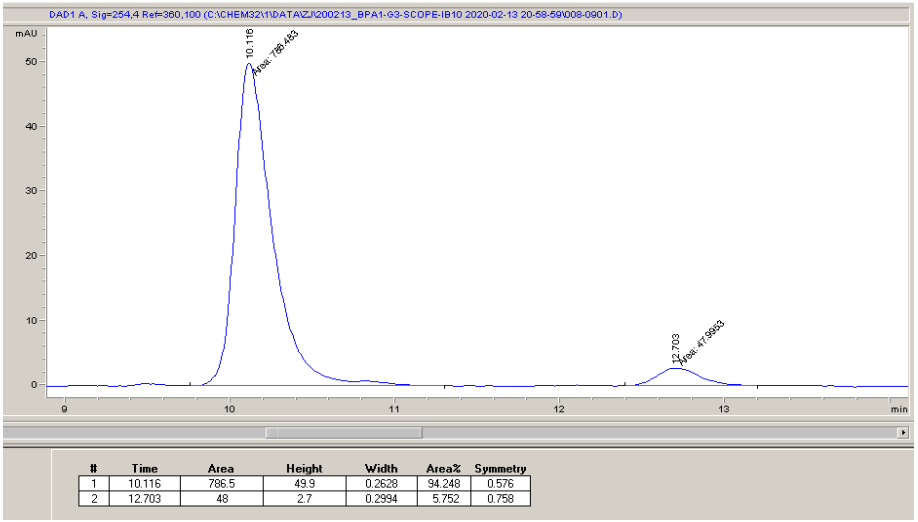
Conditions: Chiralpak IB, 10% *i*-PrOH in hexane, 1.0 mL/min, 254 nm

Racemic 3i-major:

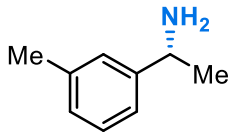


#	Time	Area	Height	Width	Area%	Symmetry
1	9.916	6428.5	378.1	0.2444	48.895	0.347
2	12.351	6718.9	313.3	0.3102	51.105	0.451

Enzymatic product 3i-major: 88% *ee*



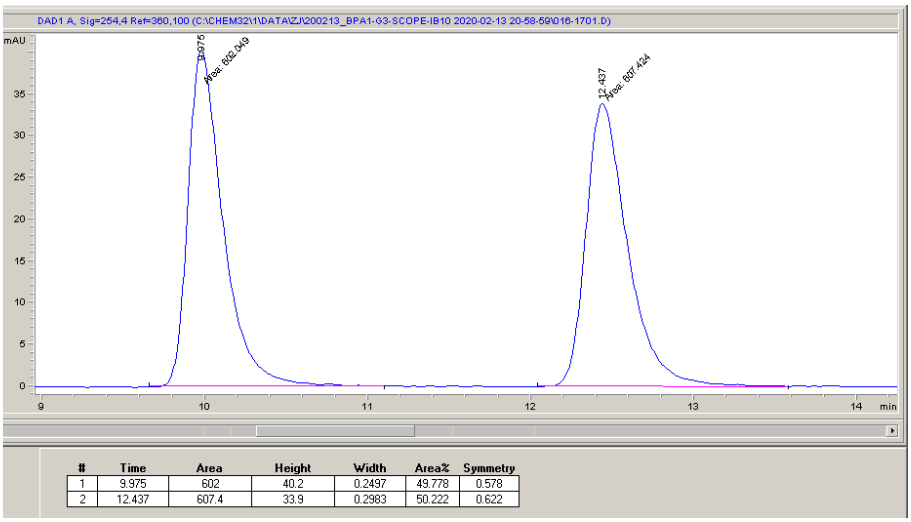
#	Time	Area	Height	Width	Area%	Symmetry
1	10.116	786.5	49.9	0.2628	94.248	0.576
2	12.703	48	2.7	0.2994	5.752	0.758



(*R*)-1-(3-methylphenyl)ethan-1-amine (3j-major)

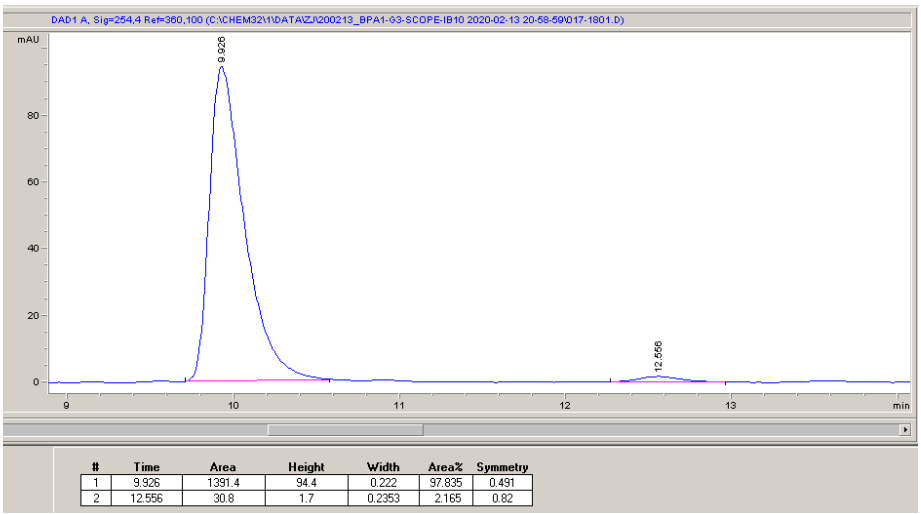
Conditions: Chiralpak IB, 10% *i*-PrOH in hexane, 1.0 mL/min, 254 nm

Racemic 3j-major:

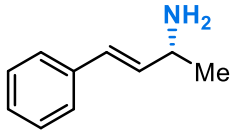


#	Time	Area	Height	Width	Area%	Symmetry
1	9.975	602	40.2	0.2497	49.778	0.578
2	12.437	607.4	33.9	0.2983	50.222	0.622

Enzymatic product 3j-major: 96% *ee*



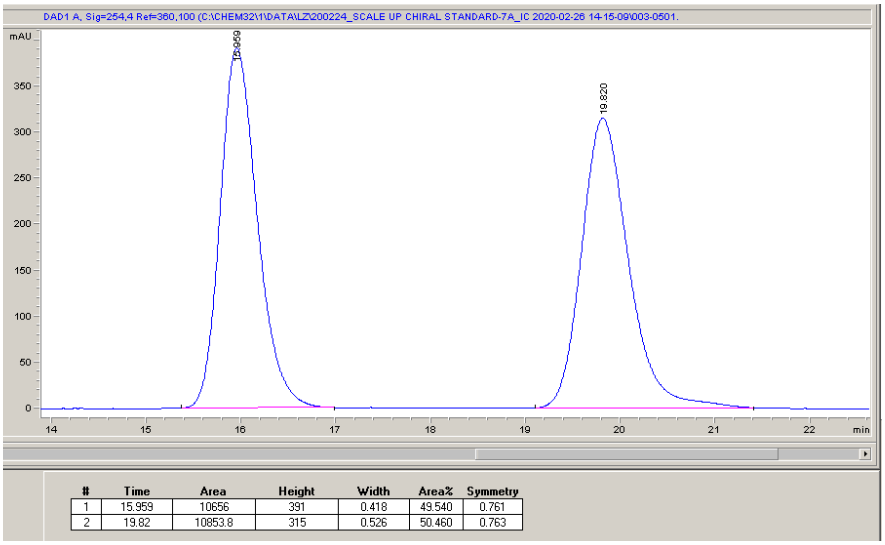
#	Time	Area	Height	Width	Area%	Symmetry
1	9.926	1391.4	94.4	0.222	97.835	0.491
2	12.556	30.8	1.7	0.2353	2.165	0.82



(*R*,*E*)-4-phenylbut-3-en-2-amine (6a)

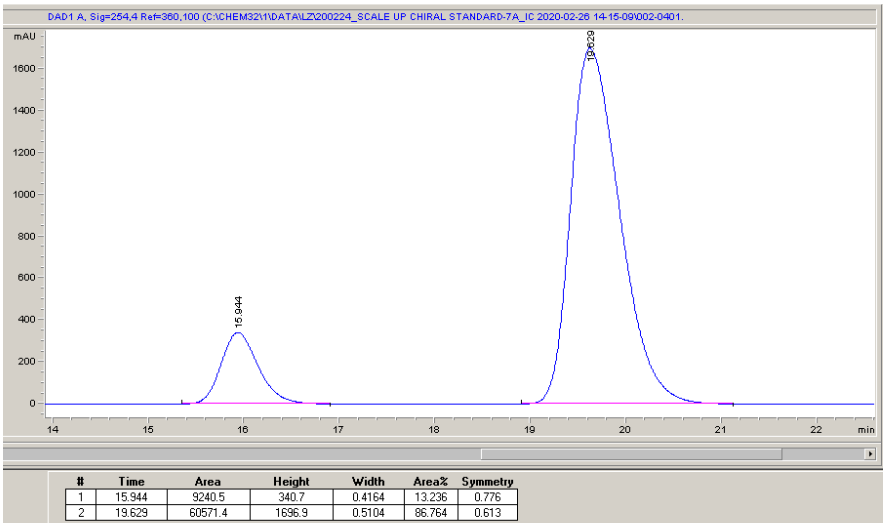
Conditions: Chiralpak IC, 10% *i*-PrOH in hexane, 1.5 mL/min, 254 nm

Racemic 6a:



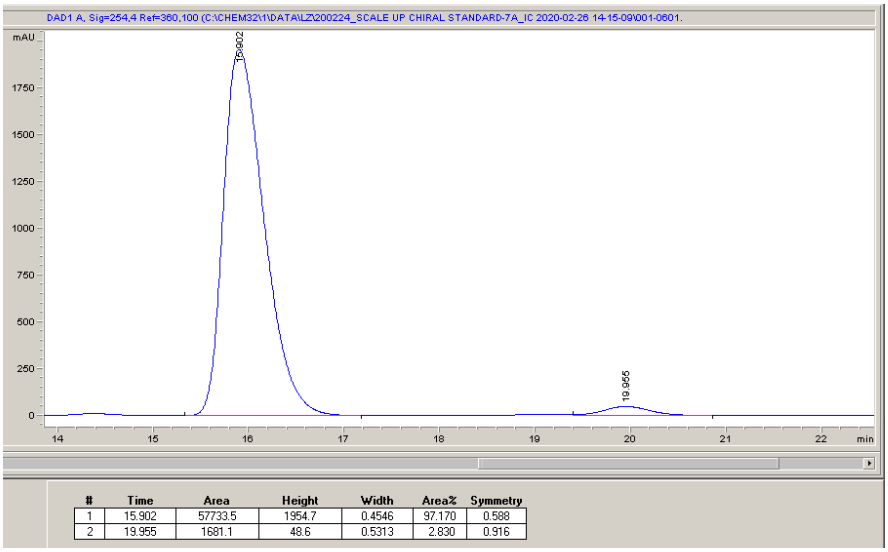
#	Time	Area	Height	Width	Area%	Symmetry
1	15.959	10656	391	0.418	49.540	0.761
2	19.82	10853.8	315	0.526	50.460	0.763

***S*-enantiomer of 6a**: –73% *ee*



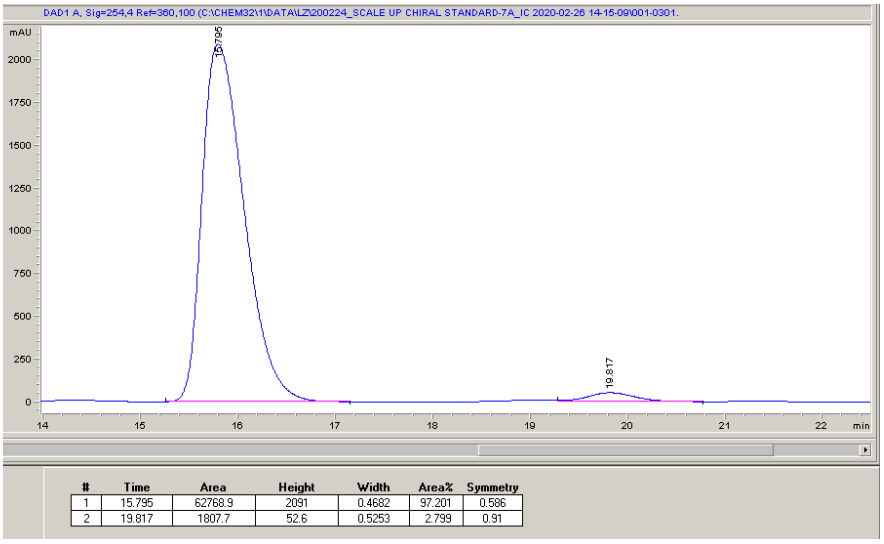
#	Time	Area	Height	Width	Area%	Symmetry
1	15.944	9240.5	340.7	0.4164	13.236	0.776
2	19.629	60571.4	1696.9	0.5104	86.764	0.613

Enzymatic product 6a: 94% *ee*

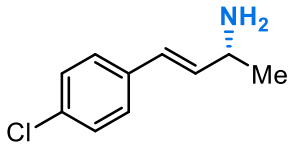


#	Time	Area	Height	Width	Area%	Symmetry
1	15.902	57733.5	1954.7	0.4546	97.170	0.588
2	19.955	1681.1	48.6	0.5313	2.830	0.916

Enzymatic product 6a from the scale-up reaction: 94% *ee*



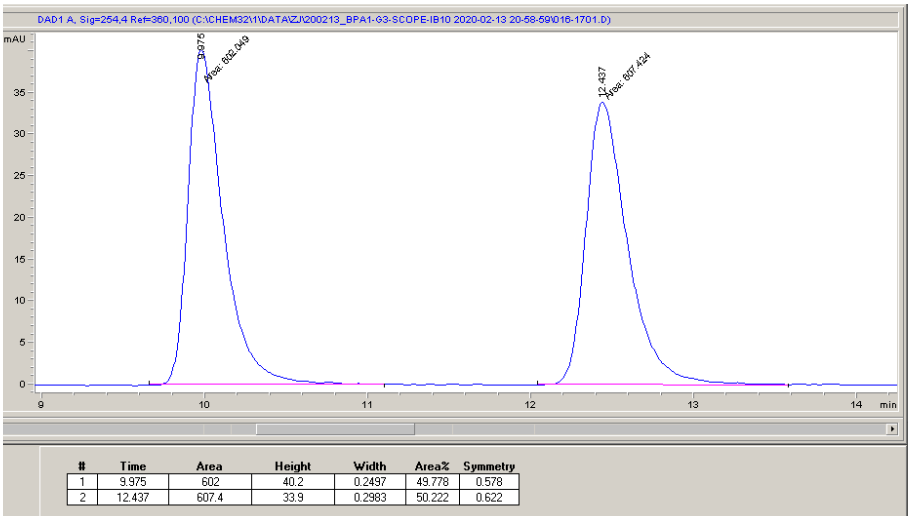
#	Time	Area	Height	Width	Area%	Symmetry
1	15.795	62768.9	2091	0.4682	97.201	0.586
2	19.817	1807.7	52.6	0.5253	2.799	0.91



(*R*,*E*)-4-(4-chlorophenyl)but-3-en-2-amine (6b)

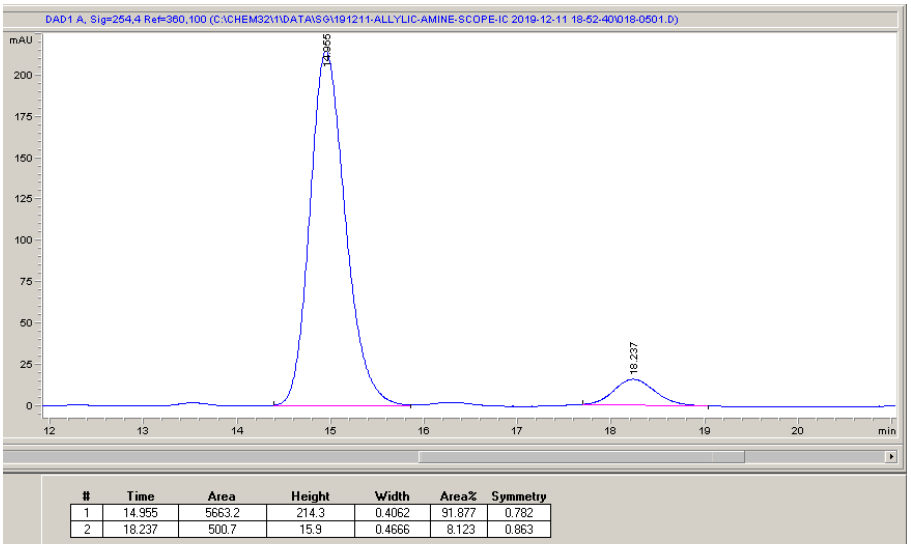
Conditions: Chiralpak IC, 10% *i*-PrOH in hexane, 1.5 mL/min, 254 nm

Racemic 6b:

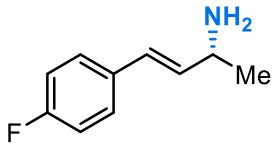


#	Time	Area	Height	Width	Area%	Symmetry
1	9.975	602	40.2	0.2497	49.778	0.578
2	12.437	607.4	33.9	0.2983	50.222	0.622

Enzymatic product 6b: 84% *ee*



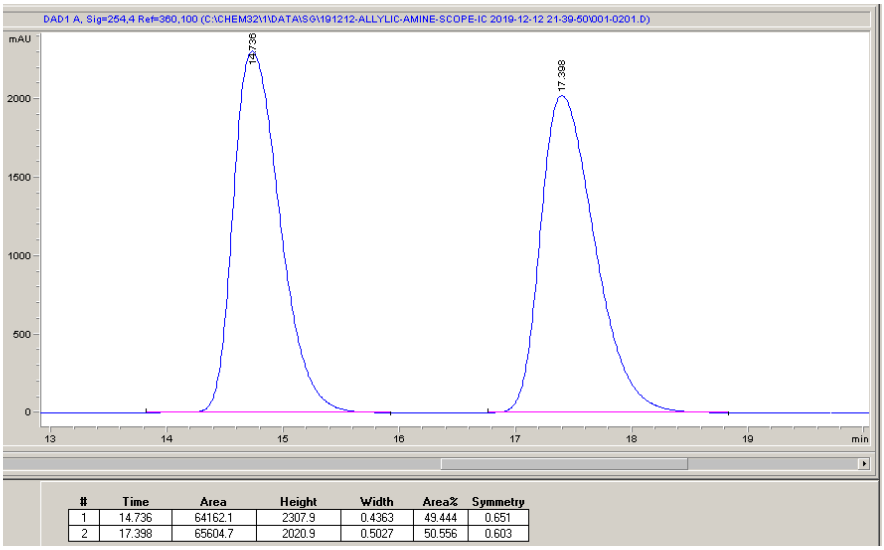
#	Time	Area	Height	Width	Area%	Symmetry
1	14.955	5663.2	214.3	0.4062	91.877	0.782
2	18.237	500.7	15.9	0.4666	8.123	0.863



(*R*,*E*)-4-(4-fluorophenyl)but-3-en-2-amine (6c)

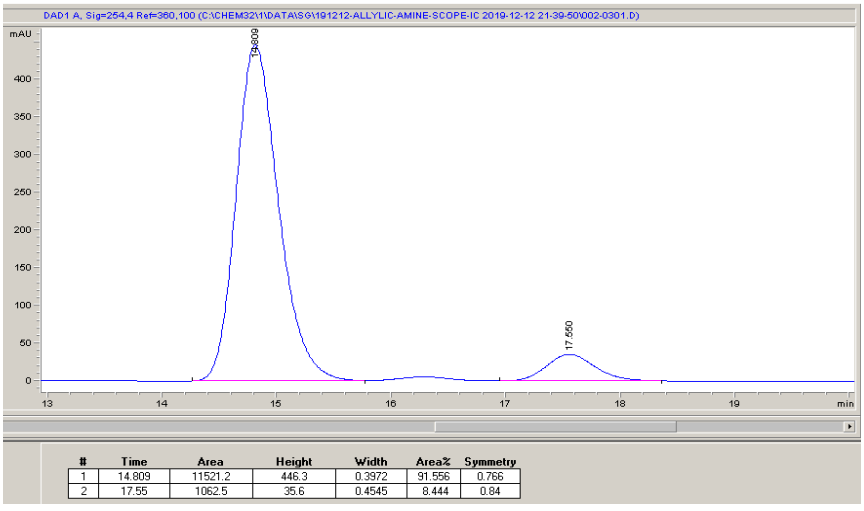
Conditions: Chiralpak IC, 10% *i*-PrOH in hexane, 1.5 mL/min, 254 nm

Racemic 6c:

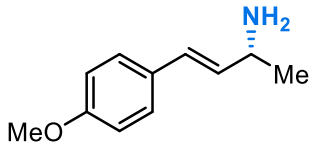


#	Time	Area	Height	Width	Area%	Symmetry
1	14.736	64162.1	2307.9	0.4363	49.444	0.651
2	17.398	65604.7	2020.9	0.5027	50.556	0.603

Enzymatic product 6c: 83% *ee*



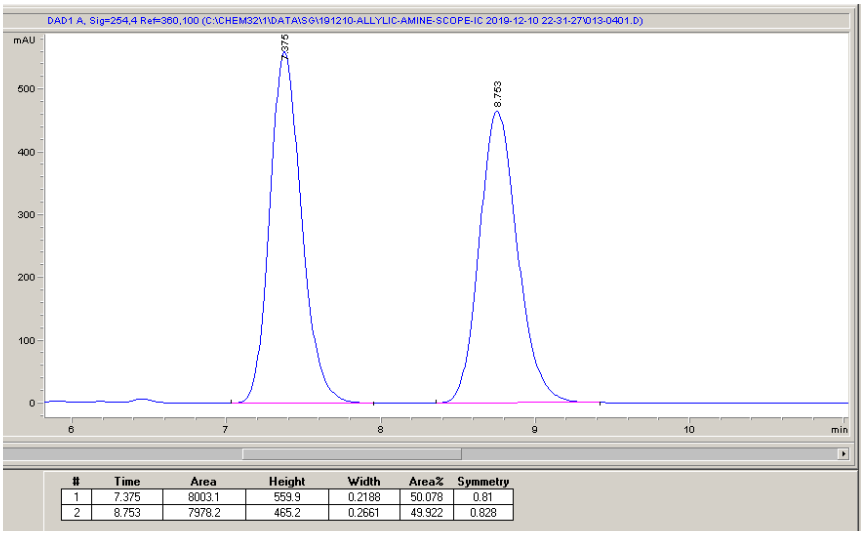
#	Time	Area	Height	Width	Area%	Symmetry
1	14.809	11521.2	446.3	0.3972	91.556	0.766
2	17.55	1062.5	35.6	0.4545	8.444	0.84



(*R*,*E*)-4-(4-methoxyphenyl)but-3-en-2-amine (6d)

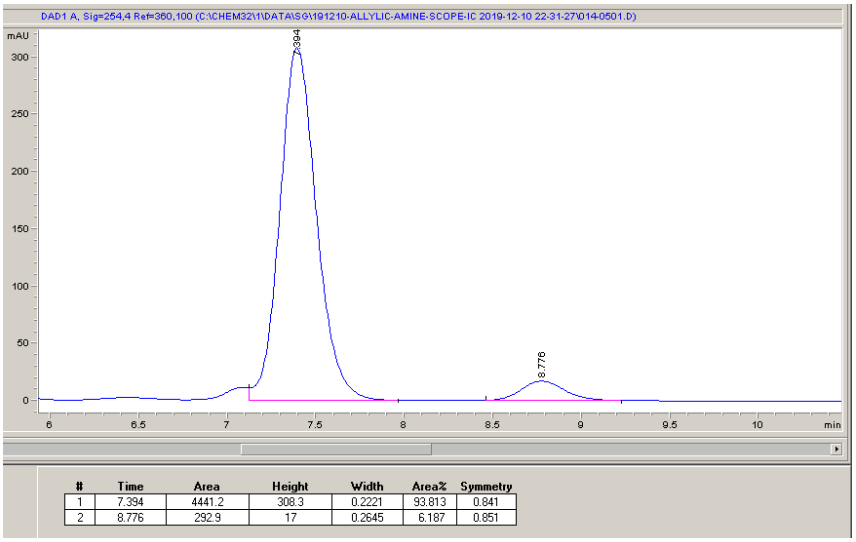
Conditions: Chiralpak IC, 25% *i*-PrOH in hexane, 1.5 mL/min, 254 nm

Racemic 6d:

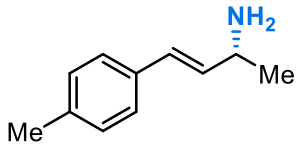


#	Time	Area	Height	Width	Area%	Symmetry
1	7.375	8003.1	559.9	0.2188	50.078	0.81
2	8.753	7978.2	465.2	0.2661	49.922	0.828

Enzymatic product 6d: 88% *ee*



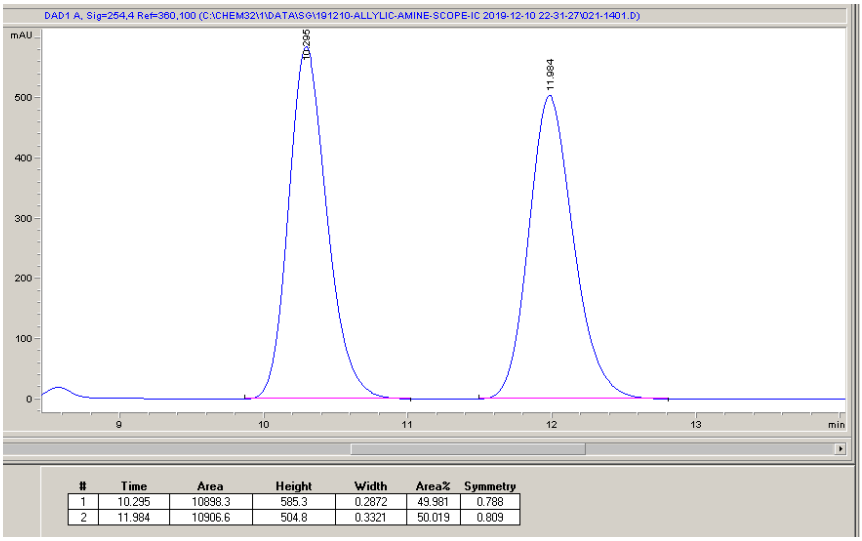
#	Time	Area	Height	Width	Area%	Symmetry
1	7.394	4441.2	308.3	0.2221	93.813	0.841
2	8.776	292.9	17	0.2645	6.187	0.851



(*R*,*E*)-4-(4-methylphenyl)but-3-en-2-amine (6e)

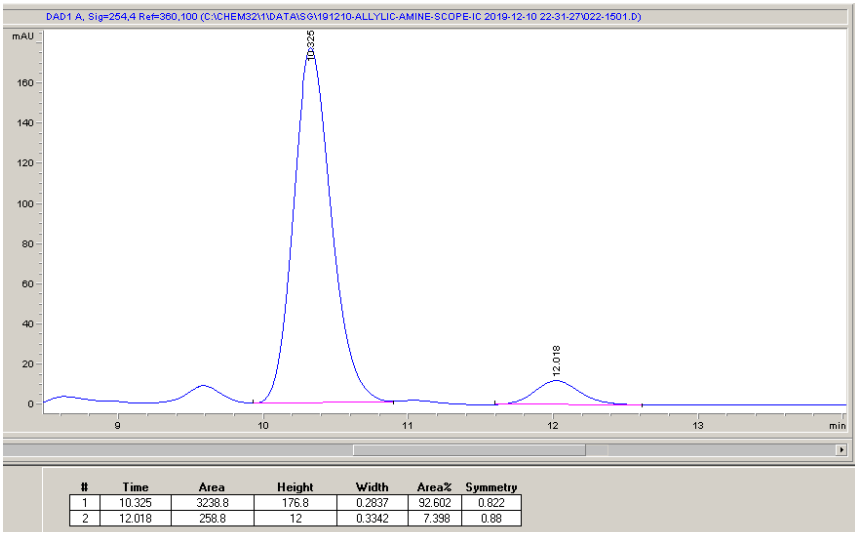
Conditions: Chiralpak IC, 15% *i*-PrOH in hexane, 1.5 mL/min, 254 nm

Racemic 6e:

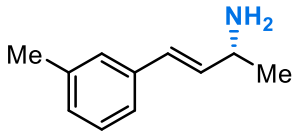


#	Time	Area	Height	Width	Area%	Symmetry
1	10.295	10898.3	585.3	0.2872	49.981	0.788
2	11.984	10906.6	504.8	0.3321	50.019	0.809

Enzymatic product 6e: 85% *ee*



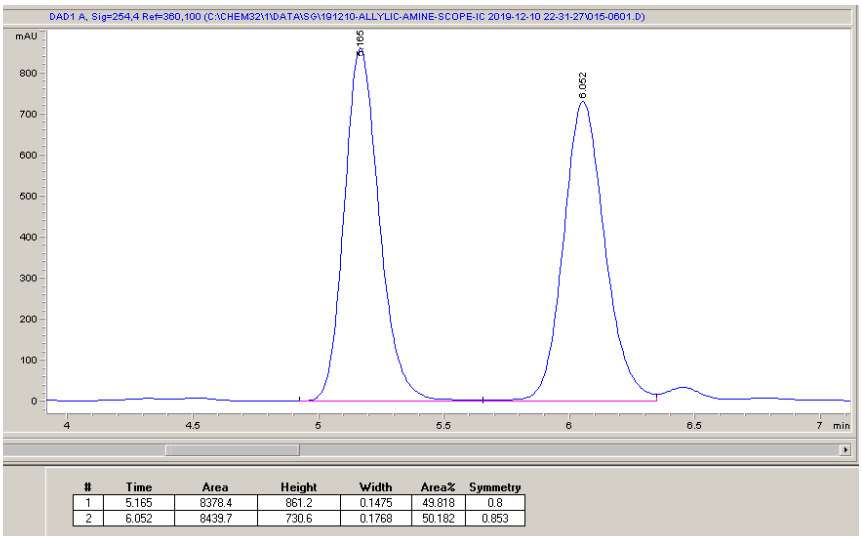
#	Time	Area	Height	Width	Area%	Symmetry
1	10.325	3238.8	176.8	0.2837	92.602	0.822
2	12.018	258.8	12	0.3342	7.398	0.88



(*R*,*E*)-4-(3-methylphenyl)but-3-en-2-amine (6f)

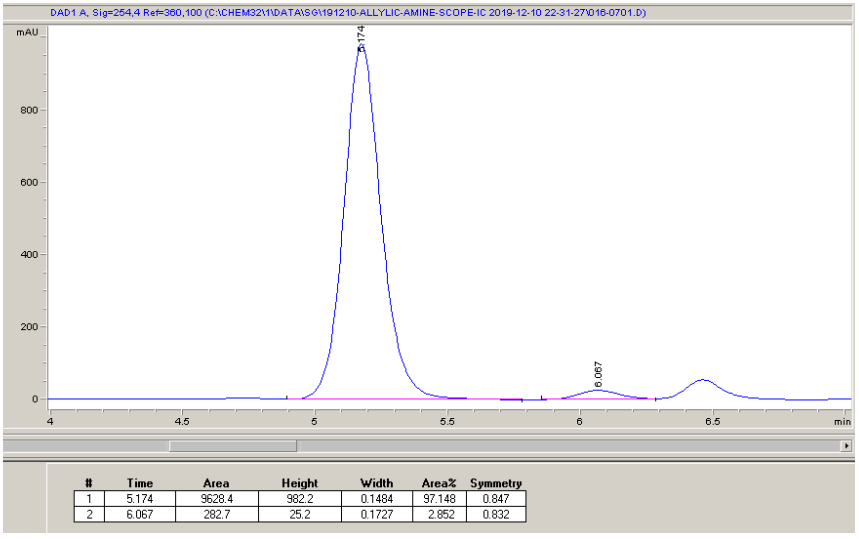
Conditions: Chiralpak IC, 25% *i*-PrOH in hexane, 1.5 mL/min, 254 nm

Racemic 6f:

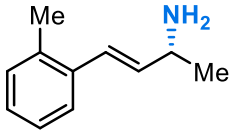


#	Time	Area	Height	Width	Area%	Symmetry
1	5.165	8378.4	861.2	0.1475	49.818	0.8
2	6.052	8439.7	730.6	0.1768	50.182	0.853

Enzymatic product 6f: 94% *ee*



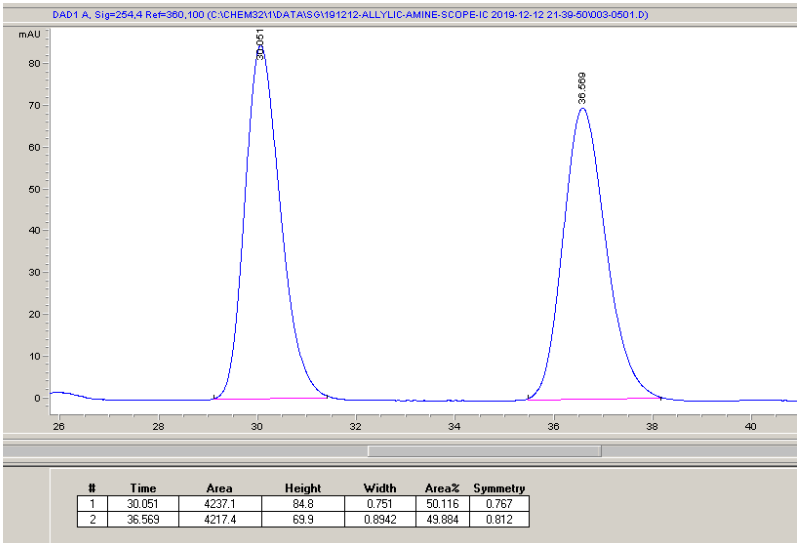
#	Time	Area	Height	Width	Area%	Symmetry
1	5.174	9628.4	982.2	0.1484	97.148	0.847
2	6.067	282.7	25.2	0.1727	2.852	0.832



(*R*,*E*)-4-(2-methylphenyl)but-3-en-2-amine (6g)

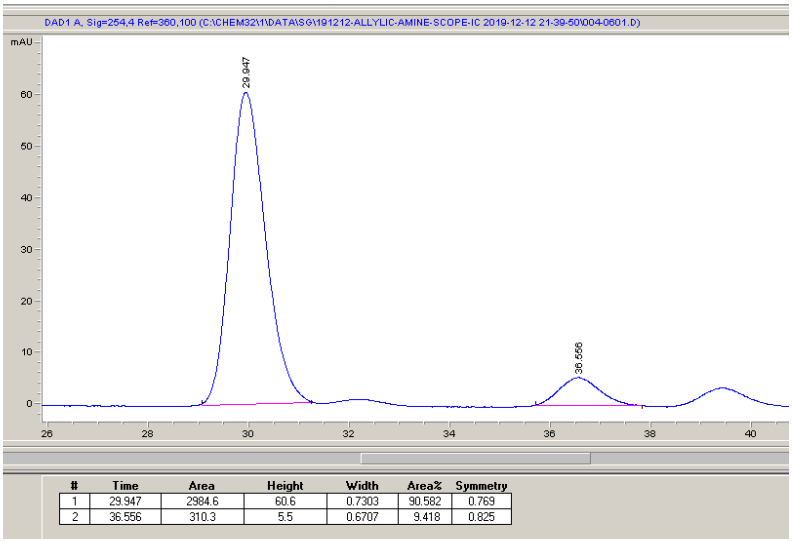
Conditions: Chiralpak IC, 5% *i*-PrOH in hexane, 1.5 mL/min, 254 nm

Racemic 6g:

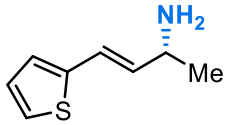


#	Time	Area	Height	Width	Area%	Symmetry
1	30.051	4237.1	84.8	0.751	50.116	0.767
2	36.569	4217.4	69.9	0.8942	49.884	0.812

Enzymatic product 6g: 81% *ee*



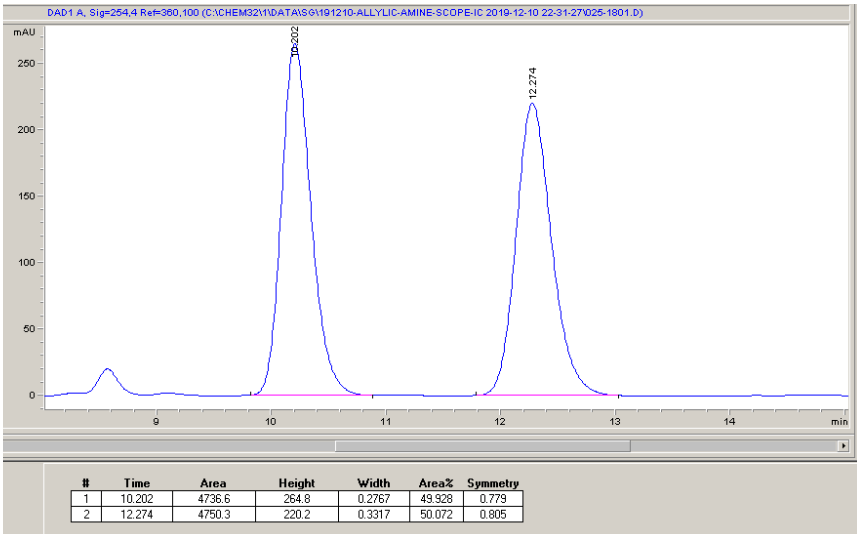
#	Time	Area	Height	Width	Area%	Symmetry
1	29.947	2984.6	60.6	0.7303	90.582	0.769
2	36.556	310.3	5.5	0.6707	9.418	0.825



(*R*,*E*)-4-(thiophen-2-yl)but-3-en-2-amine (6h)

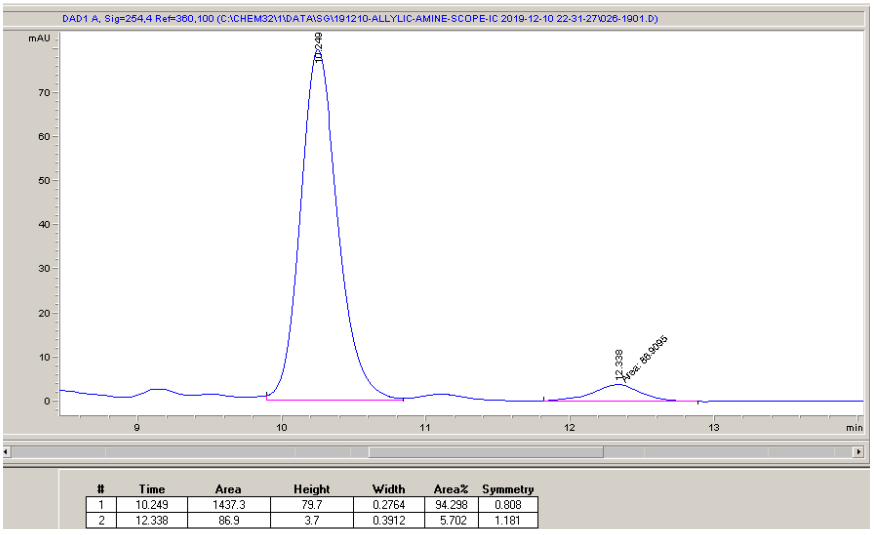
Conditions: Chiralpak IC, 15% *i*-PrOH in hexane, 1.5 mL/min, 254 nm

Racemic 6h:

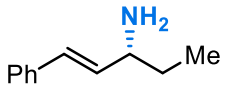


#	Time	Area	Height	Width	Area%	Symmetry
1	10.202	4736.6	264.8	0.2767	49.928	0.779
2	12.274	4750.3	220.2	0.3317	50.072	0.805

Enzymatic product 6h: 89% *ee*



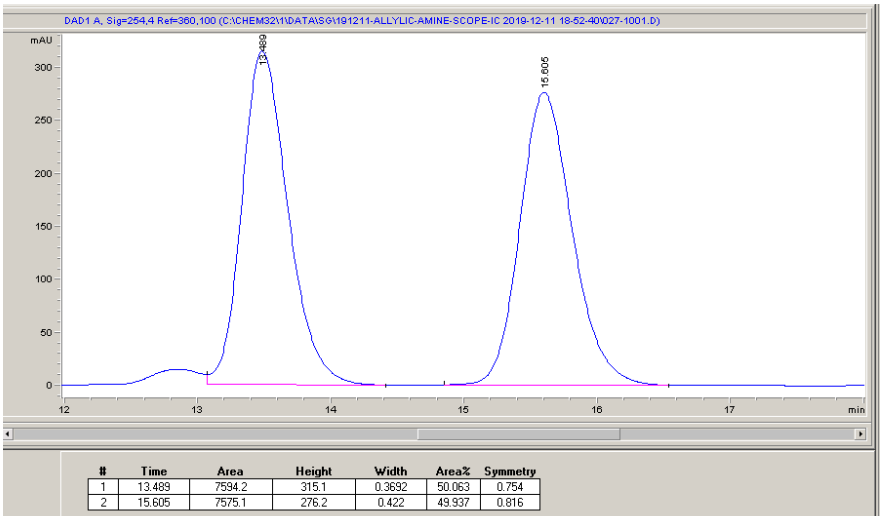
#	Time	Area	Height	Width	Area%	Symmetry
1	10.249	1437.3	79.7	0.2764	94.298	0.808
2	12.338	86.9	3.7	0.3912	5.702	1.181



(*R,E*)-1-phenylpent-1-en-3-amine (6i)

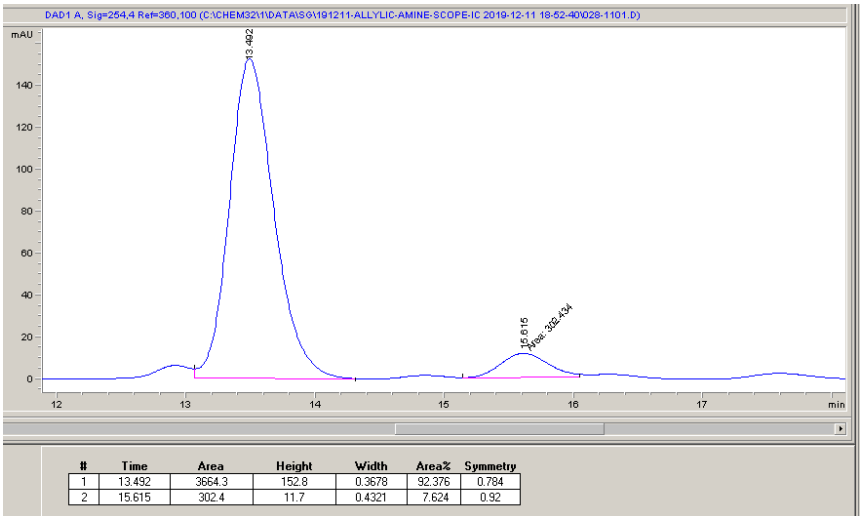
Conditions: Chiralpak IC, 10% *i*-PrOH in hexane, 1.5 mL/min, 254 nm

Racemic 6i:

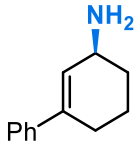


#	Time	Area	Height	Width	Area%	Symmetry
1	13.489	7594.2	315.1	0.3692	50.063	0.754
2	15.605	7575.1	276.2	0.422	49.937	0.816

Enzymatic product 6i: 85% *ee*



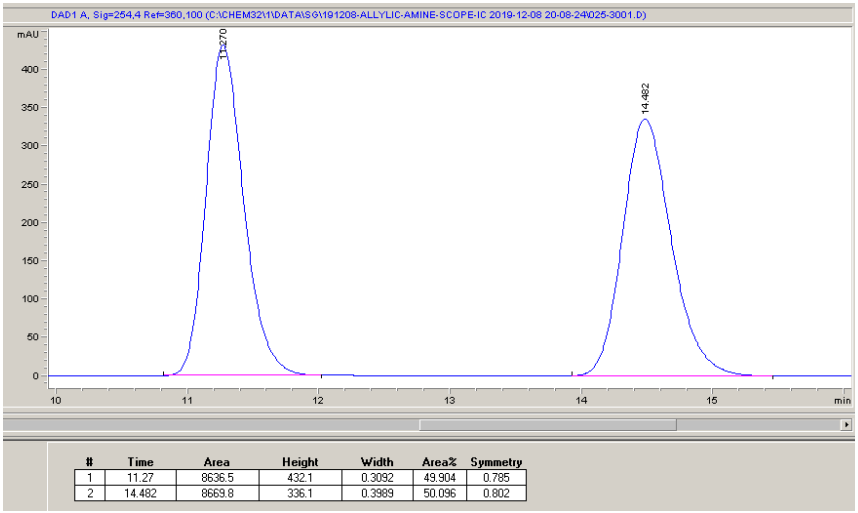
#	Time	Area	Height	Width	Area%	Symmetry
1	13.492	3664.3	152.8	0.3678	92.376	0.784
2	15.615	302.4	11.7	0.4321	7.624	0.92



(*S*)-3,4,5,6-tetrahydro-[1,1'-biphenyl]-3-amine (6j)

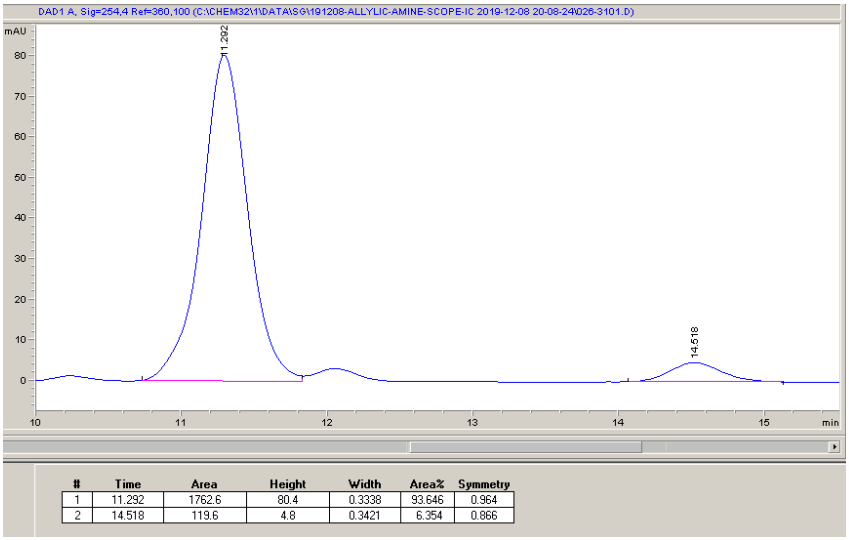
Conditions: Chiralpak IC, 15% *i*-PrOH in hexane, 1.5 mL/min, 254 nm

Racemic 6j:

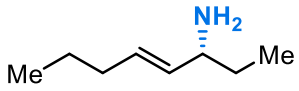


#	Time	Area	Height	Width	Area%	Symmetry
1	11.27	8636.5	432.1	0.3092	49.904	0.785
2	14.482	8669.8	336.1	0.3989	50.096	0.802

Enzymatic product 6j: 87% *ee*



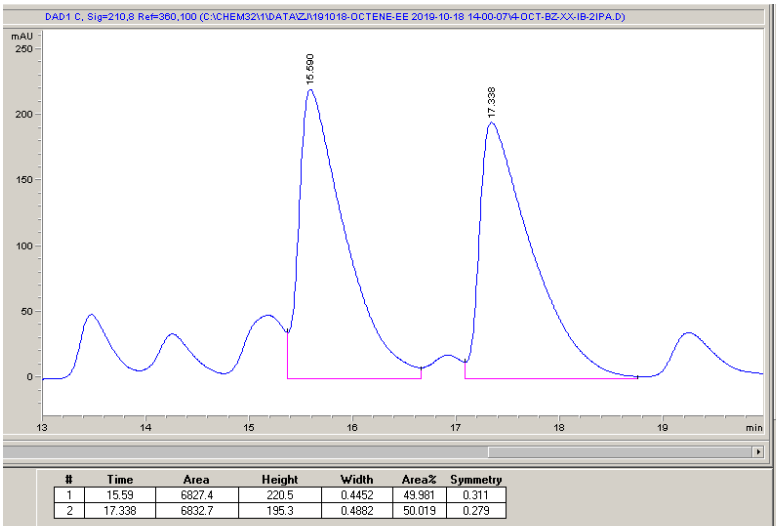
#	Time	Area	Height	Width	Area%	Symmetry
1	11.292	1762.6	80.4	0.3338	93.646	0.964
2	14.518	119.6	4.8	0.3421	6.354	0.866



(*R*,*E*)-oct-4-en-3-amine (6m)

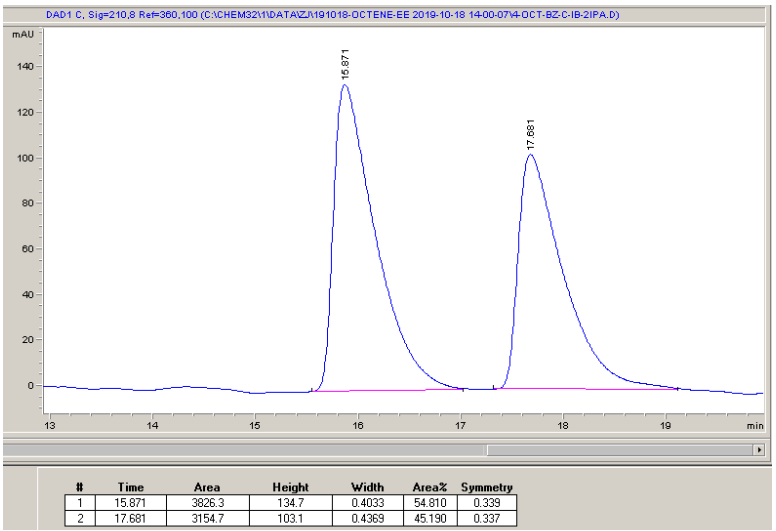
Conditions: Chiralpak IB, 2% *i*-PrOH in hexane, 1.5 mL/min, 210 nm

Racemic Bz-6m:

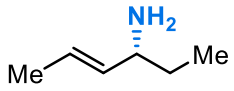


#	Time	Area	Height	Width	Area%	Symmetry
1	15.59	6827.4	220.5	0.4452	49.981	0.311
2	17.338	6832.7	195.3	0.4882	50.019	0.279

Enzymatic product Bz-6m: 10% *ee*



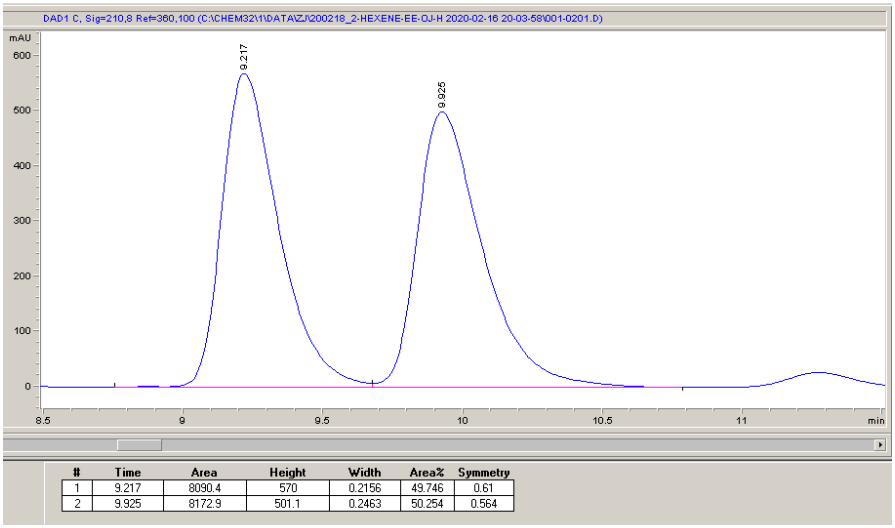
#	Time	Area	Height	Width	Area%	Symmetry
1	15.871	3826.3	134.7	0.4033	54.810	0.339
2	17.681	3154.7	103.1	0.4369	45.190	0.337



(*R,E*)-hex-4-en-3-amine (6n-major)

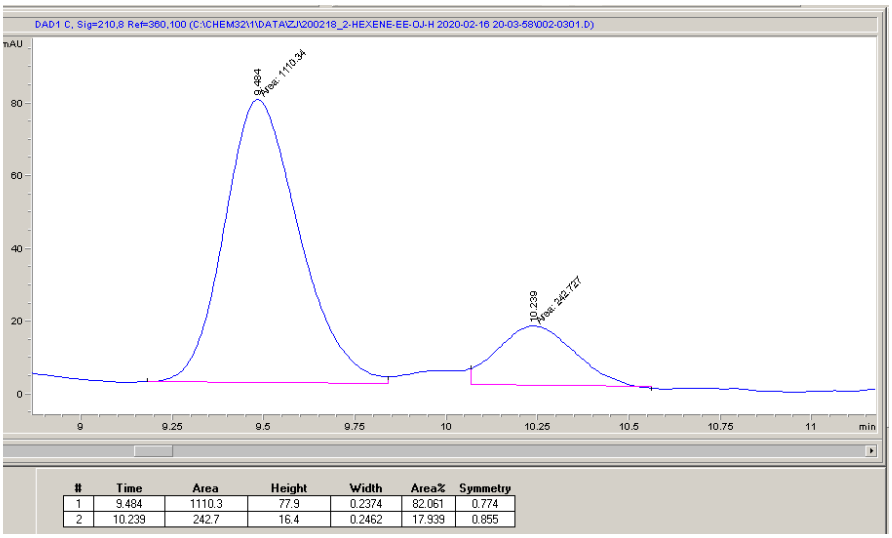
Conditions: Chiralcel OJ-H columns, 4% *i*-PrOH in hexane, 1.5 mL/min, 210 nm

Racemic 6n-major:



#	Time	Area	Height	Width	Area%	Symmetry
1	9.217	8090.4	570	0.2156	49.746	0.61
2	9.925	8172.9	501.1	0.2463	50.254	0.564

Enzymatic product 6n-major: 64% *ee*



#	Time	Area	Height	Width	Area%	Symmetry
1	9.484	1110.3	77.9	0.2374	82.061	0.774
2	10.239	242.7	16.4	0.2462	17.939	0.855

A.8. Sequence Information for Final Variants

Variant P411-B2

DNA sequence:

ATGACAATTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATTTACCGTTATTAAACACAGA
TAAACCGGTTCAAGCTTTGATGAAAATTGCGGATGAATTAGGAGAAATCTTTAAATTCGAGGCGCCTGGTC
GTGTAACGCGCTACTTATCAAGTCAGCGTCTAATTAAAGAAGCATGCGATGAATCACGCTTTGATAAAAAC
TTAAGTCAAGCGCTTAAATTTATGCGTGATTTTCTTGGAGACGGGTTAGCCACAAGCTGGACGCATGAAAA
AAATTGGAAAAAAGCGCATAATATCTTACTTCCAAGCTTTAGTCAGCAGGCAATGAAAGGCTATCATGCGA
TGATGGTCGATATCGCCGTGCAGCTTGTTCAAAAGTGGGAGCGTCTAAATGCAGATGAGCATATTGAAGTA
TCGGAAGACATGACACGTTTAACGCTTGATACAATTGGTCTTTGCGGCTTTAACTATCGCTTTAACAGCTT
TTACCGAGATCAGCCTCATCCATTTATTATAAGTATGGTCCGTGCACTGGATGAAGTAATGAACAAGCTGC
AGCGAGCAAATCCAGACGACCCAGCTTATGATGAAAACAAGCGCCAGTTTCAAGAAGATATCAAGGTGATG
AACGACCTAGTAGATAAAATTATTGCAGATCGCAAAGCAAGGGGTGAACAAAGCGATGATTTATTAACGCA
GATGCTAAACGGAAAAGATCCAGAAACGGGTGAGCCGCTTGATGACGGGAACATTCGCTATCAAATTATTA
CATTCTTAATGGCGGGACACGATCCGACAAGTGGTCTTTTATCATTTGCGCTGTATTTCTTAGTGAAAAAT
CCACATGTATTACAAAAAGTAGCAGAAGAAGCAGCACGAGTTCTAGTAGATCCTGTTCCAAGCTACAAACA
AGTCAAACAGCTTAAATATGTCGGCATGGTCTTAAACGAAGCGCTGCGCTGTTGGCCAGCGGTGCCTGCGT
TTTCCCTATATGCAAAAGAAGATACGGTGCTTGGAGGAGAATATCCTTTAGAAAAAGGCGACGAAGTAATG
GTTCTGATTCCTCAGCTTCACCGTGATAAAACAGTTTGGGGAGACGATGTGGAGGAGTTCCGTCCAGAGCG
TTTTGAAAATCCAAGTGCGATTCCGCAGCATGCGTTTAAACCGTTTGGAAATGGTCAGCGTGCGTCTATCG
GTCAGCAGTTCGCTCTTCATGAAGCAACGCTGGTACTTGGTATGATGCTAAAACACTTTGACTTTGAAGAT
CATACAAACTACGAGCTCGATATTAAAGAAACTTTTGGTTTAAAACCTAAAGGCTTTGTGGTAAAAGCAAA
ATCGAAAAAAATTCCGCTTGGCGGTATTCCTTCACCTAGCACTGAACAGTCTGCTAAAAAAGTACGCAAAA
AGGCAGAAAACGCTCATAATACGCCGCTGCTTGTGCTATACGGTTCAAATATGGGTACCGCTGAAGGAACG
GCGCGTGATTTAGCAGATATTGCAATGAGCAAAGGATTTGCACCGCAGGTCGCAACGCTTGATTCACACGC
CGGAAATCTTCCGCGCGAAGGAGCTGTATTAATTGTAACGGCGTCTTATAACGGTCATCCGCCTGATAACG
CAAAGCAATTTGTCGACTGGTTAGACCAAGCGTCTGCTGATGAAGTAAAAGGCGTTCGCTACTCCGTATTT
GGATGCGGCGATAAAAACTGGGCTACTACGTATCAAAAAGTGCCTGCTTTTATCGATGAAACGCTTGCCGC
TAAAGGGGCAGAAAACATCGCTGACCGCGGTGAAGCAGATGCAAGCGACGACTTTGAAGGCACATATGAAG
AATGGCGTGAACATATGTGGAGTGACGTAGCAGCCTACTTTAACCTCGACATTGAAAACAGTGAAGATAAT
AAATCTAC
TCTTTCACTTCAATTTGTCGACAGCGCCGCGGATATGCCGCTTGCGAAAATGCACGGTGCGTTTTCAACGA
ACGTCGTAGCAAGCAAAGAACTTCAACAGCCAGGCAGTGCACGAAGCACGCGACATCTTGAAATTGAACTT
CCAAAAGAAGCTTCTTATCAAGAAGGAGATCATTTAGGTGTTATTCCTCGCAACTATGAAGGAATAGTAAA
CCGTGTAACAGCAAGGTTCGGCCTAGATGCATCACAGCAAATCCGTCTGGAAGCAGAAGAAGAAAAATTAG
CTCATTTGCCACTCGCTAAAACAGTATCCGTAGAAGAGCTTCTGCAATACGTGGAGCTTCAAGATCCTGTT
ACGCGCACGCAGCTTCGCGCAATGGCTGCTAAAACGGTCTGCCCGCCGCATAAAGTAGAGCTTGAAGCCTT
GCTTGAAAAGCAAGCCTACAAAGAACAAGTGCTGGCAAAACGTTTAACAATGCTTGAACTGCTTGAAAAAT
ACCCGGCGTGTGAAATGAAATTCAGCGAATTTATCGCCCTTCTGCCAAGCATACGCCCGCGCTATTACTCG
ATTTCTTCATCACCTCGTGTCGATGAAAAACAAGCAAGCATCACGGTCAGCGTTGTCTCAGGAGAAGCGTG
GAGCGGATATGGAGAATATAAAGGAATTGCGTCGAACTATCTTGCCGAGCTGCAAGAAGGAGATACGATTA
CGTGCTTTATTTCCACACCGCAGTCAGAATTTACGCTGCCAAAAGACCCTGAAACGCCGCTTATCATGGTC
GGACCGGGAACAGGCGTCGCGCCGTTTAGAGGCTTTGTGCAGGCGCGCAAACAGCTAAAAGAACAA
GGACAGTCACTTGGAGAAGCACATTTATACTTCGGCTGCCGTTCACCTCATGAAGACTATCTGTATCAAGA
AGAGCTTGAAAACGCCCAAAGCGAAGGCATCATTACGCTTCATACCGCTTTTTCTCGCATGCCAAATCAGC
CGAAAACATACGTTCAGCACGTAATGGAACAAGACGGCAAGAAATTGATTGAACTTCTTGATCAAGGAGCG
CACTTCTATATTTGCGGAGACGGAAGCCAAATGGCACCTGCCGTTGAAGCAACGCTTATGAAAAGCTATGC
TGACGTTCACCAAGTGAGTGAAGCAGACGCTCGCTTATGGCTGCAGCAGCTAGAAGAAAAAGGCCGATACG
CAAAAGACGTGTGGGCTGGGCTCGAGCACC ACCACCACCACCACTGA

Variant P411$_{BPA}$

DNA sequence:

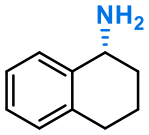
ATGACAATTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATTTACCGTTATTAAACACAGA
TAAACCGGTTCAAGCTTTGATGAAAATTGCGGATGAATTAGGAGAAATCTTTAAATTCGAGGCGCCTGGTC
GTGTAACGCGCTACTTATCAAGTCAGCGTCTAATTAAAGAAGCATGCGATGAATCACGCTTTGATAAAAAC
TTAAGTCAAGCGCTTAAATTTATGCGTGATTTTCTTGGAGACGGGTTAGCCACAAGCTGGACGCATGAAAA
AAATTGGAAAAAAGCGCATAATATCTTACTTCCAAGCTTTAGTCAGCAGGCAATGAAAGGCTATCATGCGA
TGATGGTCGATATCGCCGTGCAGCTTGTTCAAAAGTGGGAGCGTCTAAATGCAGATGAGCATATTGAAGTA
TCGGAAGACATGACACGTTTAACGCTTGATACAATTGGTCTTTGCGGCTTTAACTATCGCTTTAACAGCTT
TTACCGAGATCAGCCTCATCCATTTATTATAAGTATGGTCCGTGCACTGGATGAAGTAATGAACAAGCTGC
AGCGAGCAAATCCAGACGACCCAGCTTATGATGAAAACAAGCGCCAGTTTCAAGAAGATATCAAGGTGATG
AACGACCTAGTAGATAAAATTATTGCAGATCGCAAAGCAAGGGGTGAACAAAGCGATGATTTATTAACGCA
GATGCTAAACGGAAAAGATCCAGAAACGGGTGAGCCGCTTGATGACGGGAACATTCGCTATCAAATTATTA
CATTCTTAATGGCGGGACACGATCCGACAAGTGGTCTTTTATCATTTGCGCTGTATTTCTTAGTGAAAAAT
CCACATGTATTACAAAAAGTAGCAGAAGAAGCAGCACGAGTTCTAGTAGATCCTGTTCCAAGCTACAAACA
AGTCAAACAGCTTAAATATGTCGGCATGGTCTTAAACGAAGCGCTGCGCTGTTGGCCAGCGGTGCCTGCGT
TTTCCCTATATGCAAAAGAAGATACGGTGCTTGGAGGAGAATATCCTTTAGAAAAAGGCGACGAAGTAATG
GTTCTGATTCCTCAGCTTCACCGTGATAAAACAGTTTGGGGAGACGATGTGGAGGAGTTCCGTCCAGAGCG
TTTTGAAAATCCAAGTGCGATTCCGCAGCATGCGTTTAAACCGTTTGGAAATGGTCAGCGTGCGTCTATCG
GTCAGCAGTTCGCTCTTCATGAAGCAACGCTGGTACTTGGTATGATGCTAAAACACTTTGACTTTGAAGAT
CATACAAACTACGAGCTCGATATTAAAGAAACTTTTGGTTTAAAACCTAAAGGCTTTGTGGTAAAAGCAAA
ATCGAAAAAAATTCCGCTTGGCGGTATTCCTTCACCTAGCACTGAACAGTCTGCTAAAAAAGTACGCAAAA
AGGCAGAAAACGCTCATAATACGCCGCTGCTTGTGCTATACGGTTCAAATATGGGTACCGCTGAAGGAACG
GCGCGTGATTTAGCAGATATTGCAATGAGCAAAGGATTTGCACCGCAGGTCGCAACGCTTGATTCACACGC
CGGAAATCTTCCGCGCGAAGGAGCTGTATTAATTGTAACGGCGTCTTATAACGGTCATCCGCCTGATAACG
CAAAGCAATTTGTCGACTGGTTAGACCAAGCGTCTGCTGATGAAGTAAAAGGCGTTCGCTACTCCGTATTT
GGATGCGGCGATAAAAACTGGGCTACTACGTATCAAAAAGTGCCTGCTTTTATCGATGAAACGCTTGCCGC
TAAAGGGGCAGAAAACATCGCTGACCGCGGTGAAGCAGATGCAAGCGACGACTTTGAAGGCACATATGAAG
AATGGCGTGAACATATGTGGAGTGACGTAGCAGCCTACTTTAACCTCGACATTGAAAACAGTGAAGATAAT
AAATCTACTCTTTCACTTCAATTTGTCGACAGCGCCGCGGATATGCCGCTTGCGAAAATGCACGGTGCGTT
TTCAACGAACGTCGTAGCAAGCAAAGAACTTCAACAGCCAGGCAGTGCACGAAGCACGCGACATCTTGAAA
TTGAACTTCCAAAAGAAGCTTCTTATCAAGAAGGAGATCATTTAGGTGTTATTCCTCGCAACTATGAAGGA
ATAGTAAACCGTGTAACAGCAAGGTTCGGCCTAGATGCATCACAGCAAATCCGTCTGGAAGCAGAAGAAGA
AAAATTAGCTCATTTGCCACTCGCTAAAACAGTATCCGTAGAAGAGCTTCTGCAATACGTGGAGCTTCAAG
ATCCTGTTACGCGCACGCAGCTTCGCGCAATGGCTGCTAAAACGGTCTGCCCGCCGCATAAAGTAGAGCTT
GAAGCCTTGCTTGAAAAGCAAGCCTACAAAGAACAAGTGCTGGCAAAACGTTTAACAATGCTTGAACTGCT
TGAAAAATACCCGGCGTGTGAAATGAAATTCAGCGAATTTATCGCCCTTCTGCCAAGCATACGCCCGCGCT
ATTACTCGATTTCTTCATCACCTCGTGTCGATGAAAAACAAGCAAGCATCACGGTCAGCGTTGTCTCAGGA
GAAGCGTGGAGCGGATATGGAGAATATAAAGGAATTGCGTCGAACTATCTTGCCGAGCTGCAAGAAGGAGA
TACGATTACGTGCTTTATTTCCACACCGCAGTCAGAATTTACGCTGCCAAAAGACCCTGAAACGCCGCTTA
TCATGGTCGGACCGGGAACAGGCGTCGCGCCGTTTAGAGGCTTTGTGCAGGCGCGCAAACAGCTAAAAGAA
CAAGGACAGTCACTTGGAGAAGCACATTTATACTTCGGCTGCCGTTCACCTCATGAAGACTATCTGTATCA
AGAAGAGCTTGAAAACGCCCAAAGCGAAGGCATCATTACGCTTCATACCGCTTTTTCTCGCATGCCAAATC
AGCCGAAAACATACGTTCAGCACGTAATGGAACAAGACGGCAAGAAATTGATTGAACTTCTTGATCAAGGA
GCGCACTTCTATATTTGCGGAGACGGAAGCCAAATGGCACCTGCCGTTGAAGCAACGCTTATGAAAAGCTA
TGCTGACGTTCACCAAGTGAGTGAAGCAGACGCTCGCTTATGGCTGCAGCAGCTAGAAGAAAAAGGCCGAT
ACGCAAAAGACGTGTGGGCTGGGCTCGAGCACCACCACCACCACCACTGA

Variant P411$_{APA}$

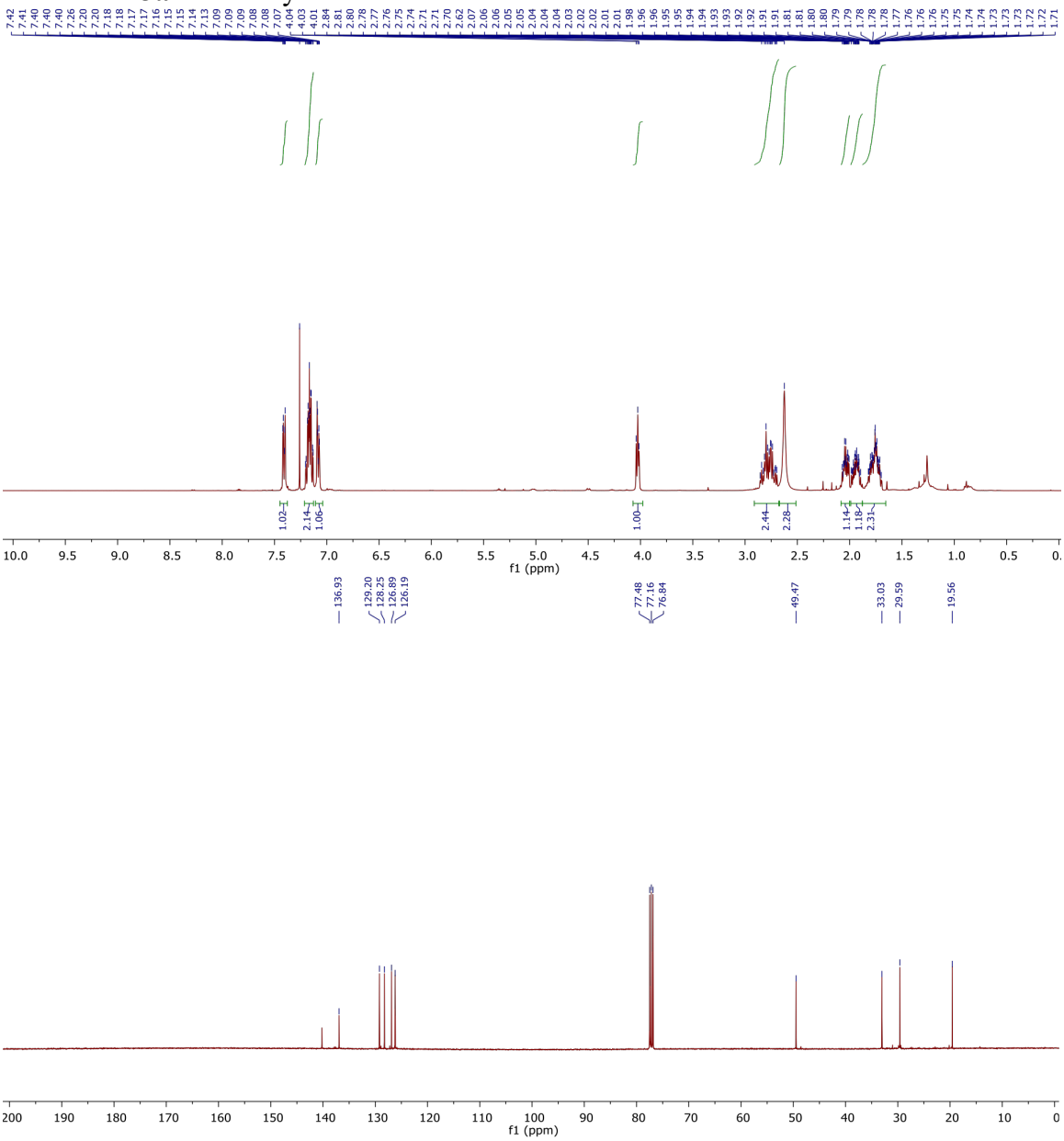
DNA sequence:

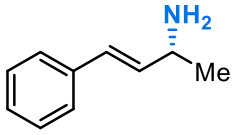
ATGACAATTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATTTACCGTTATTAAACACAGA
TAAACCGGTTCAAGCTTTGATGAAAATTGCGGATGAATTAGGAGAAATCTTTAAATTCGAGGCGCCTGGTC
GTGTAACGCGCTACTTATCAAGTCAGCGTCTAATTAAAGAAGCATGCGATGAATCACGCTTTGATAAAAAC
TTAAGTCAAGCGCTTAAATTTATGCGTGATTTTCTTGGAGACGGGTTAGCCACAAGCTGGACGCATGAAAA
AAATTGGAAAAAAGCGCATAATATCTTACTTCCAAGCTTTAGTCAGCAGGCAATGAAAGGCTATCATGCGA
TGATGGTCGATATCGCCGTGCAGCTTGTTCAAAAGTGGGAGCGTCTAAATGCAGATGAGCATATTGAAGTA
TCGGAAGACATGACACGTTTAACGCTTGATACAATTGGTCTTTGCGGCTTTAACTATCGCTTTAACAGCTT
TTACCGAGATCAGCCTCATCCATTTATTATAAGTATGGTCCGTGCACTGGATGAAGTAATGAACAAGCTGC
AGCGAGCAAATCCAGACGACCCAGCTTATGATGAAAACAAGCGCCAGTTTCAAGAAGATATCAAGGTGATG
AACGACCTAGTAGATAAAATTATTGCAGATCGCAAAGCAAGGGGTGAACAAAGCGATGATTTATTAACGCA
GATGCTAAACGGAAAAGATCCAGAAACGGGTGAGCCGCTTGATGACGGGAACATTCGCTATCAAATTATTA
CATTCTTAATGGCGGGACACGATCCGACAAGTGGTCTTTTATCATTTGCGCTGTATTTCTTAGTGAAAAAT
CCACATGTATTACAAAAAGTAGCAGAAGAAGCAGCACGAGTTCTAGTAGATCCTGTTCCAAGCTACAAACA
AGTCAAACAGCTTAAATATGTCGGCATGGTCTTAAACGAAGCGCTGCGCTTATGGCCAGCGGTGCCTGCGT
TTTCCCTATATGCAAAAGAAGATACGGTGCTTGGAGGAGAATATCCTTTAGAAAAAGGCGACGAAGTAATG
GTTCTGATTCCTCAGCTTCACCGTGATAAAACAGTTTGGGGAGACGATGTGGAGGAGTTCCGTCCAGAGCG
TTTTGAAAATCCAAGTGCGATTCCGCAGCATGCGTTTAAACCGTTTGGACGTGGTCAGCGTGCGTCTATCG
GTCAGCAGTTCGCTCTTCATGAAGCAACGCTGGTACTTGGTATGATGCTAAAACACTTTGACTTTGAAGAT
CATACAAACTACGAGCTCGATATTAAAGAAACTTTTGCGTTAAAACCTAAAGGCTTTGTGGTAAAAGCAAA
ATCGAAAAAAATTCCGCTTGGCGGTATTCCTTCACCTAGCACTGAACAGTCTGCTAAAAAAGTACGCAAAA
AGGCAGAAAACGCTCATAATACGCCGCTGCTTGTGCTATACGGTTCAAATATGGGTACCGCTGAAGGAACG
GCGCGTGATTTAGCAGATATTGCAATGAGCAAAGGATTTGCACCGCAGGTCGCAACGCTTGATTCACACGC
CGGAAATCTTCCGCGCGAAGGAGCTGTATTAATTGTAACGGCGTCTTATAACGGTCATCCGCCTGATAACG
CAAAGCAATTTGTCGACTGGTTAGACCAAGCGTCTGCTGATGAAGTAAAAGGCGTTCGCTACTCCGTATTT
GGATGCGGCGATAAAAACTGGGCTACTACGTATCAAAAAGTGCCTGCTTTTATCGATGAAACGCTTGCCGC
TAAAGGGGCAGAAAACATCGCTGACCGCGGTGAAGCAGATGCAAGCGACGACTTTGAAGGCACATATGAAG
AATGGCGTGAACATATGTGGAGTGACGTAGCAGCCTACTTTAACCTCGACATTGAAAACAGTGAAGATAAT
AAATCTACTCTTTCACTTCAATTTGTCGACAGCGCCGCGGATATGCCGCTTGCGAAAATGCACGGTGCGTT
TTCAACGAACGTCGTAGCAAGCAAAGAACTTCAACAGCCAGGCAGTGCACGAAGCACGCGACATCTTGAAA
TTGAACTTCCAAAAGAAGCTTCTTATCAAGAAGGAGATCATTTAGGTGTTATTCCTCGCAACTATGAAGGA
ATAGTAAACCGTGTAACAGCAAGGTTCGGCCTAGATGCATCACAGCAAATCCGTCTGGAAGCAGAAGAAGA
AAAATTAGCTCATTTGCCACTCGCTAAAACAGTATCCGTAGAAGAGCTTCTGCAATACGTGGAGCTTCAAG
ATCCTGTTACGCGCACGCAGCTTCGCGCAATGGCTGCTAAAACGGTCTGCCCGCCGCATAAAGTAGAGCTT
GAAGCCTTGCTTGAAAAGCAAGCCTACAAAGAACAAGTGCTGGCAAAACGTTTAACAATGCTTGAACTGCT
TGAAAAATACCCGGCGTGTGAAATGAAATTCAGCGAATTTATCGCCCTTCTGCCAAGCATACGCCCGCGCT
ATTACTCGATTTCTTCATCACCTCGTGTCGATGAAAAACAAGCAAGCATCACGGTCAGCGTTGTCTCAGGA
GAAGCGTGGAGCGGATATGGAGAATATAAAGGAATTGCGTCGAACTATCTTGCCGAGCTGCAAGAAGGAGA
TACGATTACGTGCTTTATTTCCACACCGCAGTCAGAATTTACGCTGCCAAAAGACCCTGAAACGCCGCTTA
TCATGGTCGGACCGGGAACAGGCGTCGCGCCGTTTAGAGGCTTTGTGCAGGCGCGCAAACAGCTAAAAGAA
CAAGGACAGTCACTTGGAGAAGCACATTTATACTTCGGCTGCCGTTCACCTCATGAAGACTATCTGTATCA
AGAAGAGCTTGAAAACGCCCAAAGCGAAGGCATCATTACGCTTCATACCGCTTTTTCTCGCATGCCAAATC
AGCCGAAAACATACGTTCAGCACGTAATGGAACAAGACGGCAAGAAATTGATTGAACTTCTTGATCAAGGA
GCGCACTTCTATATTTGCGGAGACGGAAGCCAAATGGCACCTGCCGTTGAAGCAACGCTTATGAAAAGCTA
TGCTGACGTTCACCAAGTGAGTGAAGCAGACGCTCGCTTATGGCTGCAGCAGCTAGAAGAAAAAGGCCGAT
ACGCAAAAGACGTGTGGGCTGGGCTCGAGCACCACCACCACCACCACTGA

A.9. NMR spectra

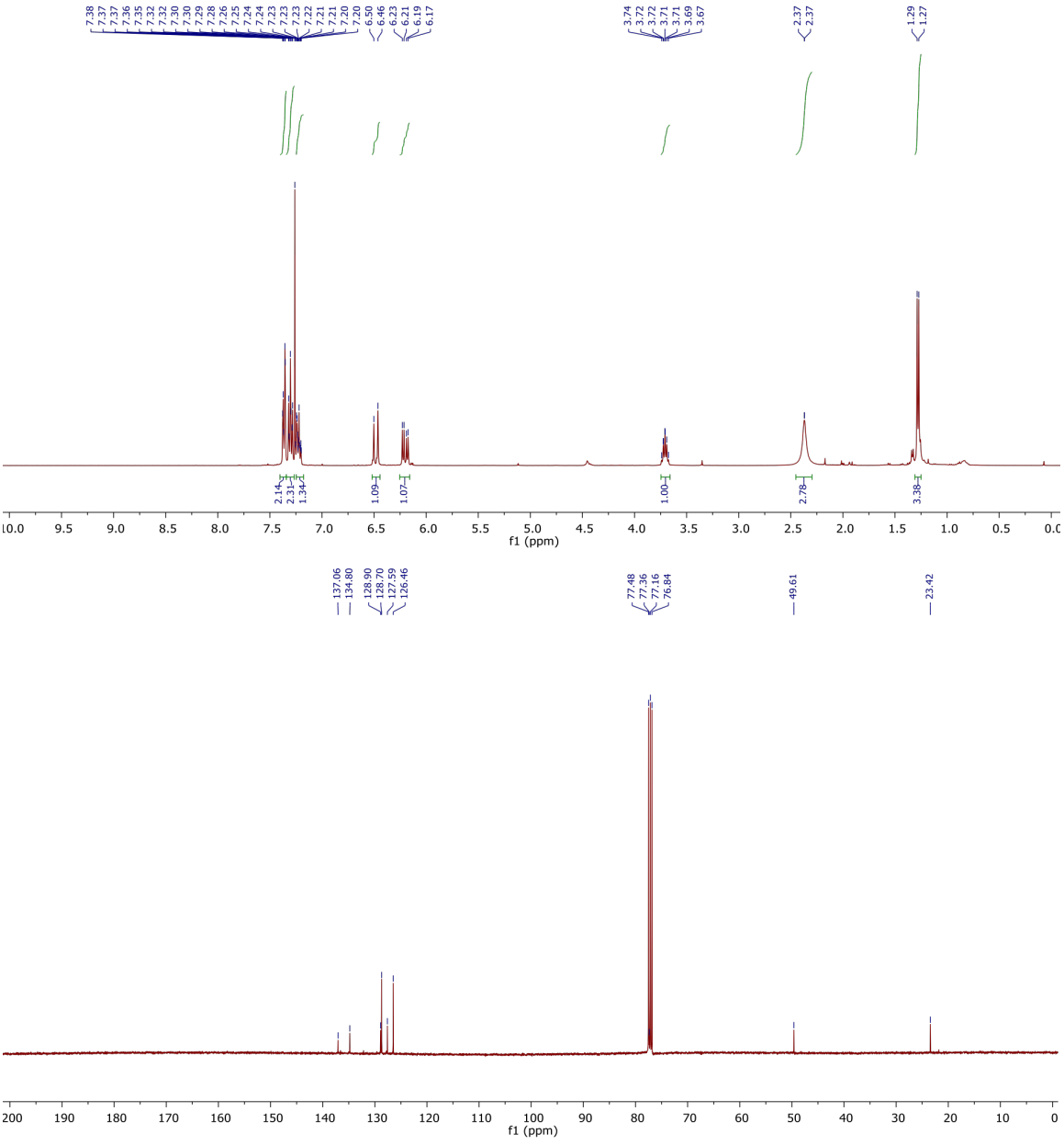


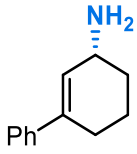
3a from enzymatic reaction



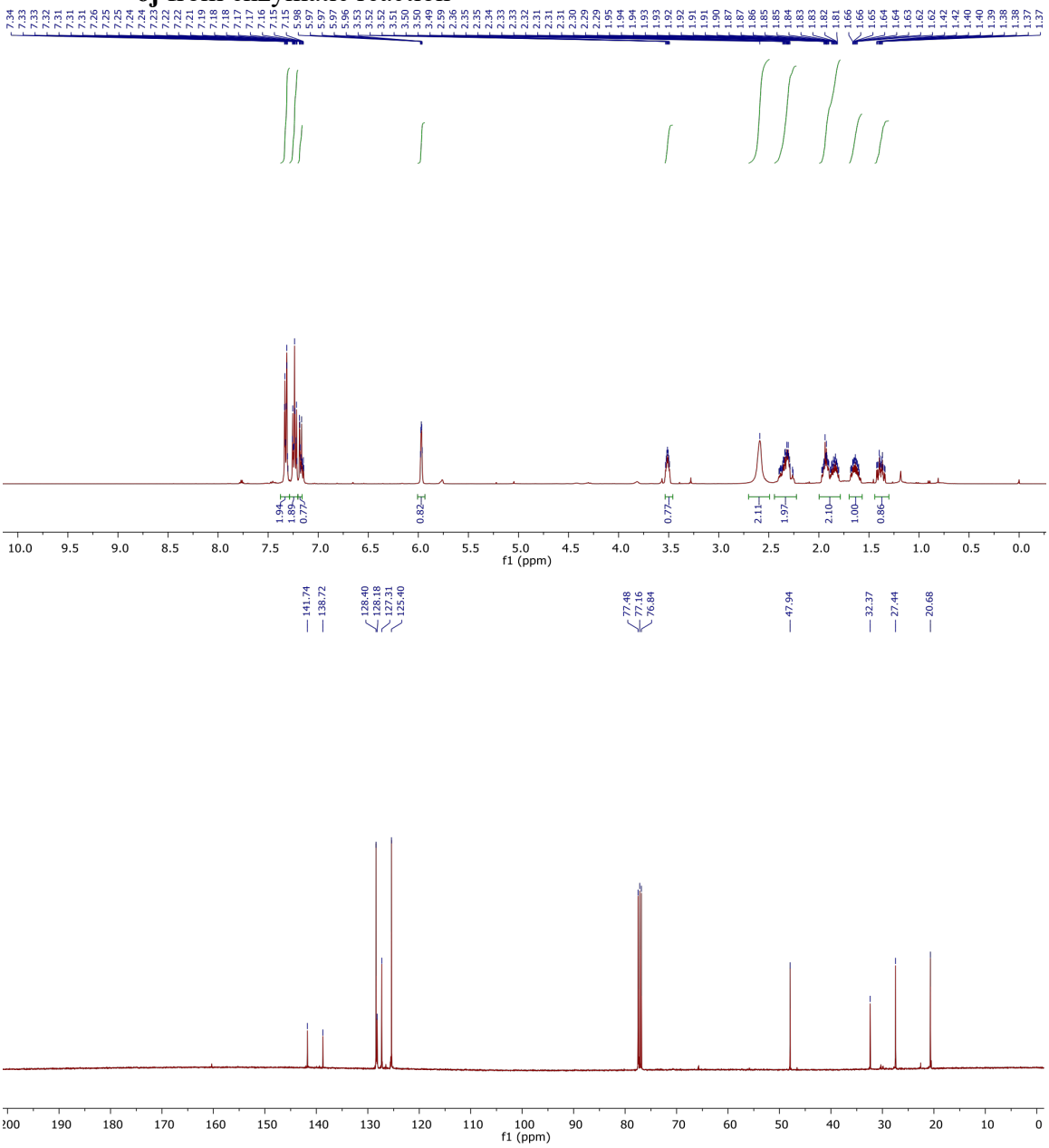


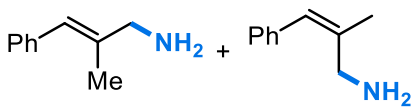
6a from enzymatic reaction



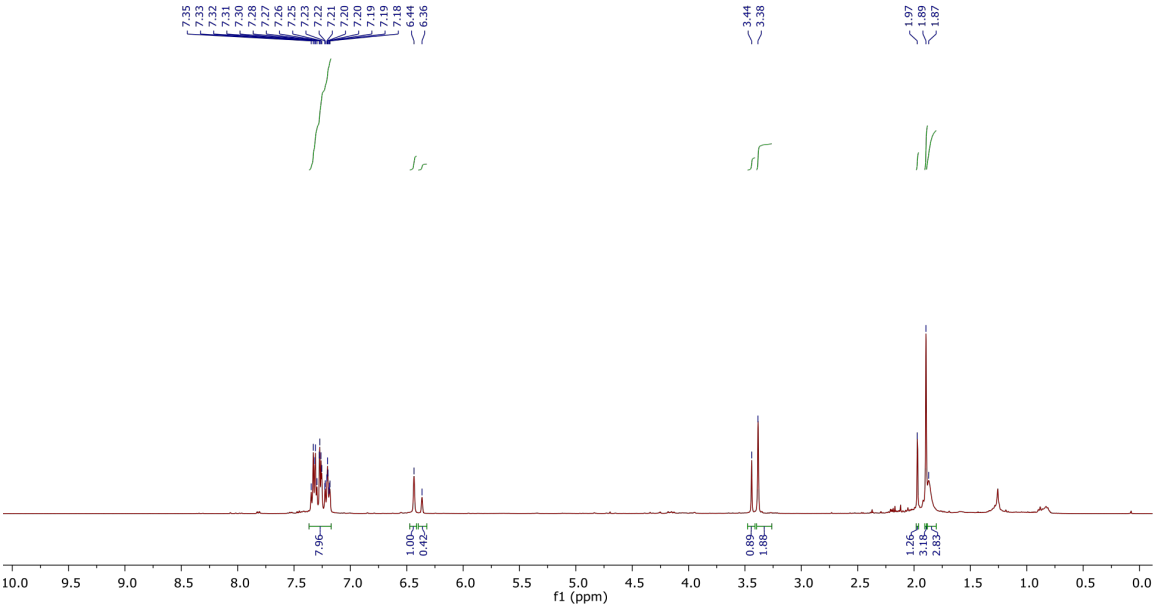


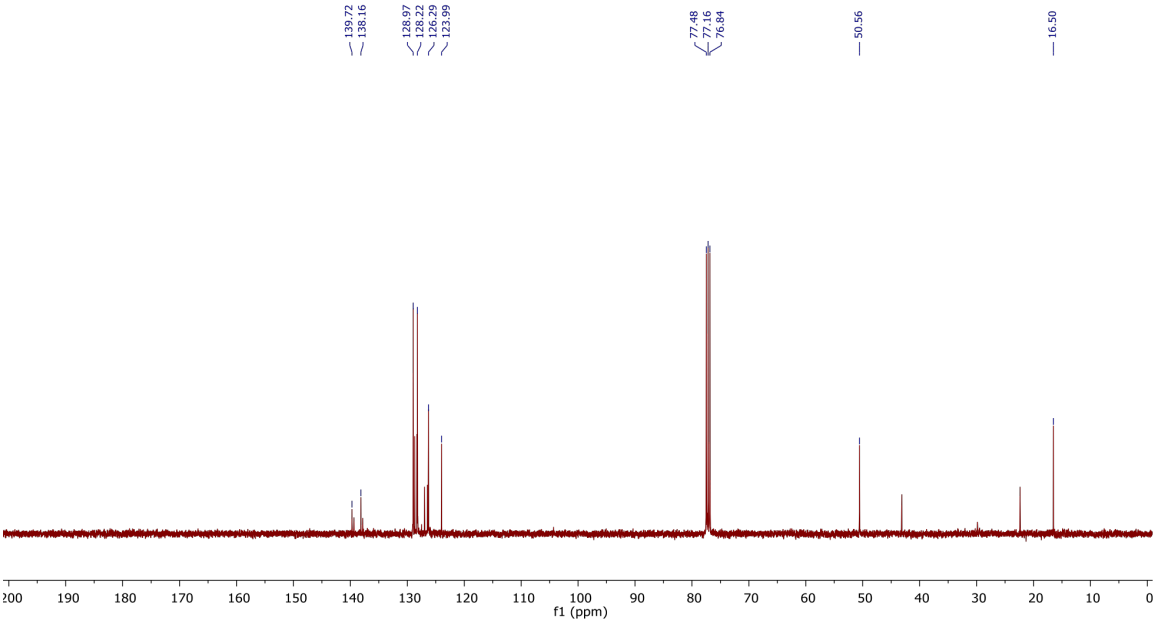
6j from enzymatic reaction

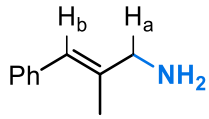


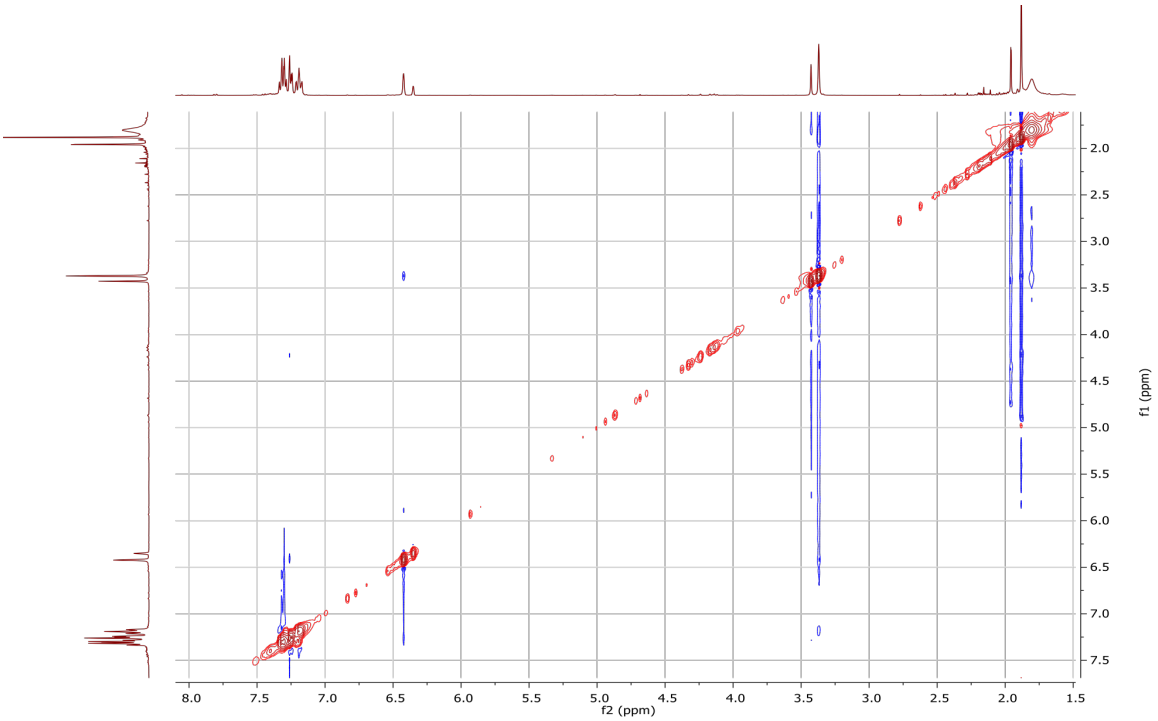


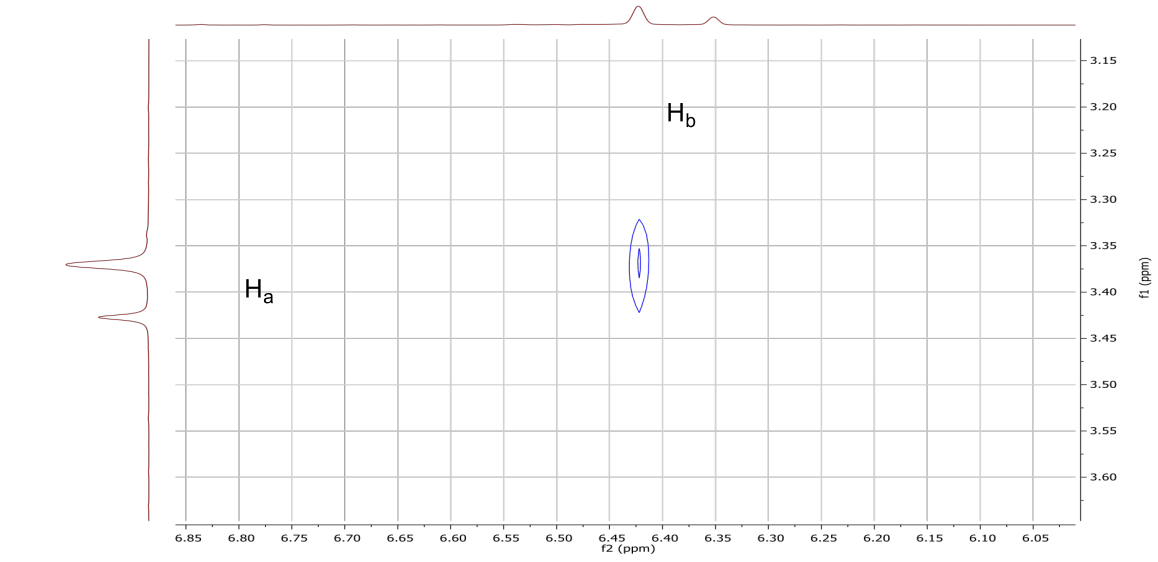
6k from enzymatic reaction

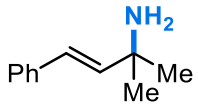




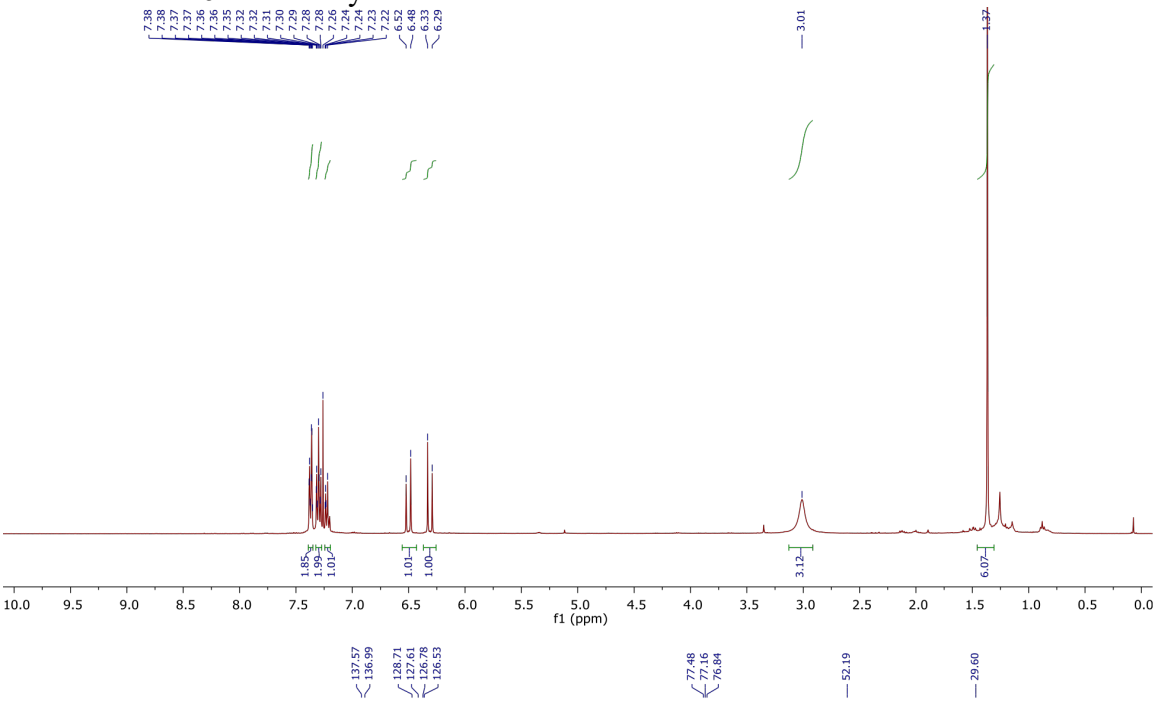

Hb
Ha
Ph
NH2

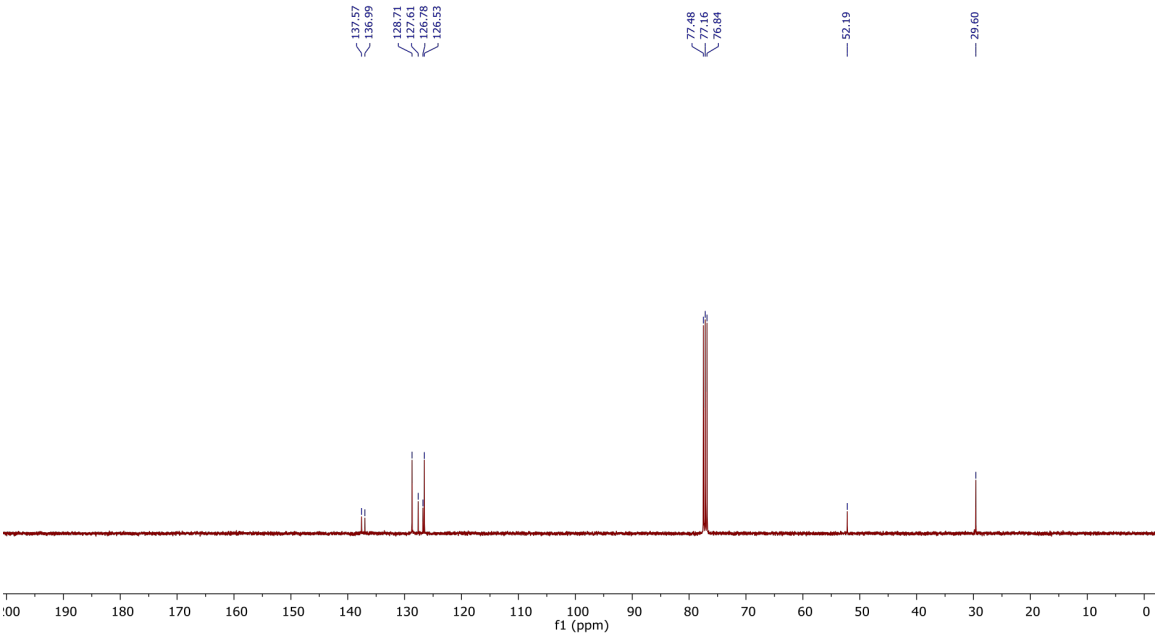

f2 (ppm)
f1 (ppm)

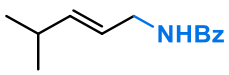

Hb
Ha
f2 (ppm)
f1 (ppm)



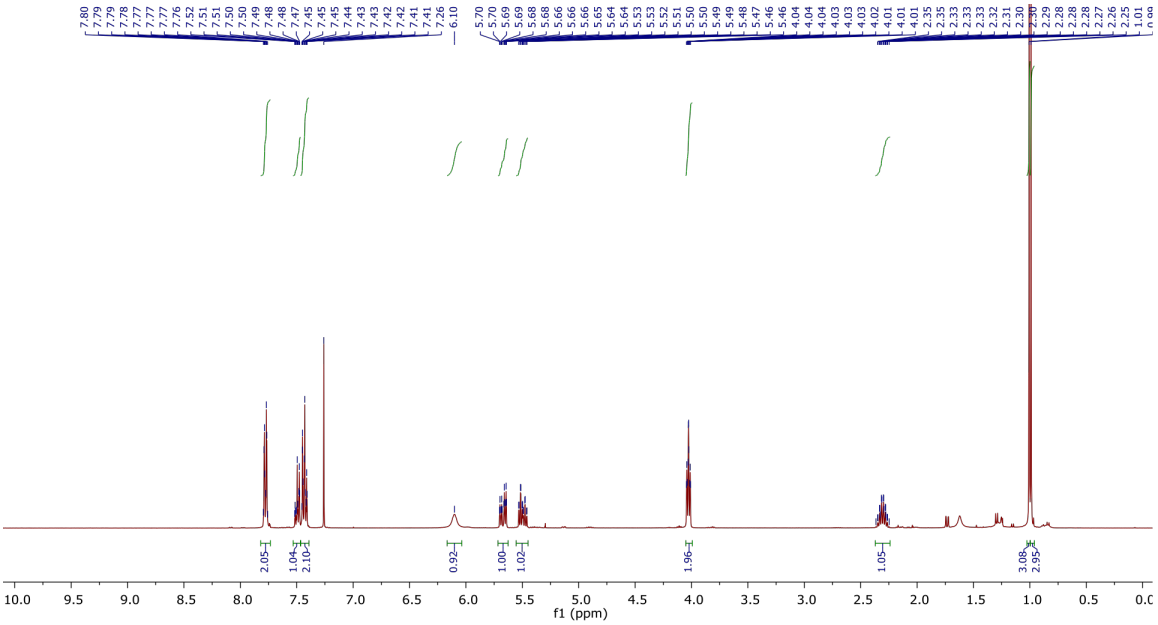
6l from enzymatic reaction

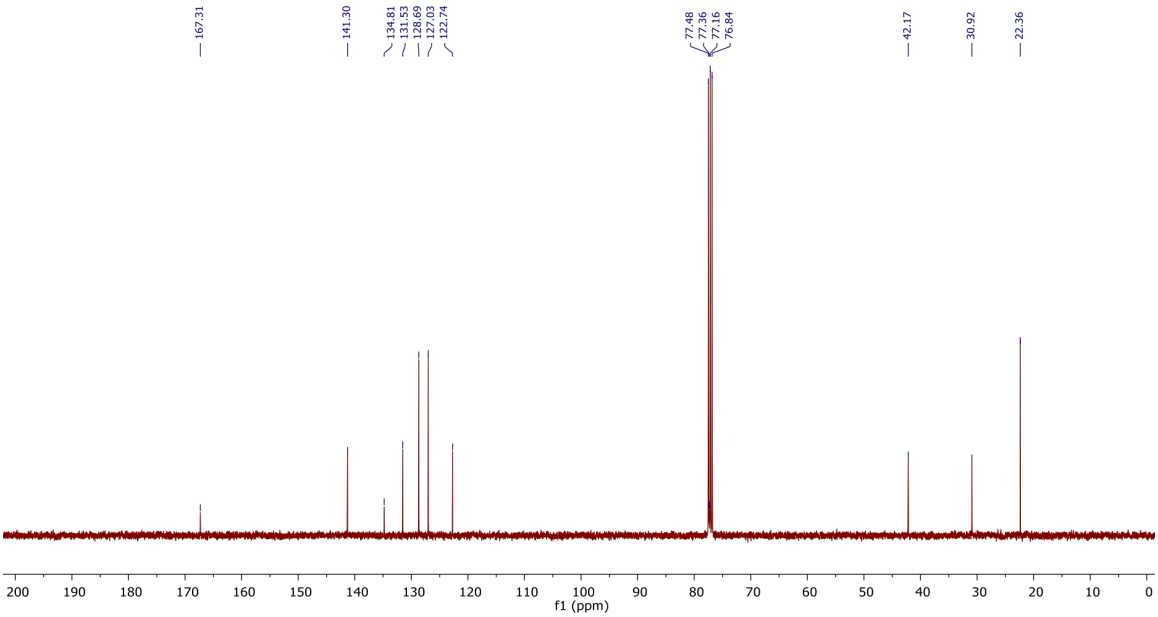


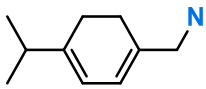




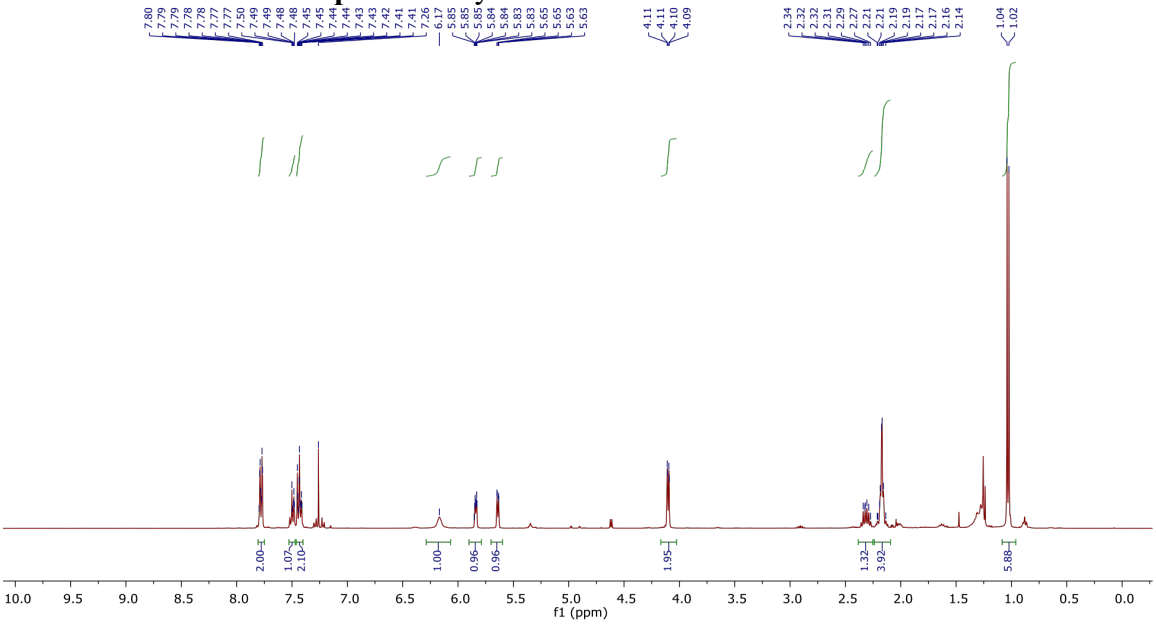
Bz-**6o** from enzymatic reaction and derivatization

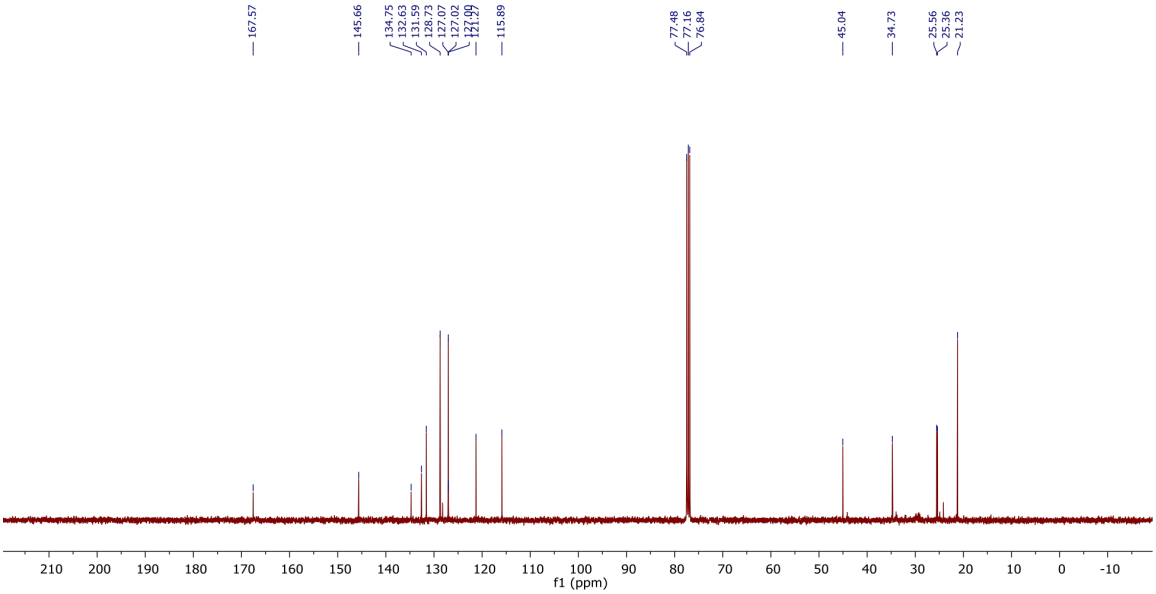






Bz-**6p** from enzymatic reaction and derivatization





A.10. References

1. S. Kille, C. G. Acevedo-Rocha, L. P. Parra, Z.-G. Zhang, D. J. Opperman, M. T. Reetz, J. P. Acevedo, Reducing Codon Redundancy and Screening Effort of Combinatorial Protein Libraries Created by Saturation Mutagenesis. *ACS Synth. Biol*. **2**, 83-92 (2013).
2. D. G. Gibson, L. Young, R-Y. Chuang, J. C. Venter, C. A. Hutchinson III, H. O. Smith, Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases. *Nat. Methods* **6**, 232-234 (2009).
3. I. Barr, F. Guo, Pyridine Hemochromagen Assay for Determining the Concentration of Heme in Purified Protein Solutions. *Bio. Protoc.* **5**, e1594 (2015).
4. J. C. Bouin, M. T. Atallah, H. O. Hultin, The Glucose Oxidase-Catalase System. *Methods Enzymol*. **44**, 478-488 (1976).
5. J. Liu *et al.*, Oxygenation of Simple Olefins through Selective Allylic C−C Bond Cleavage: A Direct Approach to Cinnamyl Aldehydes. *Angew. Chem. Int. Ed.* **56**, 11940-11944 (2017).
6. T. G. Schenck, B. Bosnich, Homogeneous catalysis. Transition-Metal-Catalyzed Claisen Rearrangements. *J. Am. Chem. Soc.* **107**, 2058-2066 (1985).
7. H. Lebel, C. Spitz, O. Leogane, C. Trudel, M. Parmentier, Stereoselective Rhodium-Catalyzed Amination of Alkenes. *Org. Lett.* **13**, 5460-5463 (2011).
8. B. Miriyala, S. Bhattacharyya, J. S. Williamson, Chemoselective Reductive Alkylation of Ammonia with Carbonyl Compounds: Synthesis of Primary and Symmetrical Secondary Amines. *Tetrahedron* **60**, 1463-1471 (2004).

Chapter III

ENZYMATIC NITROGEN INSERTION INTO UNACTIVATED C–H BONDS

Material from this chapter appears in: "Athavale, S.V.; [*,†] **Gao, S.**; [†] Das, A.; [†] Mallojjala, S. C.; Alfonzo, E.; Long, Y.; Hirschi, J. S.; [*] Arnold, F. H. [*] Enzymatic nitrogen insertion into unactivated C–H bonds. *J. Am. Chem. Soc.* **2022**, *144*(41), 19097–19105."

S. G. discovered the initial activity for unactivated C–H primary amination. S. G. participated in the design and execution of the research. S. G. participated in the preparation of the manuscript.

ABSTRACT

Selective functionalization of aliphatic C–H bonds, ubiquitous in molecular structures, could allow ready access to diverse chemical products. While enzymatic oxygenation of C–H bonds is well established, the analogous enzymatic nitrogen functionalization is still unknown; nature is reliant on pre-oxidized compounds for nitrogen incorporation. Likewise, synthetic methods for selective nitrogen derivatization of unbiased C–H bonds remain elusive. In this work, new-to-nature heme-containing nitrene transferases were used as starting points for the directed evolution of enzymes to selectively aminate and amidate unactivated C(sp^3)–H sites. The desymmetrization of methyl- and ethylcyclohexane with divergent site selectivity is offered as demonstration. The evolved enzymes in these lineages are highly promiscuous and show activity towards a wide array of substrates, providing a foundation for further evolution of nitrene transferase function. Computational studies and kinetic isotope effects (KIEs) are consistent with a stepwise radical pathway involving an irreversible, enantiodetermining hydrogen atom transfer (HAT), followed by a lower-barrier diastereoselectivity determining radical rebound step. In-enzyme molecular dynamics (MD) simulations reveal a predominantly hydrophobic pocket with favorable dispersion interactions with the substrate. By offering a direct path from saturated precursors, these enzymes present a new biochemical logic for accessing nitrogen-containing compounds.

3.1. Introduction

Various classes of enzymes catalyze challenging C–H functionalization reactions in biological systems. Iconic are the cytochromes P450 heme monooxygenases, which perform selective hydroxylation of substrates using molecular oxygen, via a reactive Fe–oxo species.[1] Conversion of a poorly reactive C–H bond to a C–OH functional group positions the product for further enzymatic manipulations in downstream metabolism and is the basis for site and stereoselective oxidative tailoring of natural compounds.[2] Intriguingly, such biochemical logic – enzymatic heteroatom C–H functionalization – is restricted to oxygen and halogen introduction.[3] Direct nitrogen incorporation would provide access to a plethora of aminated products, yet such a pathway is not known in nature. Instead, biological systems often utilize a pre-oxygenated carbon center to install nitrogen atoms, typically through an enzymatic reductive amination route (Figure 3-1, path a).[4–6]

We seek to revise the logic of biomolecular nitrogen incorporation by engineering heme enzymes to catalyze direct C–H to C–N derivatization. This long-term effort was initiated almost a decade ago with the discovery of the abiotic 'nitrene transferase' activity of heme enzymes, brought about via a nitrogen analog of the Fe-oxo species – an active site heme-nitrenoid (Figure 3-1, path b).[7,8] Utilizing appropriate nitrene precursors (typically sulfonyl azides or hydroxamate esters), P450 enzymes have been evolved to access active site Fe-nitrenoid species for biocatalytic C–H insertion and olefin functionalization reactions.[9]

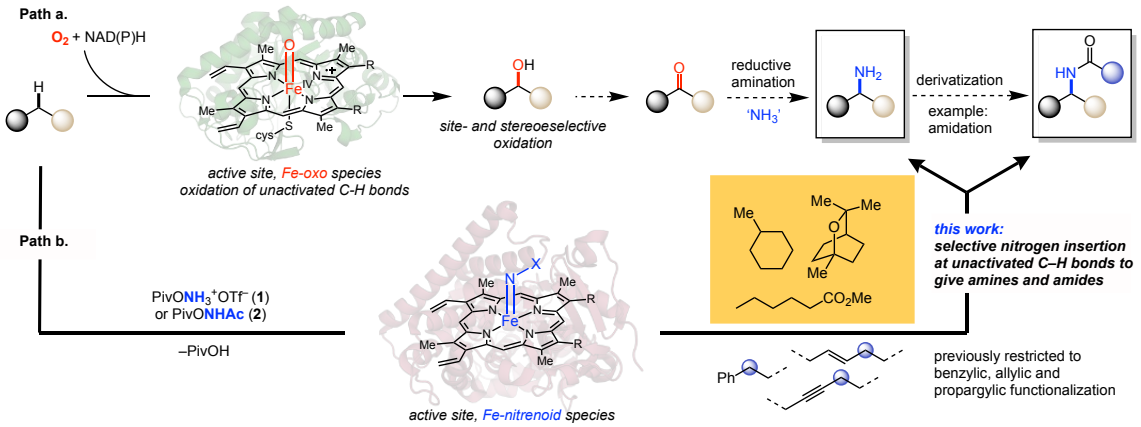


Figure 3-1. Paths for enzymatic nitrogen functionalization of C–H bonds. Path a, prevalent in extant biochemistry, utilizes oxygenation of C–H bonds as a first step. Cytochromes P450, for example, form a high valent Fe-oxo species and direct it for site- and stereo-specific oxygenation of unactivated C–H bonds. The oxygenated products can be further functionalized to nitrogen-containing compounds. In nature, such heteroatom C–H functionalization is restricted to oxygenation or halogenation. Path b depicts new-to-nature heme enzymes that catalyze C–H nitrene insertion reactions via formation of an active site Fe-nitrenoid, analogous to natural P450 chemistry's Fe-oxo species. Previously reported 'nitrene transferases' showed activity only at electronically activated benzylic, allylic and propargylic positions. This work presents enzymes that use simple hydroxylamine derivatives 1 or 2 as nitrenoid precursors for selective nitrogen functionalization of unactivated C–H bonds to give primary amines or acetamides.

Such new-to-nature activity has, however, been limited to insertions at electronically activated benzylic, allylic or propargylic positions (bond dissociation energy, BDE < 85 kcal/mol), restricting catalysis to only a minority of chemical structures.[10–12] The general heme-nitrene transferase activity for modification at the most abundant sites in organic molecules, unactivated C(sp^3)–H bonds (BDE > 95 kcal/mol), which are representative of canonical P450 chemistry, remains unsolved. Both the increased bond strength and site- and stereo- differentiation among competing locations with highly similar chemical environments pose significant challenges, as can be appreciated through an examination of the synthetic C–H functionalization literature. Such differentiation is promoted by biasing

reactivity towards a particular C–H bond through tethering (intramolecular reactions), the use of directing groups, or by exploiting inherent stereoelectronic preferences that direct reactions to electronically favored or sterically accessible positions[13–22] (see references [23–28] for relevant work on C–N bond-forming reactions). Without such bias, the only demonstration of a selective, catalyst-controlled derivgatization is the seminal report by Davies and coworkers [29–31] on rhodium-catalyzed asymmetric carbene insertion into hydrocarbon C–H bonds (also see ref [32]). Recently, selective azidation and unselective nitration of unactivated C–H bonds has been demonstrated with non-heme iron enzymes.[33–35] Here, we establish that P450 enzymes can be engineered for introduction of nitrogen at unbiased, unactivated C(sp^3)–H centers with high site and stereoselectivity, revealing a new biochemical pathway for C–H functionalization that, analogous to P450-catalyzed C–H hydroxylation, can be tuned and diversified by evolution (Figure 3-1).

3.2. Results and Discussion

3.2.1. Discovery of Initial Activity

Since the first report of P450 nitrene transferase activity, the accumulated in-lab collection of evolved enzymes has been used to discover new catalysts with promiscuous reactivity towards increasingly challenging reactions. [9] The collection includes recently evolved enzymes for allylic C–H amination (with *O*-pivaloyl-hydroxylammonium triflic salt (**1**))[10,36,37] and benzylic amidation (with *N*-Acetyl-*O*-pivaloyl hydroxylamine (**2**))[11] of benzene-containing substrates. To detect promiscuous activity at unactivated C–H bonds, the enzyme libraries were challenged with methylcyclohexane (**3**) and ethylcyclohexane (**4**), which were chosen as test substrates that might be accepted in the enzyme active site in place of the benzene ring and elicit functionalization at some position(s). Auspiciously, several tested variants exhibited trace formation of C–H insertion amine and amide product(s) in whole *Escherichia coli* cell reactions, as detected by mass spectrometry (see SI, Section B.2). Consistent with the assumption that a cyclohexane connectivity facilitates substrate binding and functionalization, no product was detected with linear alkanes like *n*-pentane, *n*-hexane, and *n*-heptane. The highest activity for C–H amination was seen with the *APA6* (*a*llylic *p*rimary *a*mination 6) variant, previously evolved for allylic amination of 1-phenylbutene, whereas the highest activity for C–H amidation was observed with the *iAMD5-Y263V* (*i*ntermolecular *amid*ation) variant evolved for benzylic amidation of ethylbenzene (these enzymes possess a coordinating Ser as the axial ligand). The two variants were renamed *uPA0* (*u*nactivated *p*rimary *a*mination 0) and *uAMD0* (*u*nactivated

*am*i*d*ation 0), respectively. Despite previous attempts, this was the first instance when such reactivity was noted, which suggests that with the *uPA* and *uAMD* lineages the enzyme had matured into a sequence space which supports unactivated C–H functionalization.

All constitutional isomers, diastereomers and enantiomers considered, the functionalization of **3** can potentially generate twelve products. This formidable selectivity challenge is nonetheless preferable to the situation for **4**, which can lead to fourteen isomeric products. Given the higher BDE of a methyl C–H bond, functionalization of **3** was hypothesized to be restricted to the cyclohexane ring. Although trace yields precluded accurate determination of product isomer distribution, we were hopeful that directed evolution starting from *uPA0* and *uAMD0* could lead to improved variants and products amenable to further analysis. Thus, enzymatic C–H amination and amidation of **3** with nitrene precursors **1** and **2,** respectively, were chosen as model reactions for the evolution campaign.

3.2.2. Directed Evolution

Directed evolution involved iterative rounds of site-saturation mutagenesis (SSM) at amino acid residues in the active site and flexible regions of the protein or error-prone polymerase-chain reaction (epPCR) to generate mutational diversity from the parent enzyme, followed by identification of daughter variants displaying improved activity as assayed by reverse-phase, liquid-chromatography mass-spectrometry (LC-MS) of whole-cell reactions (see Section 1 and 3–4 of SI for details). The LCMS assay was designed to detect improvements in product yield without resolution of isomeric distribution. A chiral-phase gas

chromatography flame-ionization detection (GC-FID) method that could resolve all twelve acetamide product isomers was also developed for more in-depth characterization of promising variants. Under whole-cell conditions, the amination of **3** (with **1**) was catalyzed by *uPA0* with trace total turnover number (< 5 TTN) and assay yield (< 10 μM product **5**) (Figure 3-2A, SI Section B.3.3.3), which is sufficient for directed evolution. Four rounds of SSM, resulting in mutations M177L, M118Q, E409S and A330H yielded variant *uPA4* that catalyzed product formation with ~35 TTN and 95 μM assay yield. A round of epPCR (to give *uPA5*), followed by two rounds of SSM (to give *uPA7*) and a further round of epPCR resulted in variant *uPA8* (with accumulated mutations L740H, L780P, M263L, L333M and D251N) with moderately improved parameters (~65 TTN, 130 μM assay yield). A final round of SSM uncovered the L188C mutant *uPA9* which catalyzed C–H amination of **3** with an additional 2.5-fold improved product yield (~ 305 μM, ~12% yield) and with ~90 TTN. Overall, the nine rounds of directed evolution resulted in a nearly 20-fold improvement in enzyme performance. To determine selectivity, the amine product was derivatized to an acetamide followed by chiral GC-FID analysis. *uPA9* converts **3** to the 1,3-aminated product with a site-selectivity (*ss*; % of 1,3-isomers) of 86% and a diastereomeric ratio (*dr*; 1,3-*cis*:1,3-*trans*) of 8:1. The transformation is enantioselective, generating (1*R,* 3*S*)-methylcyclohexylamine (**6**) with an enantiomeric ratio (*er*) of 93:7 (see SI, Section B.5–6 for further details about the evolution trajectory, selectivity and assignment of absolute configuration).

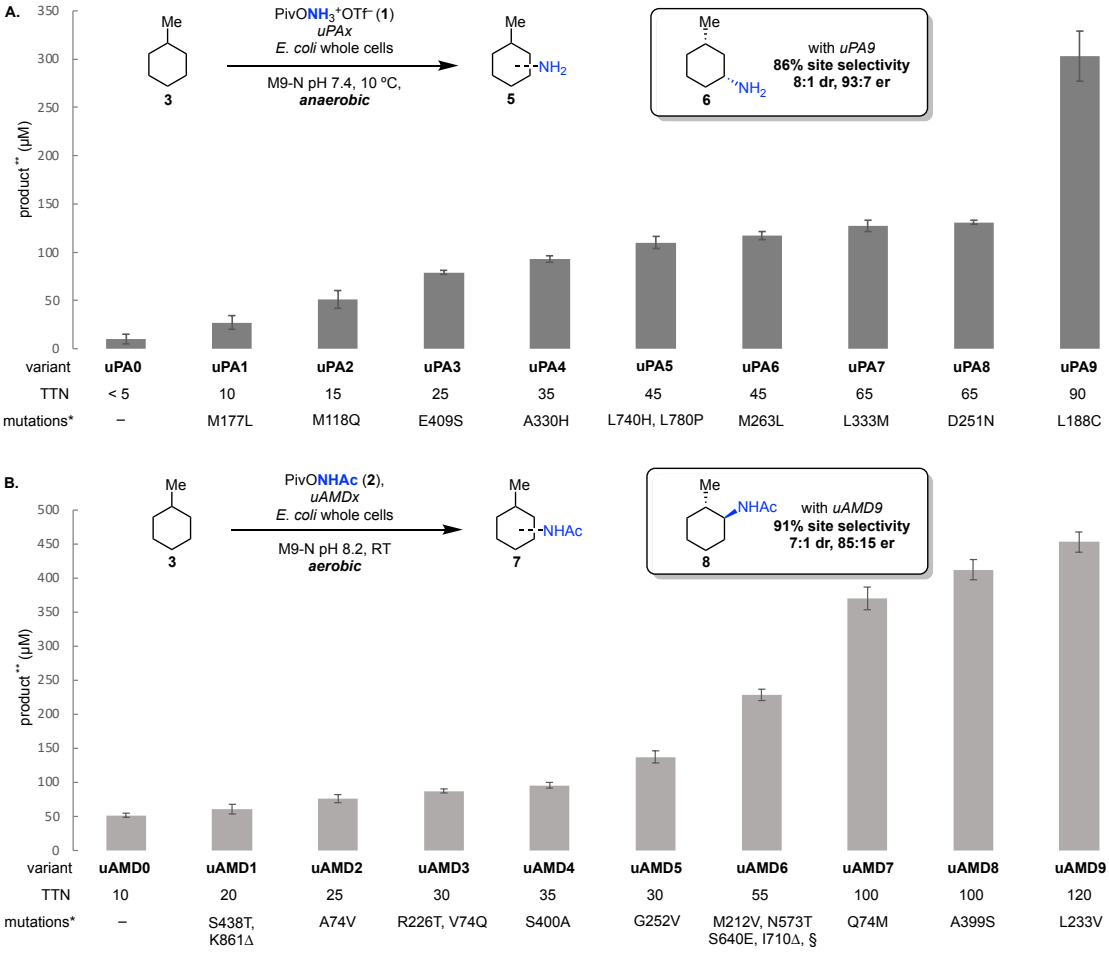


Figure 3-2. Evolution of enzymatic C–H amination (A) and amidation (B). Starting from parent enzymes *uPA0* and *uAMD0*, directed evolution improved amination or amidation of **3** as assayed by product (**5** or **7**) yield. Reactions were set up with whole cells expressing the variant of interest at an optical density (OD) of 30 and 2.5% v/v EtOH. Amination reactions were performed with 2.5 mM of **3**, 5 mM of **1**, and amidation reactions with 2.5 mM of **3**, 10 mM of **2**. Enzymatic C–H functionalization results in the regiodivergent formation of the 1,3-isomer (**6**) or the 1,2-isomer (**8**) as the major product., *with respect to the previous variant, **combined analytical yield of all product isomers, Δ stop codon, § additional frameshift or silent mutations. Refer to text and Supplementary Information for further details.

A parallel directed evolution campaign was undertaken to amplify the *uAMD0*-catalyzed trace acetamidation of **3** with **2** (~10 TTN, 50 μM product **7**) (Figure 3-2B, SI Section

B.4.3.3). Five rounds of evolution (SSM) resulted in variant *uAMD5* (incorporating mutations S438T, A74V, R226T, V74Q, S400A and G252V; in *uAMD1,* an inadvertent single nucleotide deletion also resulted in thirteen amino acid changes at the C terminal and a premature stop codon at site 861, see Supplementary Information) with moderately improved performance (~30 TTN, 140 μM). The following epPCR round (giving *uAMD6*) introduced three mutations (M212V, N573D and S640P) and a single base-pair deletion at site 706, resulting in four mutations (706-709; NYEG to TMKE) and a premature stop codon (710 stop), with an improvement in product yield to ~230 μM and 55 TTN. Two further rounds of SSM (giving *uAMD8*), followed by a final round of epPCR yielded variant *uAMD9* (with accumulated mutations Q74M, A399S and L233V) with an improved activity of ~120 TTN and 450 μM (9.0 %) yield. *uAMD9* harbors an axial alanine mutation, and delivers **8** with 91% *ss* (% of 1,2-isomers), 7:1 *dr* (1,2-*trans*: 1,2-*cis*) and 85:15 *er* (see SI, Section B.5-6 for further details regarding the evolution trajectory, selectivity and assignment of absolute configuration). The axial ligand mutation from serine to alanine shifted the absorption peak in the Soret band from 411nm to 423nm. (see SI Section B.4) With *uAMD9,* analytical yield can be increased up to ~1.6 mM (28 %) by setting up whole-cell reactions at high optical density (OD 120) (see SI, Section B.5–6 for further details).

Notably, enzymatic amination and amidation result in regiodivergent functionalization of model substrate **3**. Based on a combination of statistical chance (number of C–H bonds) and steric accessibility, the 3-position of methylcyclohexane can be expected to be the most favored site for derivatization. Whereas this is in fact observed in the amination reaction, the *contra-steric* preference of amidation for the 2-position is striking. Among these

enzymes, the difference in active site binding of **3** must be sufficient to completely select for a different C–H center. The strategy of employing distinct reagent (nitrene precursor) – enzyme combinations (in this instance **1** - *uPA* or **2 –** *uAMD*) thus allows differentiation of site-selectivity for C–H functionalization. To the best of our knowledge, such a catalytic, stereoselective nitrogen functionalization of unbiased C–H bonds has no precedent in either the synthetic or biochemistry literature. Furthermore, previously engineered P450 nitrene transferases are sensitive to oxygen and require anaerobic reaction conditions for function (note that a natural P450 nitrene transferase for aziridination has been reported to operate under aerobic conditions).[9–12,38] Consistent with this observation, the *uPA* amination reactions displayed oxygen sensitivity. Astonishingly, however, all variants in the *uAMD* lineage displayed identical reactivity and selectivity in the presence and absence of oxygen. Such fortuitous oxygen immunity greatly simplifies the practical implementation of enzymatic C–H amidation.

3.2.3. Substrate Promiscuity

A variety of substituted hydrocarbons show evidence of functionalization by the *uPA* and *uAMD* enzymes (Figure 3-3A, SI Section B.3.3.4). Ethylcyclohexane (**4**) is aminated by *uPA9* to give **9** with exquisite selectivity as a nearly single product among the 14 possible isomers (see SI Section B.5–6 and 11 for selectivity details and assignment of absolute configuration). In contrast to the model substrate **3,** amidation of **4** delivers the 1,2-functionalized product **10** with low *dr* (1:1 *cis:trans*) but excellent *er* for the *cis*-isomer (98:2, absolute configuration unassigned) and moderate *er* for the (*S,S)-trans*-isomer

(75:25). Methylcyclopentane is functionalized with the expected site preference to give **11** and **12** (site-specific diastereomers could not be analytically resolved and a qualitative inspection suggests low *dr* and *er*). Six-, seven-, and eight-membered cycloalkanes were competent towards both amination and amidation (giving achiral **13**-**18**). Linear alkanes

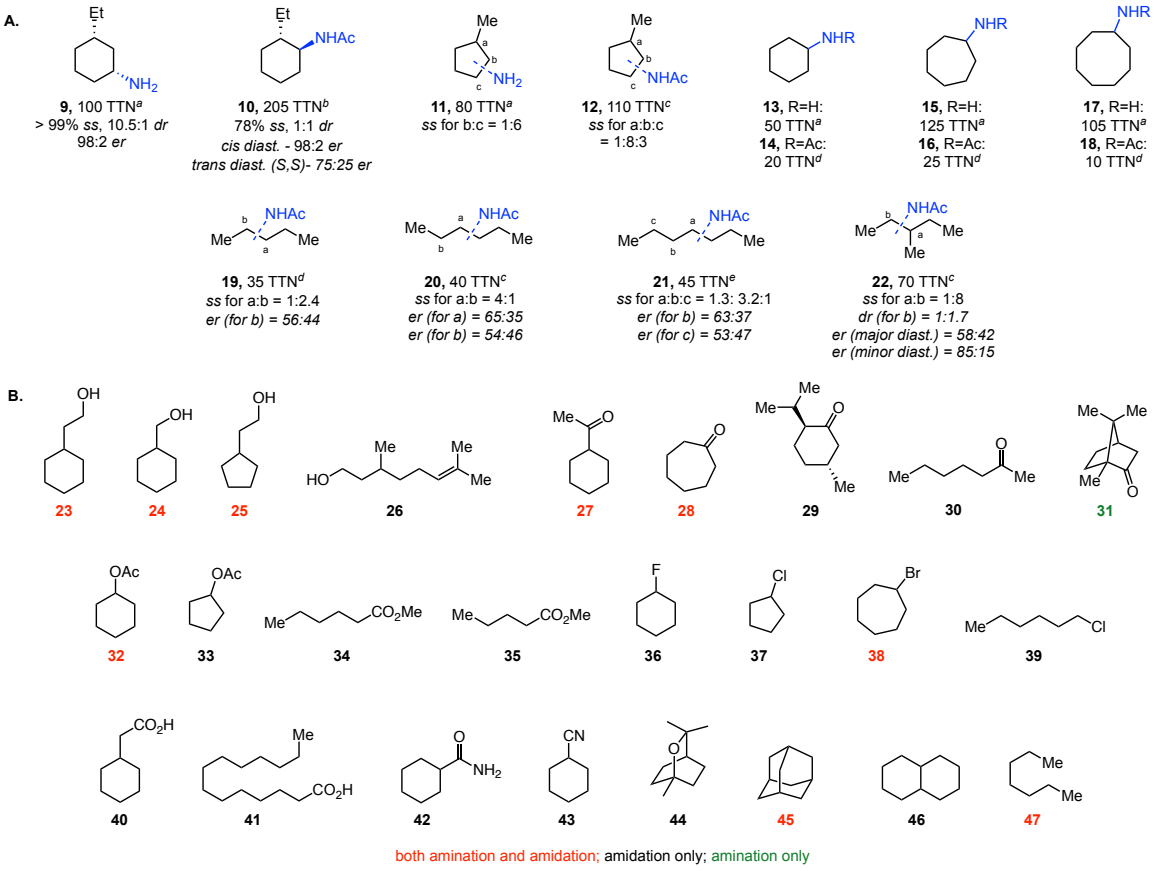


Figure 3-3. Substrate scope of enzymatic amination and amidation. **A**. Quantitative determination of product distribution for functionalization of substituted hydrocarbons. [a] with *uAPA9,* [b] with *uAMD9,* [c] with *uAMD6,* [d] with *uAMD8*,* [e] with *uAMD7-T438S* **B.** A survey of competent substrates showing proof of functionalization (amination with *uAPA9*, amidation with *uAMD8**; yields and site selectivity not determined). In all cases, reactions were performed under whole-cell conditions with *uPA* or *uAMD* variants at an optical density (OD) of 30, with 5 mM of substrate, 10 mM of **1** or **2**, and 2.5 % v/v EtOH. (see Supplementary Information for details).

were amidated with moderate site selectivity with an apparent preference for the C3 position of hexane and heptane (site *a* in **20**, *b* in **21**), but with low enantioselectivity. With 3-methylpentane, high site- selectivity for the secondary position (with 1:1.7 *dr*) was noted along with moderate enantiopreference (85:15 *er)* in case of the minor diastereomer (product **22**).

Comprehensive analysis of product distribution is predicated on the acquisition (and successful analytical resolution) of authentic amine (amide) standards at all potential sites of functionalization, which are not available. As a preliminary assessment of further substrate scope without this limitation, whole-cell reactions with a highly promiscuous variant *uAMD8** (for amidation; *uPA9* for amination) were analyzed for product detection by high resolution mass spectrometry (HRMS). Reactions expressing an inactive mutant (*uAMD8*–A400C*) acted as negative controls. With the active variants, all compounds in Figure 3-3B demonstrated product formation in nitrogen functionalization reactions (see SI, Section B.9). For the substrates tested, amidation reactions were found more tractable to product detection (Fig 3B, red and black) whereas amine functionalization could be unambiguously established for only a subset (Fig 3B, red and green). Functionalization is detected for various substituted cycloalkanes harboring alcohols (**23**–**25**), ketones (**27**–**29**), esters (**32**, **33**) and halide substitutions (**36**–**38**). Carboxylic acid **40**, amide **42** and nitrile **43** were competent. Acetamidation was also detected with terpenes **26**, **29** and **44**, while camphor (**31**) could only be aminated. Some linear substrates (**26**, **30, 34, 35, 39** and **41**) were also competent, although it appears that cyclic reactants are favored. No di-amination or di-amidation product was detected by mass spectrometry. We and others have shown

extensively how directed evolution can amplify even trace initial activity to synthetically useful levels.[39–41] Notable recent successes in industrial biocatalytic process development which started with low target activity are further testaments.[42–44] This 'initial activity' survey reveals the wide substrate range of *uPA* and *uAMD* enzymes and supports the notion that they can be founding members of an enzyme family for synthesis of diverse nitrogenated compounds by C–H functionalization.

3.2.4. Mechanistic Insights

To understand how these enzymes promote such a challenging C–H functionalization selectively, we carried out computational studies with a model porphyrin system along with in-enzyme molecular dynamics (MD) simulations (Figure 3-4). Studies were focused on the amidation (*uAMD*) system to allow comparison with earlier findings (see Section 12 of SI for a discussion of *uPA* reactions). [11] A free energy surface for the C–H amidation of **3** was calculated using density functional theory (DFT) on a simplified iron-porphyrin model system.

In agreement with previous enzymatic nitrene insertions, the reaction occurs via a stepwise pathway (Fig. 4A, blue reaction coordinate); the lowest energy transition structure (TS) for hydrogen atom transfer to the iron-nitrenoid ($^3TS_{HAT}$ 29.7 kcal/mol) was located on the triplet surface.[11] As the irreversible step that desymmetrizes the substrate, the regio- and enantioselectivity for the amidation of **3** is determined at $\mathbf{^3TS_{HAT}}$. Unsurprisingly, DFT calculations of the model system did not exhibit significant energy differences between regioisomeric hydrogen atom transfer (HAT) TSs (see SI, Figure B-1), indicating that the

experimentally observed site-selectivity is induced through interactions within the binding pocket. The HAT step is followed by a significantly lower barrier for radical rebound (^{3}**TS**$_{\text{C-N}}$ 17.4 kcal/mol), which was also located on the triplet surface and involves the equatorial approach of nitrogen at this C–N bond-forming TS. For the chair conformation placing the methyl substituent in the equatorial position, there is a >3.0 kcal/mol preference for equatorial versus axial approach of the nitrenoid in ^{3}**TS**$_{\text{C-N}}$ (the diastereoselectivity-determining step) for the model system.

Similarly, DFT studies on the amination of **3** by a simplified iron-porphyrin catalyst result in HAT as the first irreversible step that desymmetrizes the substrate (^{3}**TS**$_{\text{Amn-HAT}}$ 21.7 kcal/mol, see SI Figure B-12 and B-14 for the full energy surface). As observed in the model system studies of amidation, DFT computations did not exhibit significant energy differences between site-determining HAT TSs. The computed barrier for amination of the unactivated substrate leading to the major constitutional isomer is 7.0 kcal/mol higher in energy than the activated amination reported by Arnold and coworkers (~14.7 kcal/mol, see Figure B-14 of SI).[12] The subsequent radical rebound step has a lower barrier relative to HAT (^{3}**TS**$_{\text{Amn-CN}}$ 14.8 kcal/mol) and a preference for equatorial approach of the nitrenoid (>3.0 kcal/mol). Thus, for both amidation and amination reactions, it is most likely that the rebound step establishes the diastereoselectivity - consistent with Yang and Liu's recent proposal of a stereodetermining rebound step in a related enzymatic C–H sulfonamidation reaction.[45]

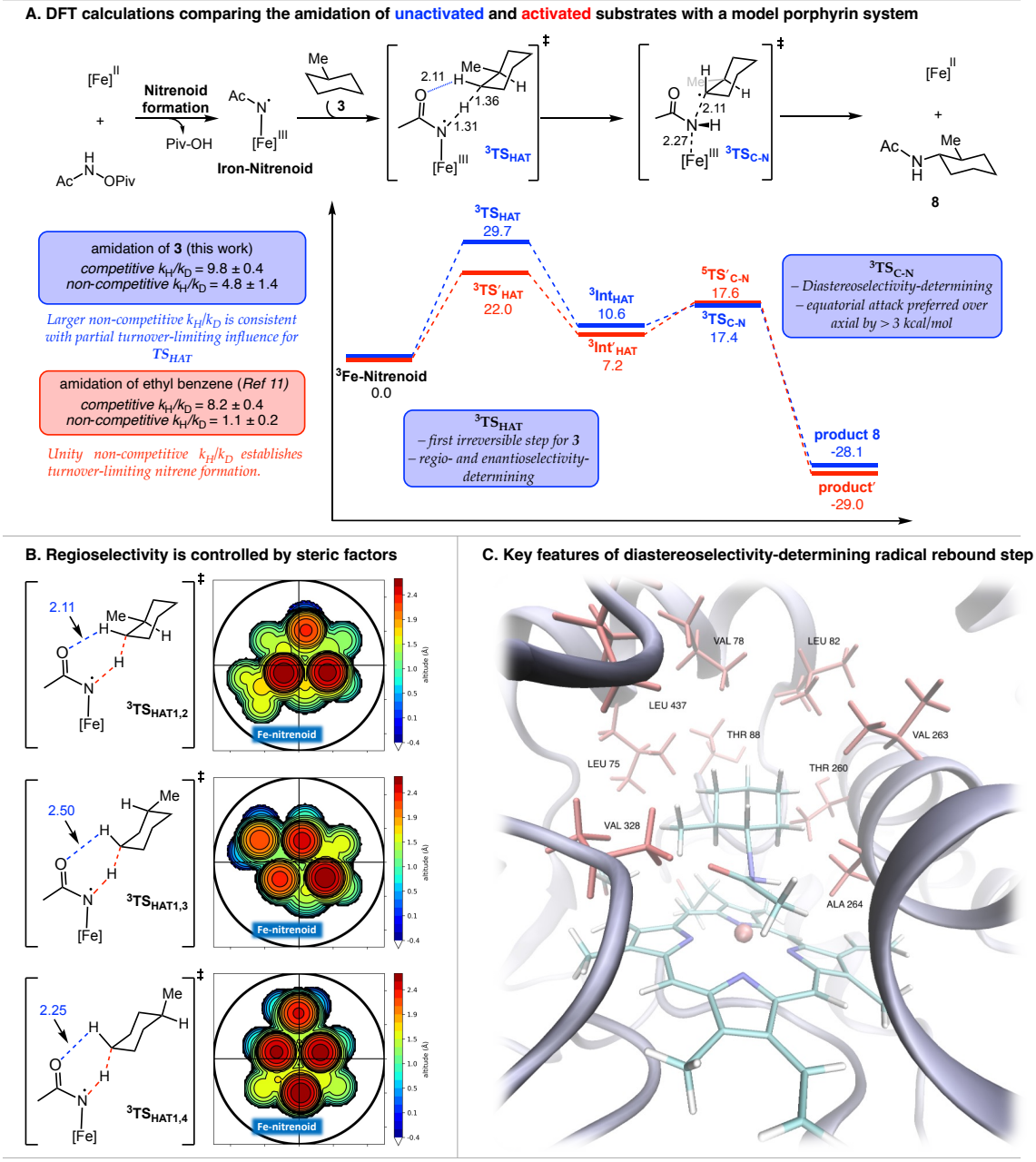


Figure 3-4. Computational modeling aided mechanistic insights into enzyme-catalyzed amidation of methylcyclohexane. A. Comparison of the DFT-computed free energy profiles and experimentally obtained kinetic isotope effects (KIEs) for the amidation of unactivated (methyl cyclohexane) and activated (ethyl benzene) substrates. **B**. Steric map analysis of the H-atom transfer steps ($^3TS_{HAT}$) leading to the formation of the 2-, 3-, and 4-regioisomeric products, as obtained from QMMM computations simulated in *u*AMD8. **C.** Depiction of the diastereoselective radical rebound step (TS_{CN}) for the amidation of **3**, obtained from MD simulations in *u*AMD8.

To determine the accuracy of the computed surface of amidation with respect to the in-enzyme reaction, kinetic isotope effects (KIEs) were measured for cyclohexane vs. cyclohexane-d_{12} (see Section 10.2 of SI). A competitive KIE of ~9.8 is consistent with irreversible HAT and a higher degree of tunneling than previously observed in these systems.[11] A non-competitive KIE of 4.8 indicates that HAT has at least partial turnover-limiting influence in the reaction of **3** (Figure 3-4A, blue). This contrasts with previously measured KIEs for the benzylic C–H amidation of ethyl benzene (Figure 3-4A, red, competitive KIE ~ 8.5 and non-competitive KIE ~1.0) – values consistent with irreversible ^{3}**TS**$'_{HAT}$ (Figure 3-4A, red reaction coordinate) preceded by nitrene formation as the sole rate- limiting step (non-competitive KIE ~1.0 since nitrene formation does not involve ethylbenzene). In the case of **3**, the higher barrier for ^{3}**TS**$_{HAT}$ for the unactivated substrate (29.7 kcal/mol, relative to 22.0 kcal/mol for ^{3}**TS**$'_{HAT}$) likely places this step on par with the The barrier for ^{3}**TS**$_{HAT}$ (29.7 kcal/mol) obtained from DFT calculations on the model system is too high for a reaction that occurs at room temperature (consistent with the experimental observation of no reaction under enzyme-free conditions). Therefore, the enzyme must significantly lower the ^{3}**TS**$_{HAT}$ barrier. To explore the origin of this catalysis as well as to probe the surprisingly high site-selectivity of the reaction, we performed MD simulations on enzyme-bound **3**. 100 ns MD simulations carried out on the docked ^{3}**TS**$_{HAT}$ reveal that the enzyme has evolved a small hydrophobic binding pocket (calculated pocket volume 399 Å^3, allowing for ~25.6% free volume with bound **3**). Furthermore, directed evolution introduced several hydrophobic side chains (V, L, and A) to line the active site and a dispersion interaction computation between the non-polar substrate and the

hydrophobic pocket gives a value of 10.1 kcal/mol. Therefore, it is reasonable that a combination of favorable dispersion interactions and proximity effects lower the effective in-enzyme barrier to a surmountable value at room temperature.

To evaluate site-selectivity for the preferred 1,2-amidation product, MD simulations were performed on the docked HAT TSs of **3** leading to radical formation at the 2-,3-, and 4-positions. We evaluated the interactions of the substrate with the active site using a steric map analysis; steric maps provide a topographic depiction of the substrate environment (Figure 3-4B). QM/MM calculations were performed on a 10 Å pocket from an average of the last 20 ns of the corresponding MD simulation for each of the three regioisomeric structures with constraints on the bonds being formed and broken at the TS. A steric analysis of the three regioisomeric transition structures (Figure 3-4B) indicates that the HAT steps leading to 1,3 and 1,4 isomeric products suffer from significant steric clashes within the pocket (clashes are shown in red contour). In contrast, the pocket can accommodate the 1,2-congener with significantly reduced steric interactions. Dispersion interactions were also more favorable for the 1,2 adduct by 1.1 kcal/mol (see SI Table B-6). In addition, a weak CH–O interaction of the amide carbonyl with the substrate (as indicated by the blue dotted lines in Figure 3-4B) also provides a greater stabilizing effect to the 1,2 TS_{HAT} (estimated stabilizing effect of ~ 0.6 kcal/mol from a fragmentation analysis, see Section 12.2–12.3 of SI for the full analysis). The docked TS for C–N bond formation leading to the major amidation product **8** is shown in Figure 3-4C. This depiction highlights the major features of the binding pocket in the diastereoselectivity-determining structure. Importantly, the “right side” of the cyclohexane is blocked by an alpha helix

causing a preferred binding mode which accommodates the methyl group on the "left side" of the pocket. In addition, a significant number of amino acids with alkyl side chains, which are unique to the *u*AMD variants, constitute a diminutive hydrophobic binding pocket that limits the conformational flexibility of this diastereoselectivity-determining transition state.

In conclusion, the amidation and amination reactions occur via stepwise, irreversible HAT followed by radical rebound, with both steps being stereodetermining. A large HAT barrier for the unactivated substrate is likely reconciled via dispersion interactions between the substrate and a highly hydrophobic pocket. These favorable dispersion interactions likely work in concert with steric clashes within the active site to dictate the stereoselectivity of the title reactions.

3.3. Conclusion

In summary, directed evolution of nitrene transferases has unveiled enzymes that can selectively introduce nitrogen at unactivated C–H bonds under mild conditions. The ability to carry out such a challenging transformation with high site and stereoselectivity was demonstrated for the regiodivergent desymmetrization of methyl- and ethylcyclohexane. Mechanistic studies indicate steric effects in an evolved hydrophobic binding pocket biases site-specific hydrogen atom transfer to promote a selective transformation. Directed evolution may allow development of variants that favor product formation at alternate positions (for instance, 1,3- or 1,4- isomers for amidation, and 1,2- or 1,4- isomers for amination). *uPA* and *uAMD* variants possess wide substrate promiscuity and may be also evolved further for the biocatalytic synthesis of myriad chiral amines and acetamides by the C–H functionalization of readily available starting materials. These evolved enzymes can be considered as the nitrogen counterparts of P450 monooxygenases, where a hydroxylamine derivative is utilized as a nitrogen atom source to modify otherwise inert substrates. By bypassing the reliance on oxygenated precursors, this single-step, enzymatic $C(sp^3)$–H functionalization provides a new route for biology to synthesize nitrogen-containing molecules.

Chapter III Bibliography

1. Ortiz de Montellano, P. R. Hydrocarbon Hydroxylation by Cytochrome P450 Enzymes. *Chem. Rev.* **2010**, *110* (2), 932–948.
2. Greule, A.; Stok, J. E.; De Voss, J. J.; Cryle, M. J. Unrivalled Diversity: The Many Roles and Reactions of Bacterial Cytochromes P450 in Secondary Metabolism. *Nat. Prod. Rep.* **2018**, *35* (8), 757–791.
3. Agarwal, V.; Miles, Z. D.; Winter, J. M.; Eustáquio, A. S.; El Gamal, A. A.; Moore, B. S. Enzymatic Halogenation and Dehalogenation Reactions: Pervasive and Mechanistically Diverse. *Chem. Rev.* **2017**, *117* (8), 5619–5674.
4. Schrewe, M.; Ladkau, N.; Bühler, B.; Schmid, A. Direct Terminal Alkylamino-Functionalization *via* Multistep Biocatalysis in One Recombinant Whole-Cell Catalyst. *Adv. Synth. Catal.* **2013**, *355* (9), 1693–1697.
5. Wu, S.; Zhou, Y.; Wang, T.; Too, H.-P.; Wang, D. I. C.; Li, Z. Highly Regio- and Enantioselective Multiple Oxy- and Amino-Functionalizations of Alkenes by Modular Cascade Biocatalysis. *Nat Commun* **2016**, *7* (1), 11917.
6. Tavanti, M.; Mangas-Sanchez, J.; Montgomery, S. L.; Thompson, M. P.; Turner, N. J. A Biocatalytic Cascade for the Amination of Unfunctionalised Cycloalkanes. *Org. Biomol. Chem.* **2017**, *15* (46), 9790–9793.
7. McIntosh, J. A.; Coelho, P. S.; Farwell, C. C.; Wang, Z. J.; Lewis, J. C.; Brown, T. R.; Arnold, F. H. Enantioselective Intramolecular C□H Amination Catalyzed by Engineered Cytochrome P450 Enzymes In Vitro and In Vivo. *Angew. Chem. Int. Ed.* **2013**, *52* (35), 9309–9312.
8. Miller, D. C.; Athavale, S. V.; Arnold, F. H. Combining Chemistry and Protein Engineering for New-to-Nature Biocatalysis. *Nat Synth* **2022**, *1* (1), 18–23.
9. 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* (5), 1209–1225.
10. Jia, Z.-J.; Gao, S.; Arnold, F. H. Enzymatic Primary Amination of Benzylic and Allylic C(Sp^3)–H Bonds. *J. Am. Chem. Soc.* **2020**, *142* (23), 10279–10283.
11. Athavale, S. V.; Gao, S.; Liu, Z.; Mallojjala, S. C.; Hirschi, J. S.; Arnold, F. H. Biocatalytic, Intermolecular C−H Bond Functionalization for the Synthesis of Enantioenriched Amides. *Angew. Chem. Int. Ed.* **2021**, *60* (47), 24864–24869.
12. Liu, Z.; Qin, Z.-Y.; Zhu, L.; Athavale, S. V.; Sengupta, A.; Jia, Z.-J.; Garcia-Borràs, M.; Houk, K. N.; Arnold, F. H. An Enzymatic Platform for Primary Amination of 1-Aryl-2-Alkyl Alkynes. *J. Am. Chem. Soc.* **2022**, *144* (1), 80–85.
13. *C-H Bond Activation and Catalytic Functionalization I*; Dixneuf, P. H., Doucet, H., Eds.; Topics in Organometallic Chemistry; Springer International Publishing: Cham, 2016; Vol. 55.
14. *C-H Activation for Asymmetric Synthesis*, 1st ed.; Colobert, F., Wencel-Delord, J., Eds.; Wiley, 2019.

15. Rogge, T.; Kaplaneris, N.; Chatani, N.; Kim, J.; Chang, S.; Punji, B.; Schafer, L. L.; Musaev, D. G.; Wencel-Delord, J.; Roberts, C. A.; Sarpong, R.; Wilson, Z. E.; Brimble, M. A.; Johansson, M. J.; Ackermann, L. C–H Activation. *Nat Rev Methods Primers* **2021**, *1* (1), 43.
16. Chen, Z.; Wang, B.; Zhang, J.; Yu, W.; Liu, Z.; Zhang, Y. Transition Metal-Catalyzed C–H Bond Functionalizations by the Use of Diverse Directing Groups. *Org. Chem. Front.* **2015**, *2* (9), 1107–1295.
17. Sambiagio, C.; Schönbauer, D.; Blieck, R.; Dao-Huy, T.; Pototschnig, G.; Schaaf, P.; Wiesinger, T.; Zia, M. F.; Wencel-Delord, J.; Besset, T.; Maes, B. U. W.; Schnürch, M. A Comprehensive Overview of Directing Groups Applied in Metal-Catalysed C–H Functionalisation Chemistry. *Chem. Soc. Rev.* **2018**, *47* (17), 6603–6743.
18. Rej, S.; Ano, Y.; Chatani, N. Bidentate Directing Groups: An Efficient Tool in C–H Bond Functionalization Chemistry for the Expedient Construction of C–C Bonds. *Chem. Rev.* **2020**, *120* (3), 1788–1887.
19. Borpatra, P. J.; Deka, B.; Deb, M. L.; Baruah, P. K. Recent Advances in Intramolecular C–O/C–N/C–S Bond Formation *via* C–H Functionalization. *Org. Chem. Front.* **2019**, *6* (20), 3445–3489.
20. Newton, C. G.; Wang, S.-G.; Oliveira, C. C.; Cramer, N. Catalytic Enantioselective Transformations Involving C–H Bond Cleavage by Transition-Metal Complexes. *Chem. Rev.* **2017**, *117* (13), 8908–8976.
21. Hartwig, J. F.; Larsen, M. A. Undirected, Homogeneous C–H Bond Functionalization: Challenges and Opportunities. *ACS Cent. Sci.* **2016**, *2* (5), 281–292.
22. Liao, G.; Zhang, T.; Lin, Z.; Shi, B. Transition Metal-Catalyzed Enantioselective C−H Functionalization via Chiral Transient Directing Group Strategies. *Angew. Chem. Int. Ed.* **2020**, *59* (45), 19773–19786.
23. Michaudel, Q.; Thevenet, D.; Baran, P. S. Intermolecular Ritter-Type C–H Amination of Unactivated Sp ^{3} Carbons. *J. Am. Chem. Soc.* **2012**, *134* (5), 2547–2550.
24. Tran, B. L.; Li, B.; Driess, M.; Hartwig, J. F. Copper-Catalyzed Intermolecular Amidation and Imidation of Unactivated Alkanes. *J. Am. Chem. Soc.* **2014**, *136* (6), 2555–2563.
25. Sharma, A.; Hartwig, J. F. Metal-Catalysed Azidation of Tertiary C–H Bonds Suitable for Late-Stage Functionalization. *Nature* **2015**, *517* (7536), 600–604.
26. Lee, J.; Jin, S.; Kim, D.; Hong, S. H.; Chang, S. Cobalt-Catalyzed Intermolecular C–H Amidation of Unactivated Alkanes. *J. Am. Chem. Soc.* **2021**, *143* (13), 5191–5200.
27. Fuentes, M. Á.; Gava, R.; Saper, N. I.; Romero, E. A.; Caballero, A.; Hartwig, J. F.; Pérez, P. J. Copper-Catalyzed Dehydrogenative Amidation of Light Alkanes. *Angew. Chem.* **2021**, *133* (34), 18615–18619.
28. Trowbridge, A.; Walton, S. M.; Gaunt, M. J. New Strategies for the Transition-Metal Catalyzed Synthesis of Aliphatic Amines. *Chem. Rev.* **2020**, *120* (5), 2613–2692.

29. Liao, K.; Negretti, S.; Musaev, D. G.; Bacsa, J.; Davies, H. M. L. Site-Selective and Stereoselective Functionalization of Unactivated C–H Bonds. *Nature* **2016**, *533* (7602), 230–234.
30. Fu, J.; Ren, Z.; Bacsa, J.; Musaev, D. G.; Davies, H. M. L. Desymmetrization of Cyclohexanes by Site- and Stereoselective C–H Functionalization. *Nature* **2018**, *564* (7736), 395–399.
31. Davies, H. M. L.; Liao, K. Dirhodium Tetracarboxylates as Catalysts for Selective Intermolecular C–H Functionalization. *Nat Rev Chem* **2019**, *3* (6), 347–360.
32. Li, Y.; Lei, M.; Gong, L. Photocatalytic Regio- and Stereoselective C(Sp3)–H Functionalization of Benzylic and Allylic Hydrocarbons as Well as Unactivated Alkanes. *Nat Catal* **2019**, *2* (11), 1016–1026.
33. Matthews, M. L.; Chang, W.; Layne, A. P.; Miles, L. A.; Krebs, C.; Bollinger, J. M. Direct Nitration and Azidation of Aliphatic Carbons by an Iron-Dependent Halogenase. *Nat Chem Biol* **2014**, *10* (3), 209–215.
34. Chan, N. H.; Gomez, C. A.; Vennelakanti, V.; Du, Q.; Kulik, H. J.; Lewis, J. C. Non-Native Anionic Ligand Binding and Reactivity in Engineered Variants of the Fe(II)- and α-Ketoglutarate-Dependent Oxygenase, SadA. *Inorg. Chem.* **2022**, DOI: acs.inorgchem.2c02872.
35. Gomez, C. A.; Mondal, D.; Du, Q.; Chan, N.; Lewis, J. C. Directed Evolution of a Fe(II)- and α-Ketoglutarate-Dependent Dioxygenase for Site-Selective Azidation of Unactivated Aliphatic C-H Bonds; preprint; Chemistry, 2022. https://doi.org/10.26434/chemrxiv-2022-ddfzp.
36. Legnani, L.; Morandi, B. Direct Catalytic Synthesis of Unprotected 2-Amino-1-Phenylethanols from Alkenes by Using Iron(II) Phthalocyanine. *Angew. Chem. Int. Ed.* **2016**, *55* (6), 2248–2251.
37. Legnani, L.; Prina Cerai, G.; Morandi, B. Direct and Practical Synthesis of Primary Anilines through Iron-Catalyzed C–H Bond Amination. *ACS Catal.* **2016**, *6* (12), 8162–8165.
38. Tsutsumi, H.; Katsuyama, Y.; Izumikawa, M.; Takagi, M.; Fujie, M.; Satoh, N.; Shin-ya, K.; Ohnishi, Y. Unprecedented Cyclization Catalyzed by a Cytochrome P450 in Benzastatin Biosynthesis. *J. Am. Chem. Soc.* **2018**, *140* (21), 6631–6639.
39. Arnold, F. H. Directed Evolution: Bringing New Chemistry to Life. *Angew. Chem. Int. Ed.* **2018**, *57* (16), 4143–4148.
40. Zhang, R. K.; Chen, K.; Huang, X.; Wohlschlager, L.; Renata, H.; Arnold, F. H. Enzymatic Assembly of Carbon–Carbon Bonds via Iron-Catalysed Sp3 C–H Functionalization. *Nature* **2019**, *565* (7737), 67–72.
41. Röthlisberger, D.; Khersonsky, O.; Wollacott, A. M.; Jiang, L.; DeChancie, J.; Betker, J.; Gallaher, J. L.; Althoff, E. A.; Zanghellini, A.; Dym, O.; Albeck, S.; Houk, K. N.; Tawfik, D. S.; Baker, D. Kemp Elimination Catalysts by Computational Enzyme Design. *Nature* **2008**, *453* (7192), 190–195.
42. Savile Christopher K.; Janey Jacob M.; Mundorff Emily C.; Moore Jeffrey C.; Tam Sarena; Jarvis William R.; Colbeck Jeffrey C.; Krebber Anke; Fleitz Fred J.; Brands Jos; Devine Paul N.; Huisman Gjalt W.; Hughes Gregory J. Biocatalytic

Asymmetric Synthesis of Chiral Amines from Ketones Applied to Sitagliptin Manufacture. *Science* **2010**, *329* (5989), 305–309.

43. Huffman, M. A.; Fryszkowska, A.; Alvizo, O.; Borra-Garske, M.; Campos, K. R.; Canada, K. A.; Devine, P. N.; Duan, D.; Forstater, J. H.; Grosser, S. T.; Halsey, H. M.; Hughes, G. J.; Jo, J.; Joyce, L. A.; Kolev, J. N.; Liang, J.; Maloney, K. M.; Mann, B. F.; Marshall, N. M.; McLaughlin, M.; Moore, J. C.; Murphy, G. S.; Nawrat, C. C.; Nazor, J.; Novick, S.; Patel, N. R.; Rodriguez-Granillo, A.; Robaire, S. A.; Sherer, E. C.; Truppo, M. D.; Whittaker, A. M.; Verma, D.; Xiao, L.; Xu, Y.; Yang, H. Design of an in Vitro Biocatalytic Cascade for the Manufacture of Islatravir. *Science* **2019**, *366* (6470), 1255–1259.
44. McIntosh, J. A.; Liu, Z.; Andresen, B. M.; Marzijarani, N. S.; Moore, J. C.; Marshall, N. M.; Borra-Garske, M.; Obligacion, J. V.; Fier, P. S.; Peng, F.; Forstater, J. H.; Winston, M. S.; An, C.; Chang, W.; Lim, J.; Huffman, M. A.; Miller, S. P.; Tsay, F.-R.; Altman, M. D.; Lesburg, C. A.; Steinhuebel, D.; Trotter, B. W.; Cumming, J. N.; Northrup, A.; Bu, X.; Mann, B. F.; Biba, M.; Hiraga, K.; Murphy, G. S.; Kolev, J. N.; Makarewicz, A.; Pan, W.; Farasat, I.; Bade, R. S.; Stone, K.; Duan, D.; Alvizo, O.; Adpressa, D.; Guetschow, E.; Hoyt, E.; Regalado, E. L.; Castro, S.; Rivera, N.; Smith, J. P.; Wang, F.; Crespo, A.; Verma, D.; Axnanda, S.; Dance, Z. E. X.; Devine, P. N.; Tschaen, D.; Canada, K. A.; Bulger, P. G.; Sherry, B. D.; Truppo, M. D.; Ruck, R. T.; Campeau, L.-C.; Bennett, D. J.; Humphrey, G. R.; Campos, K. R.; Maddess, M. L. A Kinase-CGAS Cascade to Synthesize a Therapeutic STING Activator. *Nature* **2022**, *603* (7901), 439–444.
45. Mai, B. K.; Neris, N. M.; Yang, Y.; Liu, P. C–N Bond Forming Radical Rebound Is the Enantioselectivity-Determining Step in P411-Catalyzed Enantioselective C(sp ^{3})–H Amination: A Combined Computational and Experimental Investigation. *J. Am. Chem. Soc.* **2022**, 144 (25), 11215-11225.

Appendix B
SUPPLEMENTARY INFORMATION FOR CHAPTER III

B.1. General Procedures

Unless otherwise noted, all chemicals and reagents were obtained from commercial suppliers (Sigma-Aldrich, VWR, Alfa Aesar, Ambeed, Combi-Blocks, and Enamine) and used without further purification. Silica-gel chromatography was carried out using AMD Silica Gel 60, 230–400 mesh. 1H and 13C NMR spectra were recorded on a Varian Inova 300 MHz or Bruker 400 MHz instrument in $CDCl_3$ and are referenced to residual proton solvent signals. Sonication was performed using a Qsonica Q500 sonicator. High-resolution mass spectra HRMS were acquired from the Caltech Mass Spectral Facility using electrospray ionization (TOF ES+). High-performance liquid-chromatography mass spectroscopy (HPLC-MS) for analysis was carried out using Agilent 1200 series instruments, with C8 (Agilent Poroshell 120, 4.6 × 50 mm, 4 μm) and C18 (Kromasil®, 4.6 × 50 mm, 5 μm) columns. Water and acetonitrile containing 0.1% acetic acid were used as eluents. GC chromatography (GC) was performed on an Agilent Technologies 7820A GC system equipped with a split-mode capillary injection system and flame ionization detectors with CP-Chirasil-DexCB (Agilent) as a chiral stationary phase. Distilled water was utilized for growth media while double distilled water was employed for all buffer preparations.

The M9-N buffer described in this study was prepared as follows:

A 5X stock solution was prepared by dissolving Na_2HPO_4 (34 g), KH_2PO_4 (15 g) and NaCl (2.5 g) in 1 L of water followed by sterilization by autoclaving. A 1X stock was then prepared by diluting 200 mL of the 5X solution to 1 L, followed by addition of 1 mL $CaCl_2$ (0.1 M), 2 mL $MgSO_4$ (1.0 M) and adjusting the pH to 7.4 or 8.2 by the addition of 40% aqueous NaOH.

In all cases, commercially available amines (Sigma, Enamine, Ambeed, Combi Blocks) were utilized as amine standards without further purification. These were acetylated with acetyl chloride to give the corresponding acetyl amides in 50–80% yield as white solids for use as amide standards.[1] Nitrene precursors **1** and **2** were synthesized according to literature procedures.[1,2]

B.1.1. Cloning and Mutagenesis

Electrocompetent *Escherichia coli* BL21 *E. cloni*® (Lucigen) cells were utilized for all experiments. Luria-Bertani (LB) and Hyper-Broth™ (HB, AthenaES) media were used for growth. During growth, media was supplemented with ampicillin (100 μg/mL). HB was also supplemented with Glucose-Mix (AthenaES) according to the manufacturer instructions. Protein induction was performed by supplementing with isopropyl β-D-1-thiogalactopyranoside (IPTG, 0.5 mM) and δ-aminolevulinic acid (ALA, 1.0 mM).

Expression vector pET22b(+) (Novagen) was used for cloning and expression of all variants described in this paper. Site-saturation mutagenesis was performed using a modified QuikChangeTM mutagenesis protocol using the 22-codon trick.[3] The PCR

products were purified with New England Biolabs gel purification kit, and the gaps were repaired using Gibson Mix™.[4] Without further purification, 1 μL of the Gibson product was used to transform 50 μL of electrocompetent *E. coli* cells.

Random mutations were introduced using error-prone PCR with the addition of 200–500 μM $MnCl_2$ to a Taq PCR method as previously reported[5] using the primers described below.

Primers	Sequence (5' → 3')
005	GAA ATA ATT TTG TTT AAC TTT AAG AAG GAG ATA TAC ATA TG
006	GCC GGA TCT CAG TGG TGG TGG TGG TGG TGC TCG AG
007	CAT ATG TAT ATC TCC TTC TTA AAG TTA AAC AAA ATT ATT TC
008	CTC GAG CAC CAC CAC CAC CAC CAC TGA GAT CCG GC

Primers 005 and 006 were used to amplify the region coding for the enzyme (full length) using Taq polymerase and the parent DNA template. Mutations within the full-length fragment were introduced by varying the concentration of $MnCl_2$ during PCR amplification. Similarly, primers 007 and 008 were used to amplify the backbone (i.e., pET22(b)+ vector) fragment using Phusion polymerase and the parent DNA template. The excess DNA templates were digested with DpnI, and the amplified fragments were purified by gel electrophoresis (1% agarose gel). Both the amplified fragments were then assembled into a circular plasmid using the Gibson assembly.[4] Libraries generated with 200, 300, 400,

and 500 μM $MnCl_2$ were tested (one 96-well plate each) to determine which library gave the optimal balance of high diversity and low rate of inactivation (approximately 50% variants are inactive and 50% have similar or better activity compared to the parent). The chosen library was then tested further.

Staggered extension process (StEP) PCR was conducted by pooling plasmids of the variants to be recombined in equimolar fashion and using this mixture as the template DNA.[6] A PCR was then conducted by varying the annealing temperature and also by using a shorter extension time to reshuffle the mutations present in the variants. A second PCR was then performed to amplify the recombined fragment. The recombined insert library was purified and inserted into the pET22(b)+ vector as described above.

B.1.2. Protein Expression for Whole-Cell Reactions

E. coli transformed with pET22b(+) constructs encoding various P411 variants were grown overnight in 5-mL LB medium supplemented with ampicillin (LB_{amp}). Subsequently, 1 mL of this preculture was used to inoculate 50 mL of HB medium in a 150-mL Erlenmeyer flask, supplemented with ampicillin (HB_{amp}) and Glucose Mix for expression. The expression culture was incubated at 37 °C and shaken at 220 rpm for 2.5 hours, until the optical cell density at 600 nm (OD_{600}) was ~0.9 – 1.1. Then, the expression culture was cooled in an ice bath for 30–50 minutes and was treated with ALA and IPTG. Cells were grown at 22 °C and 130 rpm for 20–24 hours, and the shaking radius was 25 mm. Once expression was finished, the cultures were centrifuged (4000*g*, 3 minutes, 4 °C) and the pellets were resuspended to an OD_{600} of 30 (or other appropriate OD as necessary) in M9-

N buffer with pH adjusted to 8.2 (supplemented with 20 mM D-glucose. This cell suspension was utilized to setup whole-cell reactions.

B.1.3. Determination of Hemeprotein Concentration

Protein concentration in the cell was determined by performing hemochrome assay on the cell lysate.[7] Lysate was obtained by sonication (6 minutes, 1 second on, 2 second 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 involume): $\varepsilon\_[557_{reduced} - 540_{oxidized}] = 23.98$ $mM^{-1}cm^{-1}$. 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.

B.1.4. Reaction Screening in 96-well Plate Whole-Cell Format

After a single-site saturation mutagenesis (SSM) or error-prone PCR (epPCR) library was generated, 88 single colonies were randomly picked and cultured in 300 µL of LB medium with 0.1 mg/mL ampicillin (LB_{amp}) in a sterilized 96-well culture plate. The plate typically contained six wells inoculated with single colonies expressing the parent enzyme, and two

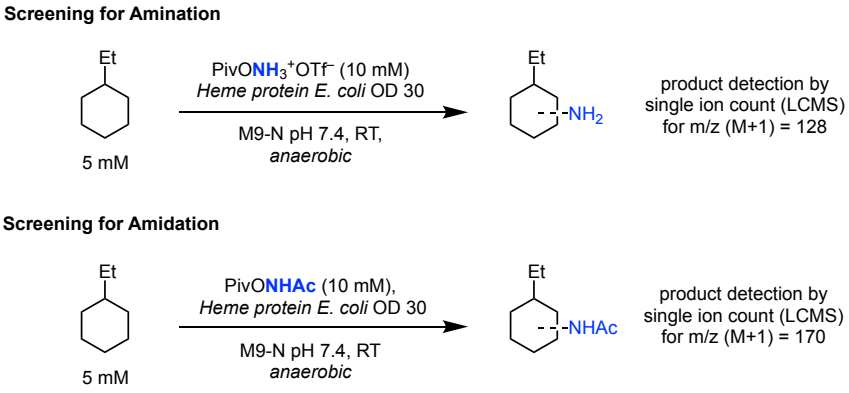
sterile wells. The cultures were grown at 37 °C, 250 rpm, and 80% relative humidity for 8–12 hours. A separate, sterilized 96-well culture plate was filled with 950 μL of Hyperbroth medium containing 0.1 mg/mL ampicillin (HB_{amp}) in each well. Likewise, a glycerol stock replica plate of the preculture (50 μL of preculture added to 50 μL of 50% glycerol per well) was also prepared and stored at -80 °C for future reference. The plate with HB_{amp} was inoculated with the LB preculture (50 μL/well), and incubated at 37 °C, 250 rpm, and 80% relative humidity for 2.5 hours. The plate was cooled on an ice bath for 1 hour, induced with 0.5 mM IPTG and 1 mM ALA (final concentrations), and then expressed at 22 °C and 220 rpm for 20–22 hours. The cells were pelleted (4,000*g*, 3 minutes), and 380 μL of M9-N (pH 8.2, 20 mM D-glucose) were added to each well. After cells were fully resuspended by shaking at 500 rpm, the 96-well plate was transferred to an anaerobic chamber. Inside the anaerobic chamber, to each well were added 10 μL of a reactant stock solution in ethanol (200 mM of hydrocarbon and 400 mM of nitrene precursor dissolved together). Thus, the final reactant concentrations in each well were 5 mM (hydrocarbon) and 10 mM (nitrene precursor). The plate was sealed with aluminum foil tape and shaken at 600 rpm in the anaerobic chamber at room temperature overnight. In case of amidation reactions, all operations were performed aerobically.

Once the plate was taken out of the anaerobic chamber and the seal was removed, ethanol (800 μL/well) was added. The resulting suspension in the wells was mixed by pipetting. The plate was then centrifuged (4,500*g*, 5 minutes) to precipitate proteins and cell debris. The supernatant (200 μL/well) was transferred to a shallow 96-well plate for reverse-phase HPLC-MS analysis. The MS product signals from each well were compared and wells

showing signals higher than the parent wells were identified. These 'hits' were recultured using the wells from the frozen replica glycerol stock plate and then sequenced. Among these, the activities of candidates showing mutations at the targeted site were revalidated in analytical scale reactions. In this way, the best variant was selected as the final hit and was used as the parent for the next round of site-saturation mutagenesis.

B.2. Discovery of Initial Activity

The Arnold lab collection of heme enzymes (including P450s, P411s, protoglobins, and cytochromes *c*) were screened in 96-well plate, whole-cell reactions for product formation (see section 1.4). The collection consisted of nearly 400 distinct variants accumulated from prior directed evolution campaigns. For amination, two variants from the previously described allylic primary amination (APA) lineage[2] showed product formation as assayed by LCMS. For amidation reactions, product formation was seen in reactions with six variants from the intermolecular benzylic amidation (iAMD) lineage.[1] No product was detected in control reactions with free heme.



Amination variants showing product formation:

Variant	Ion Count
P450-BM3 (WT)	0
APA7	181840
APA6 (*uPA0)*	219120

Amidation variants showing product formation:

Variant	Ion Count
P450-BM3 (WT)	0
iAMD6-L75M	10159
iAMD8a-L437M	15506
iAMD8b-A74T	31467
iAMD7-L181Q	45976
iAMD5-Y263L	46290
iAMD5-Y263V (*uAMD0)*	83433

B.3. Evolution for Primary Amination (*uPA* Lineage)

B.3.1 Summary of Directed Evolution

SSM Sites tested	Leading Variant	Method of Mutagenesis	Product **5** (μM)	Total Turnover Number (TTN)
-	*uPA0* (APA6)	-	-	-
F437X, S438X, L324X, A327X, F261X, D267X, V184X, M185X, M177X, L181X, S142X, T149X, R47X, H92X, N201X, L215X, R226X, M118X, R395X, E409X, F162X	*uPA1* (M177L)	SSM	27 ± 7	9 ± 2
M118X, V184X, L215X, R226X, E409X, F437X, S72X, Q73X, L75X, F77X, M78X, A264X	*uPA2* (M118Q)	SSM	51 ± 9	15 ± 1
V184X, L215X, P268X, V328X, A330X, E409X	*uPA3* (E409S)	SSM	79 ± 2	27 ± 2
V184X, L215X, A330X, I401X	*uPA4* (A330H)	SSM	93 ± 3	34 ± 4
-	*uPA5* (L740H, L780P)	epPCR	110 ± 6	47 ± 4
A74X, L82X, A87X, M263X, H266X, T269X	*uPA6* (M263L)	SSM and epPCR	117 ± 4	47 ± 4
L82X, L86X, T88X, S400X, E435X, T436X, N70X, K76X, P326X, P329X, L333X	*uPA7* (L333M)	SSM and epPCR	127 ± 6	65 ± 9

H330X, S400X, L82X, V184X, F437X, S438X	*uPA8* (D251N)	SSM and epPCR	131 ± 2	65 ± 2
K69X, A74X, L75X, M78X, L82X, A87X, Q118X, L177X, L181X, V184X, L188X, L215X, Y256X, Q257X, I259X, T260X, F261X, L263X, A264X, D267X, P268X, S270X, G271X, A327X, V328X, H330X, R395X, E409X, T436X, F437X, S438X, K440X	*uPA9* (L188C)	SSM and epPCR	303 ± 26	89 ± 5

B.3.2 DNA and Protein Sequences

3.2.1 uPA0 DNA sequence:

ACAATTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATTTAC
CGTTATTAAACACAGATAAACCGGTTCAAGCTTTGATGAAAATTGCGGATGA
ATTAGGAGAAATCTTTAAATTCGAGGCGCCTGGTCGTGTAACGCGCTACTTAT
CAAGTCAGCGTCTAATTAAAGAAGCATGCGATGAATCACGCTTTGATAAAAA
CTTAAGTCAAGCGCTTAAATTTATGCGTGATTTTCTTGGAGACGGGTTAGCCA
CAAGCTGGACGCATGAAAAAAATTGGAAAAAAGCGCATAATATCTTACTTCC
AAGCTTTAGTCAGCAGGCAATGAAAGGCTATCATGCGATGATGGTCGATATC
GCCGTGCAGCTTGTTCAAAAGTGGGAGCGTCTAAATGCAGATGAGCATATTG
AAGTATCGGAAGACATGACACGTTTAACGCTTGATACAATTGGTCTTTGCGG
CTTTAACTATCGCTTTAACAGCTTTTACCGAGATCAGCCTCATCCATTTATTAT
AAGTATGGTCCGTGCACTGGATGAAGTAATGAACAAGCTGCAGCGAGCAAAT
CCAGACGACCCAGCTTATGATGAAAACAAGCGCCAGTTTCAAGAAGATATCA
AGGTGATGAACGACCTAGTAGATAAAATTATTGCAGATCGCAAAGCAAGGG
GTGAACAAAGCGATGATTTATTAACGCAGATGCTAAACGGAAAAGATCCAGA
AACGGGTGAGCCGCTTGATGACGGGAACATTCGCTATCAAATTATTACATTCT
TAATGGCGGGACACGATCCGACAAGTGGTCTTTTATCATTTGCGCTGTATTTC
TTAGTGAAAAATCCACATGTATTACAAAAAGTAGCAGAAGAAGCAGCACGA
GTTCTAGTAGATCCTGTTCCAAGCTACAAACAAGTCAAACAGCTTAAATATGT
CGGCATGGTCTTAAACGAAGCGCTGCGCTTATGGCCAGCGGTGCCTGCGTTTT
CCCTATATGCAAAAGAAGATACGGTGCTTGGAGGAGAATATCCTTTAGAAAA

AGGCGACGAAGTAATGGTTCTGATTCCTCAGCTTCACCGTGATAAAACAGTTT
GGGGAGACGATGTGGAGGAGTTCCGTCCAGAGCGTTTTGAAAATCCAAGTGC
GATTCCGCAGCATGCGTTTAAACCGTTTGGACGTGGTCAGCGTGCGTCTATCG
GTCAGCAGTTCGCTCTTCATGAAGCAACGCTGGTACTTGGTATGATGCTAAAA
CACTTTGACTTTGAAGATCATACAAACTACGAGCTCGATATTAAAGAAACTTT
TAGTTTAAAACCTAAAGGCTTTGTGGTAAAAGCAAAATCGAAAAAAATTCCG
CTTGGCGGTATTCCTTCACCTAGCACTGAACAGTCTGCTAAAAAAGTACGCA
AAAAGGCAGAAAACGCTCATAATACGCCGCTGCTTGTGCTATACGGTTCAAA
TATGGGTACCGCTGAAGGAACGGCGCGTGATTTAGCAGATATTGCAATGAGC
AAAGGATTTGCACCGCAGGTCGCAACGCTTGATTCACACGCCGGAAATCTTC
CGCGCGAAGGAGCTGTATTAATTGTAACGGCGTCTTATAACGGTCATCCGCCT
GATAACGCAAAGCAATTTGTCGACTGGTTAGACCAAGCGTCTGCTGATGAAG
TAAAAGGCGTTCGCTACTCCGTATTTGGATGCGGCGATAAAAACTGGGCTAC
TACGTATCAAAAAGTGCCTGCTTTTATCGATGAAACGCTTGCCGCTAAAGGG
GCAGAAAACATCGCTGACCGCGGTGAAGCAGATGCAAGCGACGACTTTGAA
GGCACATATGAAGAATGGCGTGAACATATGTGGAGTGACGTAGCAGCCTACT
TTAACCTCGACATTGAAAACAGTGAAGATAATAAATCTACTCTTTCACTTCAA
TTTGTCGACAGCGCCGCGGATATGCCGCTTGCGAAAATGCACGGTGCGTTTTC
AACGAACGTCGTAGCAAGCAAAGAACTTCAACAGCCAGGCAGTGCACGAAG
CACGCGACATCTTGAAATTGAACTTCCAAAAGAAGCTTCTTATCAAGAAGGA
GATCATTTAGGTGTTATTCCTCGCAACTATGAAGGAATAGTAAACCGTGTAAC
AGCAAGGTTCGGCCTAGATGCATCACAGCAAATCCGTCTGGAAGCAGAAGAA
GAAAAATTAGCTCATTTGCCACTCGCTAAAACAGTATCCGTAGAAGAGCTTC
TGCAATACGTGGAGCTTCAAGATCCTGTTACGCGCACGCAGCTTCGCGCAAT
GGCTGCTAAAACGGTCTGCCCGCCGCATAAAGTAGAGCTTGAAGCCTTGCTT
GAAAAGCAAGCCTACAAAGAACAAGTGCTGGCAAAACGTTTAACAATGCTTG
AACTGCTTGAAAAATACCCGGCGTGTGAAATGAAATTCAGCGAATTTATCGC
CCTTCTGCCAAGCATACGCCCGCGCTATTACTCGATTTCTTCATCACCTCGTG
TCGATGAAAAACAAGCAAGCATCACGGTCAGCGTTGTCTCAGGAGAAGCGTG
GAGCGGATATGGAGAATATAAAGGAATTGCGTCGAACTATCTTGCCGAGCTG
CAAGAAGGAGATACGATTACGTGCTTTATTTCCACACCGCAGTCAGAATTTA
CGCTGCCAAAAGACCCTGAAACGCCGCTTATCATGGTCGGACCGGGAACAGG
CGTCGCGCCGTTTAGAGGCTTTGTGCAGGCGCGCAAACAGCTAAAAGAACAA
GGACAGTCACTTGGAGAAGCACATTTATACTTCGGCTGCCGTTCACCTCATGA
AGACTATCTGTATCAAGAAGAGCTTGAAAACGCCCAAAGCGAAGGCATCATT
ACGCTTCATACCGCTTTTTCTCGCATGCCAAATCAGCCGAAAACATACGTTCA
GCACGTAATGGAACAAGACGGCAAGAAATTGATTGAACTTCTTGATCAAGGA
GCGCACTTCTATATTTGCGGAGACGGAAGCCAAATGGCACCTGCCGTTGAAG
CAACGCTTATGAAAAGCTATGCTGACGTTCACCAAGTGAGTGAAGCAGACGC
TCGCTTATGGCTGCAGCAGCTAGAAGAAAAAGGCCGATACGCAAAAGACGT
GTGGGCTGGGCTCGAGCACCACCACCACCACCAC

3.2.2 uPA0 Protein sequence:

TIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQ
RLIKEACDESRFDKNLSQALKFMRDFLGDGLATSWTHEKNWKKAHNILLPSFSQ
QAMKGYHAMMVDIAVQLVQKWERLNADEHIEVSEDMTRLTLDTIGLCGFNYRF
NSFYRDQPHPFIISMVRALDEVMNKLQRANPDDPAYDENKRQFQEDIKVMNDLV
DKIIADRKARGEQSDDLLTQMLNGKDPETGEPLDDGNIRYQIITFLMAGHDPTSG
LLSFALYFLVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRL
WPAVPAFSLYAKEDTVLGGEYPLEKGDEVMVLIPQLHRDKTVWGDDVEEFRPE
RFENPSAIPQHAFKPFGRGQRASIGQQFALHEATLVLGMMLKHFDFEDHTNYEL
DIKETFSLKPKGFVVKAKSKKIPLGGIPSPSTEQSAKKVRKKAENAHNTPLLVLY
GSNMGTAEGTARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNGH
PPDNAKQFVDWLDQASADEVKGVRYSVFGCGDKNWATTYQKVPAFIDETLAA
KGAENIADRGEADASDDFEGTYEEWREHMWSDVAAYFNLDIENSEDNKSTLSL
QFVDSAADMPLAKMHGAFSTNVVASKELQQPGSARSTRHLEIELPKEASYQEGD
HLGVIPRNYEGIVNRVTARFGLDASQQIRLEAEEEKLAHLPLAKTVSVEELLQYV
ELQDPVTRTQLRAMAAKTVCPPHKVELEALLEKQAYKEQVLAKRLTMLELLEK
YPACEMKFSEFIALLPSIRPRYYSISSSPRVDEKQASITVSVVSGEAWSGYGEYKGI
ASNYLAELQEGDTITCFISTPQSEFTLPKDPETPLIMVGPGTGVAPFRGFVQARKQ
LKEQGQSLGEAHLYFGCRSPHEDYLYQEELENAQSEGIITLHTAFSRMPNQPKTY
VQHVMEQDGKKLIELLDQGAHFYICGDGSQMAPAVEATLMKSYADVHQVSEA
DARLWLQQLEEKGRYAKDVWAGLEHHHHHH

3.2.3 uPA9 DNA sequence:
ACAATTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATTTAC
CGTTATTAAACACAGATAAACCGGTTCAAGCTTTGATGAAAATTGCGGATGA
ATTAGGAGAAATCTTTAAATTCGAGGCGCCTGGTCGTGTAACGCGCTACTTAT
CAAGTCAGCGTCTAATTAAAGAAGCATGCGATGAATCACGCTTTGATAAAAA
CTTAAGTCAAGCGCTTAAATTTATGCGTGATTTTCTTGGAGACGGGTTAGCCA
CAAGCTGGACGCATGAAAAAAATTGGAAAAAAGCGCATAATATCTTACTTCC
AAGCTTTAGTCAGCAGGCAATGAAAGGCTATCATGCGCAGATGGTCGATATC
GCCGTGCAGCTTGTTCAAAAGTGGGAGCGTCTAAATGCAGATGAGCATATTG
AAGTATCGGAAGACATGACACGTTTAACGCTTGATACAATTGGTCTTTGCGG
CTTTAACTATCGCTTTAACAGCTTTTACCGAGATCAGCCTCATCCATTTATTAT
AAGTCTGGTCCGTGCACTGGATGAAGTAATGAACAAGTGTCAGCGAGCAAAT
CCAGACGACCCAGCTTATGATGAAAACAAGCGCCAGTTTCAAGAAGATATCA
AGGTGATGAACGACCTAGTAGATAAAATTATTGCAGATCGCAAAGCAAGGG
GTGAACAAAGCGATGATTTATTAACGCAGATGCTAAACGGAAAAGATCCAGA
AACGGGTGAGCCGCTTGATAACGGGAACATTCGCTATCAGATTATTACATTCT
TACTTGCGGGACACGATCCGACAAGTGGTCTTTTATCATTTGCGCTGTATTTC
TTAGTGAAAAATCCACATGTATTACAAAAAGTAGCAGAAGAAGCAGCACGA
GTTCTAGTAGATCCTGTTCCAAGCTACAAACAAGTCAAACAGCTTAAATATGT
CGGCATGGTCTTAAACGAAGCGCTGCGCTTATGGCCAGCGGTGCCTCATTTTT
CCATGTATGCAAAAGAAGATACGGTGCTTGGAGGAGAATATCCTTTAGAAAA
AGGCGACGAAGTAATGGTTCTGATACCTCAGCTTCACCGTGATAAAACAGTT
TGGGGAGACGATGTGGAGGAGTTCCGTCCAGAGCGTTTTGAAAATCCAAGTG

CGATTCCGCAGCATGCGTTTAAACCGTTTGGACGTGGTCAGCGTGCGTCTATC
GGTCAGCAGTTCGCTCTTCATAGTGCAACGCTGGTACTTGGTATGATGCTAAA
ACACTTTGACTTTGAAGATCATACAAACTACGAGCTCGATATTAAAGAAACT
TTTAGTTTAAAACCTAAAGGCTTTGTGGTAAAAGCAAAATCGAAAAAAATTC
CGCTTGGCGGTATTCCTTCACCTAGCACTGAACAGTCTGCTAAAAAAGTACGC
AAAAAGGCAGAAAACGCTCATAATACGCCGCTGCTTGTGCTATACGGTTCAA
ATATGGGTACCGCTGAAGGAACGGCGCGTGATTTAGCAGATATTGCAATGAG
CAAAGGATTTGCACCGCAGGTCGCAACGCTTGATTCACACGCCGGAAATCTT
CCGCGCGAAGGAGCTGTATTAATTGTAACGGCGTCTTATAACGGTCATCCGC
CTGATAACGCAAAGCAATTTGTCGACTGGTTAGACCAAGCGTCTGCTGATGA
AGTAAAAGGCGTTCGCTACTCCGTATTTGGATGCGGCGATAAAAACTGGGCT
ACTACGTATCAAAAAGTGCCTGCTTTTATCGATGAAACGCTTGCCGCTAAAG
GGGCAGAAAACATCGCTGACCGCGGTGAAGCAGATGCAAGCGACGACTTTG
AAGGCACATATGAAGAATGGCGTGAACATATGTGGAGTGACGTAGCAGCCTA
CTTTAACCTCGACATTGAAAACAGTGAAGATAATAAATCTACTCTTTCACTTC
AATTTGTCGACAGCGCCGCGGATATGCCGCTTGCGAAAATGCACGGTGCGTT
TTCAACGAACGTCGTAGCAAGCAAAGAACTTCAACAGCCAGGCAGTGCACG
AAGCACGCGACATCTTGAAATTGAACTTCCAAAAGAAGCTTCTTATCAAGAA
GGAGATCATTTAGGTGTTATTCCTCGCAACTATGAAGGAATAGTAAACCGTG
TAACAGCAAGGTTCGGCCTAGATGCATCACAGCAAATCCGTCTGGAAGCAGA
AGAAGAAAAATTAGCTCATTTGCCACACGCTAAAACAGTATCCGTAGAAGAG
CTTCTGCAATACGTGGAGCTTCAAGATCCTGTTACGCGCACGCAGCTTCGCGC
AATGGCTGCTAAAACGGTCTGCCCGCCGCATAAAGTAGAGCCTGAAGCCTTG
CTTGAAAAGCAAGCCTACAAAGAACAAGTGCTGGCAAAACGTTTAACAATGC
TTGAACTGCTTGAAAAATACCCGGCGTGTGAAATGAAATTCAGCGAATTTAT
CGCCCTTCTGCCAAGCATACGCCCGCGCTATTACTCGATTTCTTCATCACCTC
GTGTCGATGAAAAACAAGCAAGCATCACGGTCAGCGTTGTCTCAGGAGAAGC
GTGGAGCGGATATGGAGAATATAAAGGAATTGCGTCGAACTATCTTGCCGAG
CTGCAAGAAGGAGATACGATTACGTGCTTTATTTCCACACCGCAGTCAGAAT
TTACGCTGCCAAAAGACCCTGAAACGCCGCTTATCATGGTCGGACCGGGAAC
AGGCGTCGCGCCGTTTAGAGGCTTTGTGCAGGCGCGCAAACAGCTAAAAGAA
CAAGGACAGTCACTTGGAGAAGCACATTTATACTTCGGCTGCCGTTCACCTCA
CGAAGACTATCTGTATCAAGAAGAGCTTGAAAACGCCCAAAGCGAAGGCATC
ATTACGCTTCATACCGCTTTTTCTCGCATGCCAAATCAGCCGAAAACATACGT
TCAGCACGTAATGGAACAAGACGGCAAGAAATTGATTGAACTTCTTGATCAA
GGAGCGCACTTCTATATTTGCGGAGACGGAAGCCAAATGGCACCTGCCGTTG
AAGCAACGCTTATGAAAAGCTATGCTGACGTTCACCAAGTGAGTGAAGCAGA
CGCTCGCTTATGGCTGCAGCAGCTAGAAGAAAAAGGCCGATACGCAAAAGA
CGTGTGGGCTGGGCTCGAGCACCACCACCACCACCAC

3.2.4 uPA9 Protein sequence:
TIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQ
RLIKEACDESRFDKNLSQALKFMRDFLGDGLATSWTHEKNWKKAHNILLPSFSQ
QAMKGYHAQMVDIAVQLVQKWERLNADEHIEVSEDMTRLTLDTIGLCGFNYRF

NSFYRDQPHPFIISLVRALDEVMNKCQRANPDDPAYDENKRQFQEDIKVMNDLV
DKIIADRKARGEQSDDLLTQMLNGKDPETGEPLDNGNIRYQIITFLLAGHDPTSGL
LSFALYFLVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRL
WPAVPHFSMYAKEDTVLGGEYPLEKGDEVMVLIPQLHRDKTVWGDDVEEFRPE
RFENPSAIPQHAFKPFGRGQRASIGQQFALHSATLVLGMMLKHFDFEDHTNYEL
DIKETFSLKPKGFVVKAKSKKIPLGGIPSPSTEQSAKKVRKKAENAHNTPLLVLY
GSNMGTAEGTARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNGH
PPDNAKQFVDWLDQASADEVKGVRYSVFGCGDKNWATTYQKVPAFIDETLAA
KGAENIADRGEADASDDFEGTYEEWREHMWSDVAAYFNLDIENSEDNKSTLSL
QFVDSAADMPLAKMHGAFSTNVVASKELQQPGSARSTRHLEIELPKEASYQEGD
HLGVIPRNYEGIVNRVTARFGLDASQQIRLEAEEEKLAHLPHAKTVSVEELLQYV
ELQDPVTRTQLRAMAAKTVCPPHKVEPEALLEKQAYKEQVLAKRLTMLELLEK
YPACEMKFSEFIALLPSIRPRYYSISSSPRVDEKQASITVSVVSGEAWSGYGEYKGI
ASNYLAELQEGDTITCFISTPQSEFTLPKDPETPLIMVGPGTGVAPFRGFVQARKQ
LKEQGQSLGEAHLYFGCRSPHEDYLYQEELENAQSEGIITLHTAFSRMPNQPKTY
VQHVMEQDGKKLIELLDQGAHFYICGDGSQMAPAVEATLMKSYADVHQVSEA
DARLWLQQLEEKGRYAKDVWAGLEHHHHHH

B.3.3 Product Quantitation and TTNs

B.3.3.1 Analytical Scale Reaction Setup

M9-N medium (pH 7.4) and D-glucose solution (500 mM in M9-N, pH 7.4) were placed in the anaerobic chamber with oxygen concentration below 20 ppm at least 24 hours. Harvested cells (see section 1.2) were resuspended with M9-N buffer (pH 7.4) in coy chamber to OD_{600} = 40, and 1.2 mL of resuspended cells were aliquoted to 2-mL screw cap vials. The screw cap vials with cells were precooled to 0 ℃ on an ice bath. Unless otherwise specified, 280 μL of the pre-cooled glucose solution and M9-N medium were added to make the total volume to 1480 μL. The resulting mixtures stayed on the ice bath for another 10 minutes, and then the hydrocarbon substrate (40 μL, 0.1 M stock in ethanol) and the nitrene precursor (80 μL, 0.1 M stock in water) were added in a sequential manner. Unless otherwise noted, the reactions were then shaken at 10 ℃ for 12 hours at 250 rpm.

To 1.7-mL Eppendorf microcentrifuge tubes, 1.6 mL ethanol (2 mg/100 mL acetanilide as internal standard) were added, and 1.0 mL of the mixtures were subsequently transferred. After vortexing the mixtures, the microcentrifuge tubes were centrifuged at 14,000*g* for 5 minutes to precipitate coagulated proteins and cell debris. To identify the products and determine yields, 200 μL of the supernatants were transferred to an HPLC vial with an insert and analyzed by reverse-phase HPLC-MS. Product formation was quantified by HPLC-MS based on the calibration curve of the corresponding racemic reference compound.

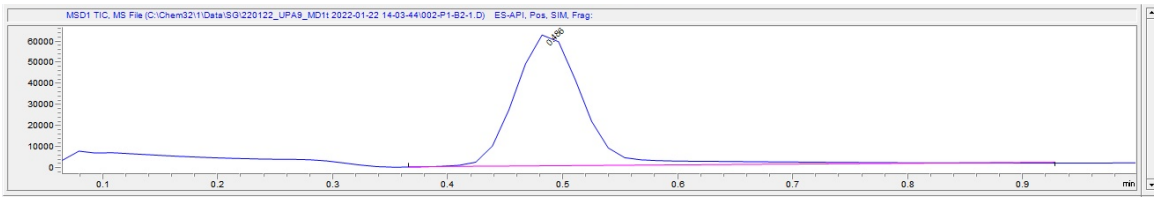


Figure B-1: LC-MS trace for enzymatic amination of methylcyclohexane with uPA9, before acetylation.The product (M+1 ion) is detected in the single ion mass channel.

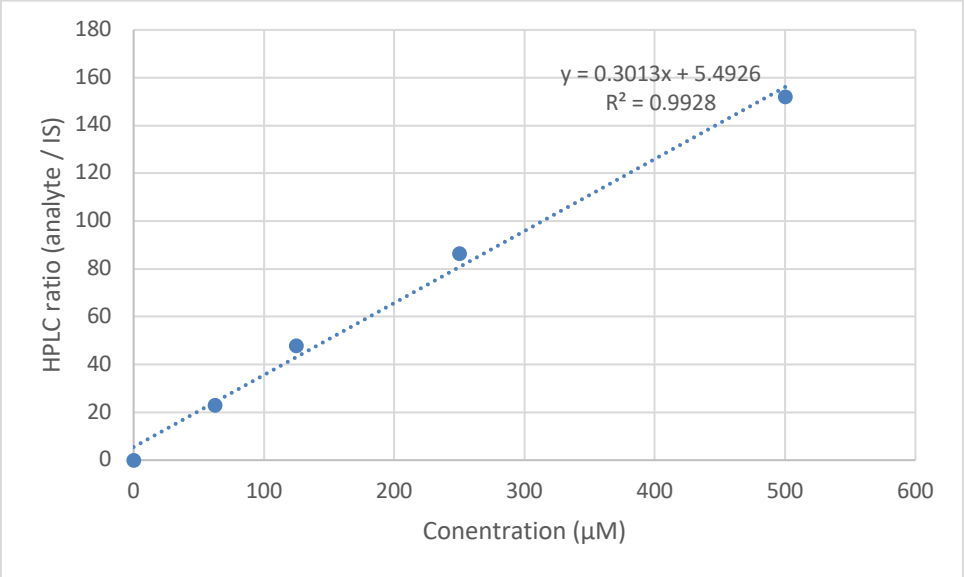
B.3.3.2 Calibration Curves for Standard Products and Reaction Product Quantitation

Enzymatic reactions on analytic scale were performed following the general procedure described in Section 3.3.1. Product formation was quantified by HPLC-MS based on the calibration curve of the corresponding racemic reference compound. TTN is defined as the concentration of product divided by the concentration of heme protein measured by the hemochrome assay (Section 1.3). Calibration curves of commercially available or synthesized reference compounds were created for the determination of yield and TTN. Harvested cells were resuspended with M9-N buffer (pH 7.4) to OD_{600} = 40, and 1.2 mL of resuspended cells were aliquoted to 2-mL screw cap vials. M9-N medium and 280 µL of the precooled glucose solution were added to make the total volume to 1480 µL. Product standard stock (40 µL, 0.1 M stock in ethanol) and 80 µL of water were added. The mixture was shaken overnight at RT and then was quenched with 1.6 mL of ethanol (2 mg/100 mL acetanilide as internal standard), vortexed, and then analyzed by LC-MS. The ion counts at selected ion monitoring mode of product over the indicated internal standard were measured. Cell impurities shared the same retention time. Methods were adjusted so that none of the products elute at the same time with any of the cell impurities. For all the

analyses, water and acetonitrile containing 0.1% acetic acid were used as eluents for a Kromasil® C18 column. The methods used 5–80% acetonitrile (0.5–5.5 minutes) and 80-5% (6.0–6.5 minutes) with 5-µL sample injections. The flow rate was 1.0 mL/minute, and the column was maintained at 22 °C.

B.3.3.3 Yields and TTNs for Methylcyclohexane Reaction across *uPA* Lineage (Figure 3-2A)

Based on ion account in 114 m/z single ion mode for product and UV absorbance at 254 nm for acetanilide as internal standard.



Variant	Pdt	IS	Pdt/IS	[Pdt], µM	[PC], µM	TTN	Avg. TTN	SD TTN
	727.89	160.91	4.52					
	666.26	160.99	4.14					
	784.71	158.22	4.96					
	802.89	156.04	5.15					
uPA0	831.58	158.70	5.24	n.d.	n.d.	n.d.	n.d.	n.d.
	2188.13	198.84	11.00	18.30	2.85	6.43		
	2324.07	154.42	15.05	31.73	3.20	9.93		
uPA1	2217.25	172.06	12.89	24.54	2.85	8.62		
	2426.95	157.76	15.38	32.83	3.20	10.27	8.81	1.74
	3045.19	164.85	18.47	43.08	2.96	14.55		
	3153.31	161.05	19.58	46.76	2.96	15.80		
uPA2	3309.97	157.03	21.08	51.74	3.79	13.67		
	3778.57	153.68	24.59	63.38	3.79	16.74	15.19	1.35
	4571.07	159.17	28.72	77.10	2.82	27.35		
	4854.50	161.18	30.12	81.75	2.82	29.00		
uPA3	4983.06	170.41	29.24	78.84	3.04	25.91		
	5002.14	174.55	28.66	76.89	3.04	25.27	26.88	1.66

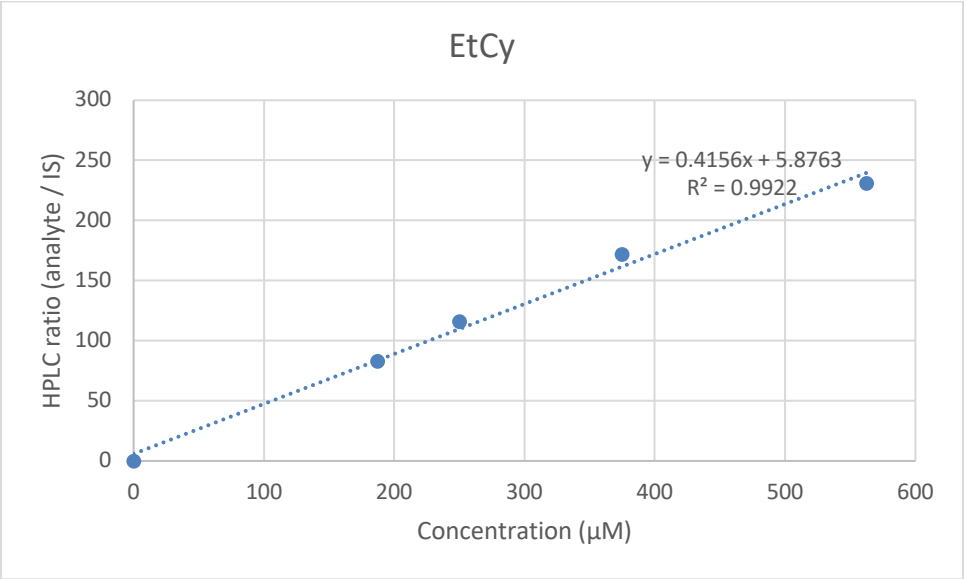
	5331.45	156.57	34.05	94.80	2.56	36.97		
	5454.23	161.10	33.86	94.15	2.56	36.71		
uPA4	5519.64	170.75	32.33	89.07	3.00	29.70		
	6035.09	180.61	33.41	92.68	3.00	30.90	33.57	3.81
	6041.19	156.28	38.66	110.08	2.29	48.10		
	6079.49	147.25	41.29	118.82	2.29	51.92		
uPA5	6297.52	169.82	37.08	104.87	2.44	42.94		
	6331.68	168.36	37.61	106.60	2.44	43.65	46.65	4.19
	6402.63	156.41	40.94	117.65	2.40	49.06		
	6711.37	158.43	42.36	122.38	2.40	51.03		
uPA6	6834.95	171.66	39.82	113.94	2.64	43.19		
	6912.16	173.74	39.79	113.83	2.64	43.15	46.60	4.05
	7041.73	157.14	44.81	130.52	1.80	72.62		
	7067.12	155.41	45.47	132.72	1.80	73.85		
uPA7	7082.61	167.91	42.18	121.78	2.11	57.59		
	7130.15	168.42	42.34	122.30	2.11	57.84	65.47	8.98
	7541.76	168.28	44.82	130.54	2.02	64.68		
	7677.92	171.93	44.66	130.01	2.02	64.42		
uPA8	7683.55	171.76	44.73	130.26	2.00	65.28		
	8111.42	175.63	46.19	135.08	2.00	67.69	65.51	1.49
	12827.01	156.04	82.21	254.64	3.22	79.06		
	13315.89	155.83	85.45	265.43	3.22	82.41		
uPA9	15818.64	160.98	98.26	307.95	3.45	89.24		
	16385.52	162.59	100.78	316.29	3.45	91.65		
	16689.48	161.62	103.26	324.54	3.45	94.05	89.3	5.02

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. The yield of *uPA0* was too low to be accurately determined. n.d. = not determined

B.3.3.4 Yields and TTNs for Substrates in Fig. 3A:

*Ethylcyclohexane (**9**)*

Based on ion account in 128 m/z single ion mode for product and UV absorbance at 254 nm for acetanilide as internal standard.

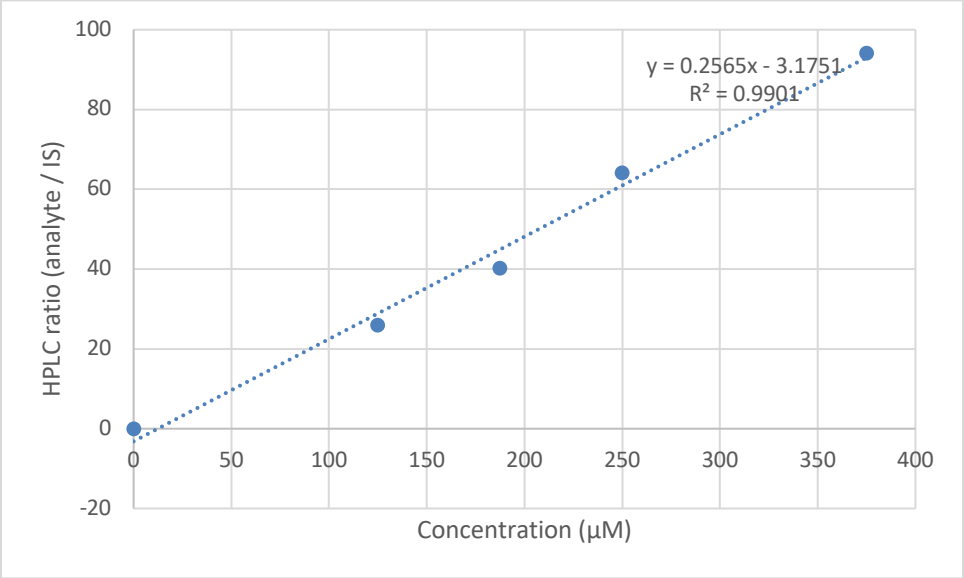


Variant	Pdt	IS	Pdt/IS	yield	[Pdt], µM	[PC], µM	TTN	Avg. TTN	SD TTN
uPA9	18381.70	156.80	117.23	0.1072	267.94	2.69	99.59	101.2	3.4
	17864.10	155.70	114.73	0.1048	261.93	2.69	97.36		
	18070.40	153.20	117.95	0.1079	269.67	2.61	103.16		
	18411.90	153.70	119.79	0.1096	274.10	2.61	104.85		

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.

*Methylcyclopentane (**11**)*

Based on ion account in 100 m/z single ion mode for product and UV absorbance at 254 nm for acetanilide as internal standard.

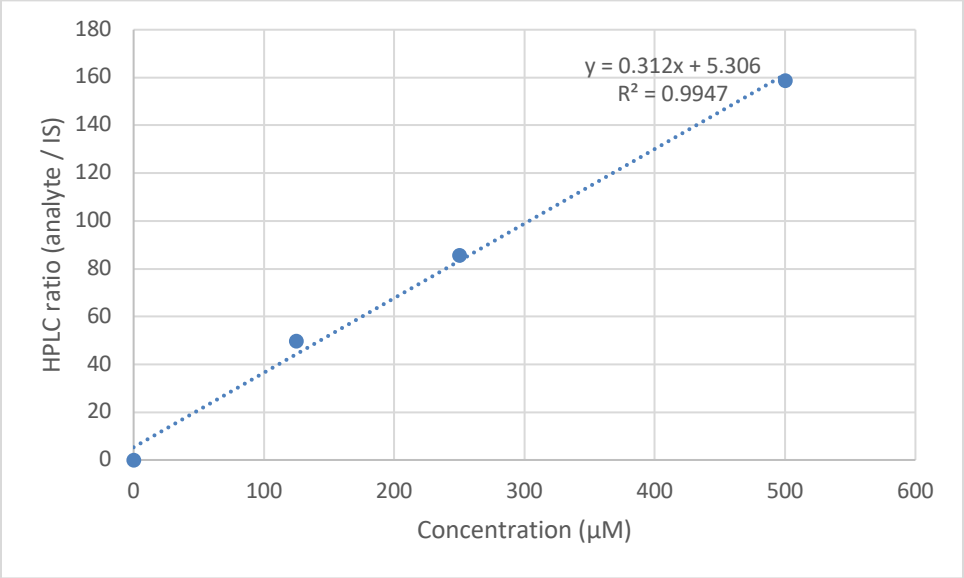


Variant	Pdt	IS	Pdt/IS	yield	[Pdt], µM	[PC], µM	TTN	Avg. TTN	SD TTN
uPA9	8231.10	159.50	51.61	0.0854	213.54	2.69	79.37		
	8129.90	159.70	50.91	0.0843	210.81	2.69	78.36		
	8134.70	158.50	51.32	0.0850	212.44	2.61	81.26		
	8184.60	156.80	52.20	0.0863	215.84	2.61	82.57	80.4	1.9

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.

*Cyclohexane (**13**)*

Based on ion account in 100 m/z single ion mode for product and UV absorbance at 254 nm for acetanilide as internal standard.

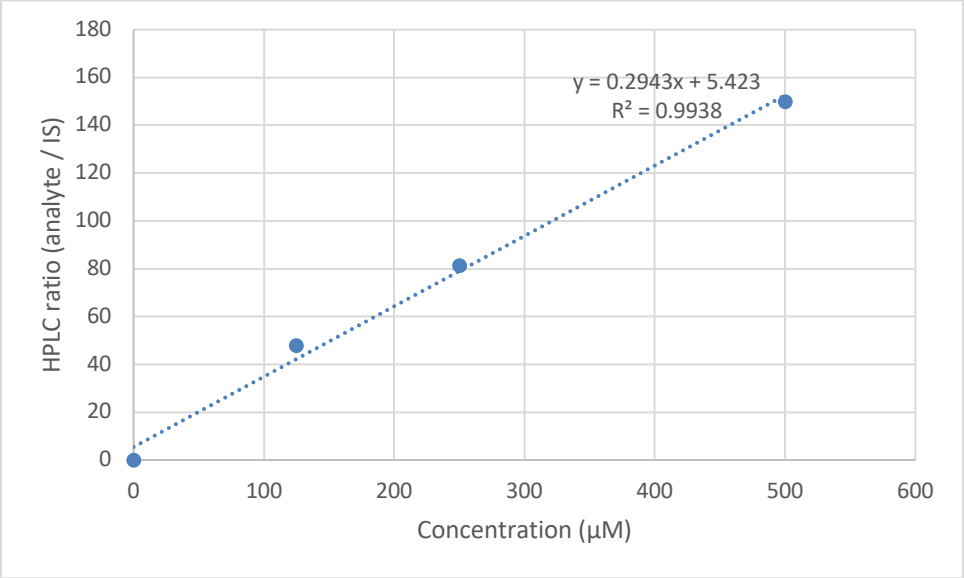


Variant	Pdt	IS	Pdt/IS	yield	[Pdt], µM	[PC], µM	TTN	Avg. TTN	SD TTN
uPA9	6985.60	158.10	44.18	0.0498	124.62	2.69	46.32	46.9	1.5
	7116.20	152.65	46.62	0.0530	132.42	2.69	49.22		
	6788.10	158.50	42.83	0.0481	120.27	2.61	46.01		
	6755.00	157.30	42.94	0.0483	120.64	2.61	46.15		

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.

*Cycloheptane (**15**)*

Based on ion account in 114 m/z single ion mode for product and UV absorbance at 254 nm for acetanilide as internal standard.

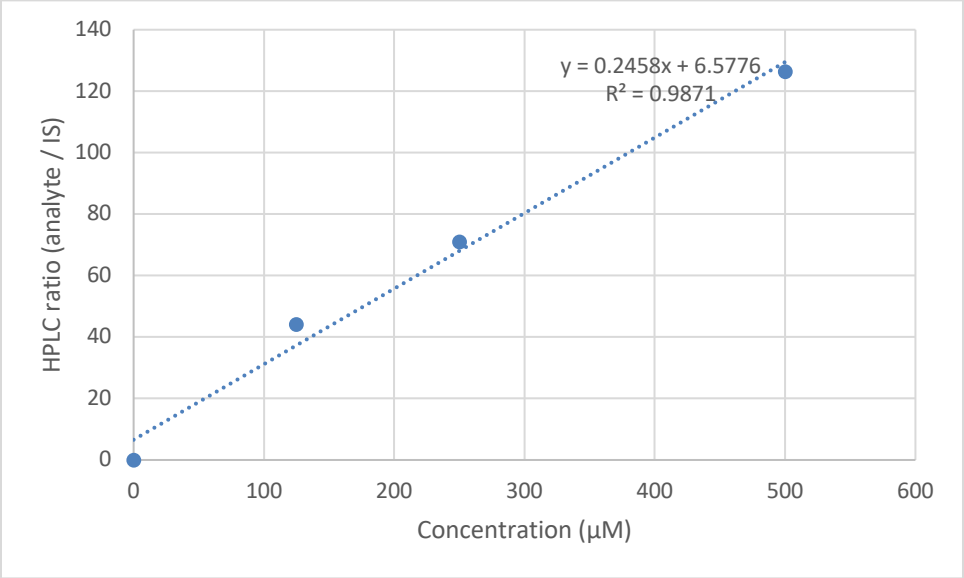


Variant	Pdt	IS	Pdt/IS	yield	[Pdt], µM	[PC], µM	TTN	Avg. TTN	SD TTN
uPA9	16969.10	159.00	106.72	0.1377	344.17	2.69	127.93	127.3	4.9
	16369.30	158.70	103.15	0.1328	332.01	2.69	123.41		
	16003.10	158.80	100.78	0.1296	323.96	2.61	123.92		
	17134.20	157.80	108.58	0.1402	350.48	2.61	134.07		

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.

*Cyclooctane (**17**)*

Based on ion account in 128 m/z single ion mode for product and UV absorbance at 254 nm for acetanilide as internal standard.



Variant	Pdt	IS	Pdt/IS	yield	[Pdt], µM	[PC], µM	TTN	Avg. TTN	SD TTN
uPA9	12134.00	157.30	77.14	0.1148	287.05	2.69	106.69	107.0	4.1
	12214.10	157.20	77.70	0.1157	289.32	2.69	107.54		
	12433.80	158.50	78.45	0.1169	292.36	2.61	111.84		
	11402.60	158.20	72.08	0.1066	266.45	2.61	101.92		

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.

B.4. Evolution for Amidation (*uAmD* Lineage)

B.4.1 Summary of Directed Evolution

SSM Sites Tested	Leading Variant	Mutation Method	Mutations obtained	TTN	Product Yield / mM
-	*uAmD0*	-	-	10 ± 1	51 ± 3
74X, 82X, 181X, 327X, 438X	*uAmD1*	SSM	S438T VSGEAWSGYGEYK(849-861)RLRRSVERIWRI[STOP]	18 ± 2	60 ± 7
70X, 74X, 75X, 78X, 184X, 401X	*uAmD2*	SSM	A74V	25 ± 4	76 ± 6
92X, 107X, 149X, 177X, 226X, 267X, 328X, 330X, 370X, 436X, 437X, 440X	*uAmD3*	SSM	A74Q, R226T	28 ± 4	87 ± 3
47X, 87X, 142X, 252X, 353X, 366X, 400X	*uAmD4*	SSM	S400A	34 ± 7	95 ± 4
190X, 252X, 280X, 290X, 366X, 451X	*uAmD5*	SSM	G252V	28 ± 1	137 ± 9

-	*uAmD6*	epPCR and StEP	M212V, V366V, N573D, S640P, Q673Q, NYEGI(706-710)TMKE[STOP], A1026A	54 ± 4	228 ± 8
73X, 74X, 75X, 225X, 226X, 227X, 252X, 279X, 399X, 400X, 401X, 437X, 438X, 439X	*uAmD7*	SSM	Q74M	99 ± 27	370 ± 17
226X, 227X, 252X, 399X	*uAmD8*	SSM	A399S	102 ± 8	412 ± 15
-	*uAmD9*	epPCR	L233V, E244E	122 ± 8	453 ± 15

Note: In uAMD1, an inadvertent single base pair deletion led to a frameshift mutation and introduced a premature stop codon. Similarly, a single base pair deletion was also seen in uAMD6. In some cases, a silent mutation was noted, for example, A1026A in uAMD6, or E244E in uAMD9.

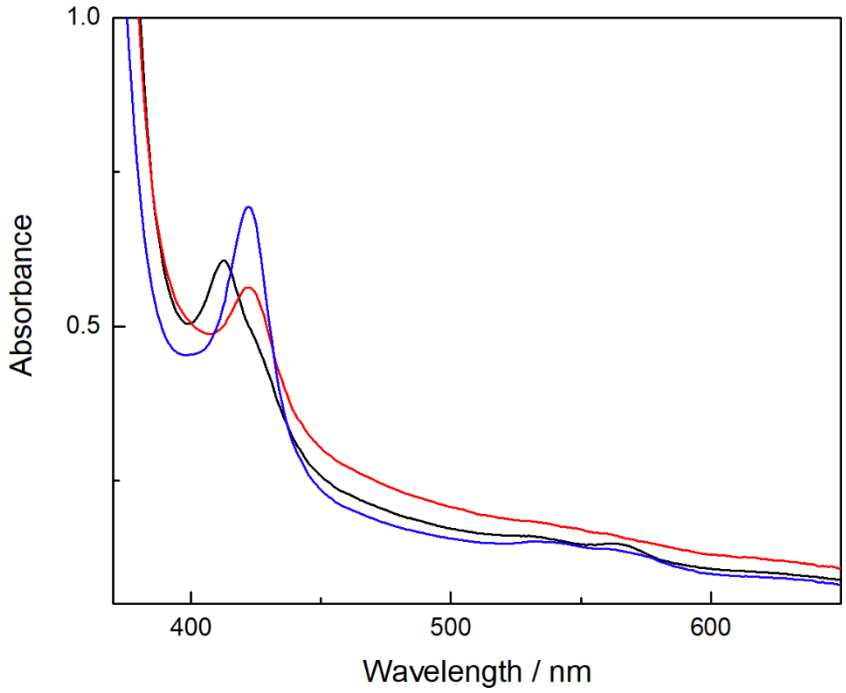


Figure B-2. UV/Vis spectra of the enzyme variants obtained from CO-binding assay. . The starting variant uAmD0 (black –) has serine as the axial ligand and l_{max} at 411 nm. Variant uAmD4 (red –), which has alanine as the axial ligand, has l_{max} at 423 nm. The final variant uAmD9 (blue –) also has alanine as the axial ligand and the l_{max} is at 423 nm.

B.4.2 DNA and Protein Sequences

4.2.1 uAmD0 DNA Sequence:

ACAATTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATTTAC
CGTTATTAAACACAGATAAACCGGTTCAAGCTTTGATGAAAATTGCGGATGA
ATTAGGAGAAATCTTTAAATTCGAGGCGCCTGGTCGTGTAACGCGCTACTTAT
CAAGTCAGCGTCTAATTAAAGAAGCATGCGATGAATCACGCTTTGATAAAAG
TTTAAGTCAAGCGCTGAAATTTCTGCGTGATTTTCTTGGAGACGGGTTAGCCA
CAAGCTGGACGCATGAAAAAAATTGGAAAAAAGCGCATAATATCTTACTTCC
AAGCTTTAGTCAGCAGGCAATGAAAGGCTATCATGCGATGATGGTCGATATC
GCCGTGCAGCTTGTTCAAAAGTGGGAGCGTCTAAATGCAGATGAGCATATTG
AAGTATCGGAAGACATGACACGTTTAACGCTTGATACAATTGGTCTTTGCGG
CTTTAACTATCGCTTTAACAGCTTTTACCGAGATCAGCCTCATCCATTTATTAT
AAGTATGGTCCGTGCACTGGATGAAGTAATGAACAAGCTGCAGCGAGCAAAT
CCAGACGACCCAGCTTATGATGAAAACAAGCGCCAGTTTCAAGAAGATATCA
AGGTGATGAACGACCTAGTAGATAAAATTATTGCAGATCGCAAAGCAAGGG
GTGAACAAAGCGATGATTTATTAACGCAGATGCTAAACGGAAAAGATCCAGA
AACGGGTGAGCCGCTTGATGACGGGAACATTCGCTATCAAATTATTACATTCT
TAGTGGCGGGACACGAAGGTACAAGTGGTCTTTTATCATTTGCGCTGTATTTC
TTAGTGAAAAATCCACATGTATTACAAAAAGTAGCAGAAGAAGCAGCACGA
GTTCTAGTAGATCCTGTTCCAAGCTACAAACAAGTCAAACAGCTTAAATATGT
CGGCATGGTCTTAAACGAAGCGCTGCGCTTATGGCCAACTACGCCTGCGTTTT
CCCTATATGCAAAAGAAGATACGGTGCTTGGAGGAGAATATCCTTTAGAAAA
AGGCGACGAAGTAATGGTTCTGATTCCTCAGCTTCACCGTGATAAAACAGTTT
GGGGAGACGATGTGGAGGAGTTCCGTCCAGAGCGTTTTGAAAATCCAAGTGC
GATTCCGCAGCATGCGTTTAAACCGTTTGGAAACGGTCAGCGTGCGTCTCTTG
GTCAGCAGTTCGCTCTTCATGAAGCAACGCTGGTACTTGGTATGATGCTAAAA
CACTTTGACTTTGAAGATCATACAAACTACGAGCTCGATATTAAAGAAACTTT
AAGCTTAAAACCTAAAGGCTTTGTGGTAAAAGCAAAATCGAAAAAAATTCCG
CTTGGCGGTATTCCTTCACCTAGCACTGAACAGTCTGCTAAAAAAGTACGCA
AAAAGGCAGAAAACGCTCATAATACGCCGCTGCTTGTGCTATACGGTTCAAA
TATGGGTACCGCTGAAGGAACGGCGCGTGATTTAGCAGATATTGCAATGAGC
AAAGGATTTGCACCGCAGGTCGCAACGCTTGATTCACACGCCGGAAATCTTC
CGCGCGAAGGAGCTGTATTAATTGTAACGGCGTCTTATAACGGTCATCCGCCT
GATAACGCAAAGCAATTTGTCGACTGGTTAGACCAAGCGTCTGCTGATGAAG
TAAAAGGCGTTCGCTACTCCGTATTTGGATGCGGCGATAAAAACTGGGCTAC
TACGTATCAAAAAGTGCCTGCTTTTATCGATGAAACGCTTGCCGCTAAAGGG
GCAGAAAACATCGCTGACCGCGGTGAAGCAGATGCAAGCGACGACTTTGAA
GGCACATATGAAGAATGGCGTGAACATATGTGGAGTGACGTAGCAGCCTACT
TTAACCTCGACATTGAAAACAGTGAAGATAATAAATCTACTCTTTCACTTCAA
TTTGTCGACAGCGCCGCGGATATGCCGCTTGCGAAAATGCACGGTGCGTTTTC
AACGAACGTCGTAGCAAGCAAAGAACTTCAACAGCCAGGCAGTGCACGAAG
CACGCGACATCTTGAAATTGAACTTCCAAAAGAAGCTTCTTATCAAGAAGGA
GATCATTTAGGTGTTATTCCTCGCAACTATGAAGGAATAGTAAACCGTGTAAC

AGCAAGGTTCGGCCTAGATGCATCACAGCAAATCCGTCTGGAAGCAGAAGAA
GAAAAATTAGCTCATTTGCCACTCGCTAAAACAGTATCCGTAGAAGAGCTTC
TGCAATACGTGGAGCTTCAAGATCCTGTTACGCGCACGCAGCTTCGCGCAAT
GGCTGCTAAAACGGTCTGCCCGCCGCATAAAGTAGAGCTTGAAGCCTTGCTT
GAAAAGCAAGCCTACAAAGAACAAGTGCTGGCAAAACGTTTAACAATGCTTG
AACTGCTTGAAAAATACCCGGCGTGTGAAATGAAATTCAGCGAATTTATCGC
CCTTCTGCCAAGCATACGCCCGCGCTATTACTCGATTTCTTCATCACCTCGTG
TCGATGAAAAACAAGCAAGCATCACGGTCAGCGTTGTCTCAGGAGAAGCGTG
GAGCGGATATGGAGAATATAAAGGAATTGCGTCGAACTATCTTGCCGAGCTG
CAAGAAGGAGATACGATTACGTGCTTTATTTCCACACCGCAGTCAGAATTTA
CGCTGCCAAAAGACCCTGAAACGCCGCTTATCATGGTCGGACCGGGAACAGG
CGTCGCGCCGTTTAGAGGCTTTGTGCAGGCGCGCAAACAGCTAAAAGAACAA
GGACAGTCACTTGGAGAAGCACATTTATACTTCGGCTGCCGTTCACCTCATGA
AGACTATCTGTATCAAGAAGAGCTTGAAAACGCCCAAAGCGAAGGCATCATT
ACGCTTCATACCGCTTTTTCTCGCATGCCAAATCAGCCGAAAACATACGTTCA
GCACGTAATGGAACAAGACGGCAAGAAATTGATTGAACTTCTTGATCAAGGA
GCGCACTTCTATATTTGCGGAGACGGAAGCCAAATGGCACCTGCCGTTGAAG
CAACGCTTATGAAAAGCTATGCTGACGTTCACCAAGTGAGTGAAGCAGACGC
TCGCTTATGGCTGCAGCAGCTAGAAGAAAAAGGCCGATACGCAAAAGACGT
GTGGGCTGGGCTCGAGCACCACCACCACCACCACTGA

4.2.2 uAmD0 Protein Sequence:
TIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQ
RLIKEACDESRFDKSLSQALKFLRDFLGDGLATSWTHEKNWKKAHNILLPSFSQQ
AMKGYHAMMVDIAVQLVQKWERLNADEHIEVSEDMTRLTLDTIGLCGFNYRFN
SFYRDQPHPFIISMVRALDEVMNKLQRANPDDPAYDENKRQFQEDIKVMNDLVD
KIIADRKARGEQSDDLLTQMLNGKDPETGEPLDDGNIRYQIITFLVAGHEGTSGLL
SFALYFLVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRLW
PTTPAFSLYAKEDTVLGGEYPLEKGDEVMVLIPQLHRDKTVWGDDVEEFRPERF
ENPSAIPQHAFKPFGNGQRASLGQQFALHEATLVLGMMLKHFDFEDHTNYELDI
KETLSLKPKGFVVKAKSKKIPLGGIPSPSTEQSAKKVRKKAENAHNTPLLVLYGS
NMGTAEGTARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNGHPP
DNAKQFVDWLDQASADEVKGVRYSVFGCGDKNWATTYQKVPAFIDETLAAKG
AENIADRGEADASDDFEGTYEEWREHMWSDVAAYFNLDIENSEDNKSTLSLQFV
DSAADMPLAKMHGAFSTNVVASKELQQPGSARSTRHLEIELPKEASYQEGDHLG
VIPRNYEGIVNRVTARFGLDASQQIRLEAEEEKLAHLPLAKTVSVEELLQYVELQ
DPVTRTQLRAMAAKTVCPPHKVELEALLEKQAYKEQVLAKRLTMLELLEKYPA
CEMKFSEFIALLPSIRPRYYSISSSPRVDEKQASITVSVVSGEAWSGYGEYKGIASN
YLAELQEGDTITCFISTPQSEFTLPKDPETPLIMVGPGTGVAPFRGFVQARKQLKE
QGQSLGEAHLYFGCRSPHEDYLYQEELENAQSEGIITLHTAFSRMPNQPKTYVQH
VMEQDGKKLIELLDQGAHFYICGDGSQMAPAVEATLMKSYADVHQVSEADARL
WLQQLEEKGRYAKDVWAGLEHHHHHH

4.2.3 uAmD1 DNA Sequence:

ACAATTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATTTAC
CGTTATTAAACACAGATAAACCGGTTCAAGCTTTGATGAAAATTGCGGATGA
ATTAGGAGAAATCTTTAAATTCGAGGCGCCTGGTCGTGTAACGCGCTACTTAT
CAAGTCAGCGTCTAATTAAAGAAGCATGCGATGAATCACGCTTTGATAAAAG
TTTAAGTCAAGCGCTGAAATTTCTGCGTGATTTTCTTGGAGACGGGTTAGCCA
CAAGCTGGACGCATGAAAAAAATTGGAAAAAAGCGCATAATATCTTACTTCC
AAGCTTTAGTCAGCAGGCAATGAAAGGCTATCATGCGATGATGGTCGATATC
GCCGTGCAGCTTGTTCAAAAGTGGGAGCGTCTAAATGCAGATGAGCATATTG
AAGTATCGGAAGACATGACACGTTTAACGCTTGATACAATTGGTCTTTGCGG
CTTTAACTATCGCTTTAACAGCTTTTACCGAGATCAGCCTCATCCATTTATTAT
AAGTATGGTCCGTGCACTGGATGAAGTAATGAACAAGCTGCAGCGAGCAAAT
CCAGACGACCCAGCTTATGATGAAAACAAGCGCCAGTTTCAAGAAGATATCA
AGGTGATGAACGACCTAGTAGATAAAATTATTGCAGATCGCAAAGCAAGGG
GTGAACAAAGCGATGATTTATTAACGCAGATGCTAAACGGAAAAGATCCAGA
AACGGGTGAGCCGCTTGATGACGGGAACATTCGCTATCAAATTATTACATTCT
TAGTGGCGGGACACGAAGGTACAAGTGGTCTTTTATCATTTGCGCTGTATTTC
TTAGTGAAAAATCCACATGTATTACAAAAAGTAGCAGAAGAAGCAGCACGA
GTTCTAGTAGATCCTGTTCCAAGCTACAAACAAGTCAAACAGCTTAAATATGT
CGGCATGGTCTTAAACGAAGCGCTGCGCTTATGGCCAACTACGCCTGCGTTTT
CCCTATATGCAAAAGAAGATACGGTGCTTGGAGGAGAATATCCTTTAGAAAA
AGGCGACGAAGTAATGGTTCTGATTCCTCAGCTTCACCGTGATAAAACAGTTT
GGGGAGACGATGTGGAGGAGTTCCGTCCAGAGCGTTTTGAAAATCCAAGTGC
GATTCCGCAGCATGCGTTTAAACCGTTTGGAAACGGTCAGCGTGCGTCTCTTG
GTCAGCAGTTCGCTCTTCATGAAGCAACGCTGGTACTTGGTATGATGCTAAAA
CACTTTGACTTTGAAGATCATACAAACTACGAGCTCGATATTAAAGAAACTTT
AACGTTAAAACCTAAAGGCTTTGTGGTAAAAGCAAAATCGAAAAAAATTCCG
CTTGGCGGTATTCCTTCACCTAGCACTGAACAGTCTGCTAAAAAAGTACGCA
AAAAGGCAGAAAACGCTCATAATACGCCGCTGCTTGTGCTATACGGTTCAAA
TATGGGTACCGCTGAAGGAACGGCGCGTGATTTAGCAGATATTGCAATGAGC
AAAGGATTTGCACCGCAGGTCGCAACGCTTGATTCACACGCCGGAAATCTTC
CGCGCGAAGGAGCTGTATTAATTGTAACGGCGTCTTATAACGGTCATCCGCCT
GATAACGCAAAGCAATTTGTCGACTGGTTAGACCAAGCGTCTGCTGATGAAG
TAAAAGGCGTTCGCTACTCCGTATTTGGATGCGGCGATAAAAACTGGGCTAC
TACGTATCAAAAAGTGCCTGCTTTTATCGATGAAACGCTTGCCGCTAAAGGG
GCAGAAAACATCGCTGACCGCGGTGAAGCAGATGCAAGCGACGACTTTGAA
GGCACATATGAAGAATGGCGTGAACATATGTGGAGTGACGTAGCAGCCTACT
TTAACCTCGACATTGAAAACAGTGAAGATAATAAATCTACTCTTTCACTTCAA
TTTGTCGACAGCGCCGCGGATATGCCGCTTGCGAAAATGCACGGTGCGTTTTC
AACGAACGTCGTAGCAAGCAAAGAACTTCAACAGCCAGGCAGTGCACGAAG
CACGCGACATCTTGAAATTGAACTTCCAAAAGAAGCTTCTTATCAAGAAGGA
GATCATTTAGGTGTTATTCCTCGCAACTATGAAGGAATAGTAAACCGTGTAAC
AGCAAGGTTCGGCCTAGATGCATCACAGCAAATCCGTCTGGAAGCAGAAGAA
GAAAAATTAGCTCATTTGCCACTCGCTAAAACAGTATCCGTAGAAGAGCTTC
TGCAATACGTGGAGCTTCAAGATCCTGTTACGCGCACGCAGCTTCGCGCAAT

GGCTGCTAAAACGGTCTGCCCGCCGCATAAAGTAGAGCTTGAAGCCTTGCTT
GAAAAGCAAGCCTACAAAGAACAAGTGCTGGCAAAACGTTTAACAATGCTTG
AACTGCTTGAAAAATACCCGGCGTGTGAAATGAAATTCAGCGAATTTATCGC
CCTTCTGCCAAGCATACGCCCGCGCTATTACTCGATTTCTTCATCACCTCGTG
TCGATGAAAAACAAGCAAGCATCACGGTCAGCGTGCGTCTCAGGAGAAGCGT
GGAGCGGATATGGAGAATATAA

4.2.4 uAmD1 Protein Sequence:
TIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQ
RLIKEACDESRFDKSLSQALKFLRDFLGDGLATSWTHEKNWKKAHNILLPSFSQQ
AMKGYHAMMVDIAVQLVQKWERLNADEHIEVSEDMTRLTLDTIGLCGFNYRFN
SFYRDQPHPFIISMVRALDEVMNKLQRANPDDPAYDENKRQFQEDIKVMNDLVD
KIIADRKARGEQSDDLLTQMLNGKDPETGEPLDDGNIRYQIITFLVAGHEGTSGLL
SFALYFLVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRLW
PTTPAFSLYAKEDTVLGGEYPLEKGDEVMVLIPQLHRDKTVWGDDVEEFRPERF
ENPSAIPQHAFKPFGNGQRASLGQQFALHEATLVLGMMLKHFDFEDHTNYELDI
KETLTLKPKGFVVKAKSKKIPLGGIPSPSTEQSAKKVRKKAENAHNTPLLVLYGS
NMGTAEGTARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNGHPP
DNAKQFVDWLDQASADEVKGVRYSVFGCGDKNWATTYQKVPAFIDETLAAKG
AENIADRGEADASDDFEGTYEEWREHMWSDVAAYFNLDIENSEDNKSTLSLQFV
DSAADMPLAKMHGAFSTNVVASKELQQPGSARSTRHLEIELPKEASYQEGDHLG
VIPRNYEGIVNRVTARFGLDASQQIRLEAEEEKLAHLPLAKTVSVEELLQYVELQ
DPVTRTQLRAMAAKTVCPPHKVELEALLEKQAYKEQVLAKRLTMLELLEKYPA
CEMKFSEFIALLPSIRPRYYSISSSPRVDEKQASITVSVRLRRSVERIWRI

4.2.5 uAmD6 DNA Sequence:
ACAATTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATTTAC
CGTTATTAAACACAGATAAACCGGTTCAAGCTTTGATGAAAATTGCGGATGA
ATTAGGAGAAATCTTTAAATTCGAGGCGCCTGGTCGTGTAACGCGCTACTTAT
CAAGTCAGCGTCTAATTAAAGAAGCATGCGATGAATCACGCTTTGATAAAAG
TTTAAGTCAACAGCTGAAATTTCTGCGTGATTTTCTTGGAGACGGGTTAGCCA
CAAGCTGGACGCATGAAAAAAATTGGAAAAAAGCGCATAATATCTTACTTCC
AAGCTTTAGTCAGCAGGCAATGAAAGGCTATCATGCGATGATGGTCGATATC
GCCGTGCAGCTTGTTCAAAAGTGGGAGCGTCTAAATGCAGATGAGCATATTG
AAGTATCGGAAGACATGACACGTTTAACGCTTGATACAATTGGTCTTTGCGG
CTTTAACTATCGCTTTAACAGCTTTTACCGAGATCAGCCTCATCCATTTATTAT
AAGTATGGTCCGTGCACTGGATGAAGTAATGAACAAGCTGCAGCGAGCAAAT
CCAGACGACCCAGCTTATGATGAAAACAAGCGCCAGTTTCAAGAAGATATCA
AGGTGGTGAACGACCTAGTAGATAAAATTATTGCAGATCGCAAAGCAACGGG
TGAACAAAGCGATGATTTATTAACGCAGATGCTAAACGGAAAAGATCCAGAA
ACGGGTGAGCCGCTTGATGACGTGAACATTCGCTATCAAATTATTACATTCTT
AGTGGCGGGACACGAAGGTACAAGTGGTCTTTTATCATTTGCGCTGTATTTCT
TAGTGAAAAATCCACATGTATTACAAAAAGTAGCAGAAGAAGCAGCACGAG
TTCTAGTAGATCCTGTTCCAAGCTACAAACAAGTCAAACAGCTTAAATATGTC

GGCATGGTCTTAAACGAAGCGCTGCGCTTATGGCCAACTACGCCTGCGTTTTC
CCTATATGCAAAAGAAGATACGGTGCTTGGAGGAGAATATCCTTTAGAAAAA
GGCGACGAAGTAATGGTTCTGATTCCTCAGCTTCACCGTGATAAAACAGTCT
GGGGAGACGATGTGGAGGAGTTCCGTCCAGAGCGTTTTGAAAATCCAAGTGC
GATTCCGCAGCATGCGTTTAAACCGTTTGGAAACGGTCAGCGTGCGGCGCTT
GGTCAGCAGTTCGCTCTTCATGAAGCAACGCTGGTACTTGGTATGATGCTAAA
ACACTTTGACTTTGAAGATCATACAAACTACGAGCTCGATATTAAAGAAACT
TTAACGTTAAAACCTAAAGGCTTTGTGGTAAAAGCAAAATCGAAAAAAATTC
CGCTTGGCGGTATTCCTTCACCTAGCACTGAACAGTCTGCTAAAAAAGTACGC
AAAAAGGCAGAAAACGCTCATAATACGCCGCTGCTTGTGCTATACGGTTCAA
ATATGGGTACCGCTGAAGGAACGGCGCGTGATTTAGCAGATATTGCAATGAG
CAAAGGATTTGCACCGCAGGTCGCAACGCTTGATTCACACGCCGGAAATCTT
CCGCGCGAAGGAGCTGTATTAATTGTAACGGCGTCTTATAACGGTCATCCGC
CTGATAACGCAAAGCAATTTGTCGACTGGTTAGACCAAGCGTCTGCTGATGA
AGTAAAAGGCGTTCGCTACTCCGTATTTGGATGCGGCGATAAAGACTGGGCT
ACTACGTATCAAAAAGTGCCTGCTTTTATCGATGAAACGCTTGCCGCTAAAG
GGGCAGAAAACATCGCTGACCGCGGTGAAGCAGATGCAAGCGACGACTTTG
AAGGCACATATGAAGAATGGCGTGAACATATGTGGAGTGACGTAGCAGCCTA
CTTTAACCTCGACATTGAAAACAGTGAAGATAATAAACCTACTCTTTCACTTC
AATTTGTCGACAGCGCCGCGGATATGCCGCTTGCGAAAATGCACGGTGCGTT
TTCAACGAACGTCGTAGCAAGCAAAGAACTTCAGCAGCCAGGCAGTGCACG
AAGCACGCGACATCTTGAAATTGAACTTCCAAAAGAAGCTTCTTATCAAGAA
GGAGATCATTTAGGTGTTATTCCTCGCACTATGAAGGAATAG

4.2.6 uAmD6 Protein Sequence:
TIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQ
RLIKEACDESRFDKSLSQQLKFLRDFLGDGLATSWTHEKNWKKAHNILLPSFSQQ
AMKGYHAMMVDIAVQLVQKWERLNADEHIEVSEDMTRLTLDTIGLCGFNYRFN
SFYRDQPHPFIISMVRALDEVMNKLQRANPDDPAYDENKRQFQEDIKVVNDLVD
KIIADRKATGEQSDDLLTQMLNGKDPETGEPLDDVNIRYQIITFLVAGHEGTSGLL
SFALYFLVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRLW
PTTPAFSLYAKEDTVLGGEYPLEKGDEVMVLIPQLHRDKTVWGDDVEEFRPERF
ENPSAIPQHAFKPFGNGQRAALGQQFALHEATLVLGMMLKHFDFEDHTNYELDI
KETLTLKPKGFVVKAKSKKIPLGGIPSPSTEQSAKKVRKKAENAHNTPLLVLYGS
NMGTAEGTARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNGHPP
DNAKQFVDWLDQASADEVKGVRYSVFGCGDKDWATTYQKVPAFIDETLAAKG
AENIADRGEADASDDFEGTYEEWREHMWSDVAAYFNLDIENSEDNKPTLSLQFV
DSAADMPLAKMHGAFSTNVVASKELQQPGSARSTRHLEIELPKEASYQEGDHLG
VIPRTMKE

4.2.7 uAmD9 DNA Sequence:
ACAATTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATTTAC
CGTTATTAAACACAGATAAACCGGTTCAAGCTTTGATGAAAATTGCGGATGA
ATTAGGAGAAATCTTTAAATTCGAGGCGCCTGGTCGTGTAACGCGCTACTTAT

CAAGTCAGCGTCTAATTAAAGAAGCATGCGATGAATCACGCTTTGATAAAAG
TTTAAGTCAAATGCTGAAATTTCTGCGTGATTTTCTTGGAGACGGGTTAGCCA
CAAGCTGGACGCATGAAAAAAATTGGAAAAAAGCGCATAATATCTTACTTCC
AAGCTTTAGTCAGCAGGCAATGAAAGGCTATCATGCGATGATGGTCGATATC
GCCGTGCAGCTTGTTCAAAAGTGGGAGCGTCTAAATGCAGATGAGCATATTG
AAGTATCGGAAGACATGACACGTTTAACGCTTGATACAATTGGTCTTTGCGG
CTTTAACTATCGCTTTAACAGCTTTTACCGAGATCAGCCTCATCCATTTATTAT
AAGTATGGTCCGTGCACTGGATGAAGTAATGAACAAGCTGCAGCGAGCAAAT
CCAGACGACCCAGCTTATGATGAAAACAAGCGCCAGTTTCAAGAAGATATCA
AGGTGGTGAACGACCTAGTAGATAAAATTATTGCAGATCGCAAAGCAACGGG
TGAACAAAGCGATGATGTATTAACGCAGATGCTAAACGGAAAAGATCCAGA
GACGGGTGAGCCGCTTGATGACGTGAACATTCGCTATCAAATTATTACATTCT
TAGTGGCGGGACACGAAGGTACAAGTGGTCTTTTATCATTTGCGCTGTATTTC
TTAGTGAAAAATCCACATGTATTACAAAAAGTAGCAGAAGAAGCAGCACGA
GTTCTAGTAGATCCTGTTCCAAGCTACAAACAAGTCAAACAGCTTAAATATGT
CGGCATGGTCTTAAACGAAGCGCTGCGCTTATGGCCAACTACGCCTGCGTTTT
CCCTATATGCAAAAGAAGATACGGTGCTTGGAGGAGAATATCCTTTAGAAAA
AGGCGACGAAGTAATGGTTCTGATTCCTCAGCTTCACCGTGATAAAACAGTC
TGGGGAGACGATGTGGAGGAGTTCCGTCCAGAGCGTTTTGAAAATCCAAGTG
CGATTCCGCAGCATGCGTTTAAACCGTTTGGAAACGGTCAGCGTAGTGCGCTT
GGTCAGCAGTTCGCTCTTCATGAAGCAACGCTGGTACTTGGTATGATGCTAAA
ACACTTTGACTTTGAAGATCATACAAACTACGAGCTCGATATTAAAGAAACT
TTAACGTTAAAACCTAAAGGCTTTGTGGTAAAAGCAAAATCGAAAAAAATTC
CGCTTGGCGGTATTCCTTCACCTAGCACTGAACAGTCTGCTAAAAAAGTACGC
AAAAAGGCAGAAAACGCTCATAATACGCCGCTGCTTGTGCTATACGGTTCAA
ATATGGGTACCGCTGAAGGAACGGCGCGTGATTTAGCAGATATTGCAATGAG
CAAAGGATTTGCACCGCAGGTCGCAACGCTTGATTCACACGCCGGAAATCTT
CCGCGCGAAGGAGCTGTATTAATTGTAACGGCGTCTTATAACGGTCATCCGC
CTGATAACGCAAAGCAATTTGTCGACTGGTTAGACCAAGCGTCTGCTGATGA
AGTAAAAGGCGTTCGCTACTCCGTATTTGGATGCGGCGATAAAGACTGGGCT
ACTACGTATCAAAAAGTGCCTGCTTTTATCGATGAAACGCTTGCCGCTAAAG
GGGCAGAAAACATCGCTGACCGCGGTGAAGCAGATGCAAGCGACGACTTTG
AAGGCACATATGAAGAATGGCGTGAACATATGTGGAGTGACGTAGCAGCCTA
CTTTAACCTCGACATTGAAAACAGTGAAGATAATAAACCTACTCTTTCACTTC
AATTTGTCGACAGCGCCGCGGATATGCCGCTTGCGAAAATGCACGGTGCGTT
TTCAACGAACGTCGTAGCAAGCAAAGAACTTCAGCAGCCAGGCAGTGCACG
AAGCACGCGACATCTTGAAATTGAACTTCCAAAAGAAGCTTCTTATCAAGAA
GGAGATCATTTAGGTGTTATTCCTCGCACTATGAAGGAATAG

4.2.8 uAmD9 Protein Sequence:
TIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQ
RLIKEACDESRFDKSLSQMLKFLRDFLGDGLATSWTHEKNWKKAHNILLPSFSQ
QAMKGYHAMMVDIAVQLVQKWERLNADEHIEVSEDMTRLTLDTIGLCGFNYRF
NSFYRDQPHPFIISMVRALDEVMNKLQRANPDDPAYDENKRQFQEDIKVVNDLV

DKIIADRKATGEQSDDVLTQMLNGKDPETGEPLDDVNIRYQIITFLVAGHEGTSG
LLSFALYFLVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRL
WPTTPAFSLYAKEDTVLGGEYPLEKGDEVMVLIPQLHRDKTVWGDDVEEFRPER
FENPSAIPQHAFKPFGNGQRSALGQQFALHEATLVLGMMLKHFDFEDHTNYELD
IKETLTLKPKGFVVKAKSKKIPLGGIPSPSTEQSAKKVRKKAENAHNTPLLVLYGS
NMGTAEGTARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNGHPP
DNAKQFVDWLDQASADEVKGVRYSVFGCGDKDWATTYQKVPAFIDETLAAKG
AENIADRGEADASDDFEGTYEEWREHMWSDVAAYFNLDIENSEDNKPTLSLQFV
DSAADMPLAKMHGAFSTNVVASKELQQPGSARSTRHLEIELPKEASYQEGDHLG
VIPRTMKE

*4.2.9. uAmD8** DNA Sequence:

ACAATTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATTTAC
CGTTATTAAACACAGATAAACCGGTTCAAGCTTTGATGAAAATTGCGGATGA
ATTAGGAGAAATCTTTAAATTCGAGGCGCCTGGTCGTGTAACGCGCTACTTAT
CAAGTCAGCGTCTAATTAAAGAAGCATGCGATGAATCACGCTTTGATAAAAG
TTTAAGTCAAATGCTGAAATTTCTGCGTGATTTTCTTGGAGACGGGTTAGCCA
CAAGCTGGACGCATGAAAAAAATTGGAAAAAAGCGCATAATATCTTACTTCC
AAGCTTTAGTCAGCAGGCAATGAAAGGCTATCATGCGATGATGGTCGATATC
GCCGTGCAGCTTGTTCAAAAGTGGGAGCGTCTAAATGCAGATGAGCATATTG
AAGTATCGGAAGACATGACACGTTTAACGCTTGATACAATTGGTCTTTGCGG
CTTTAACTATCGCTTTAACAGCTTTTACCGAGATCAGCCTCATCCATTTATTAT
AAGTATGGTCCGTGCACTGGATGAAGTAATGAACAAGCTGCAGCGAGCAAAT
CCAGACGACCCAGCTTATGATGAAAACAAGCGCCAGTTTCAAGAAGATATCA
AGGTGGTGAACGACCTAGTAGATAAAATTATTGCAGATCGCAAAGCAGGTGG
TGAACAAAGCGATGATTTATTAACGCAGATGCTAAACGGAAAAGATCCAGAA
ACGGGTGAGCCGCTTGATGACGTGAACATTCGCTATCAAATTATTACATTCTT
AGTGGCGGGACACGAAGGTACAAGTGGTCTTTTATCATTTGCGCTGTATTTCT
TAGTGAAAAATCCACATGTATTACAAAAAGTAGCAGAAGAAGCAGCACGAG
TTCTAGTAGATCCTGTTCCAAGCTACAAACAAGTCAAACAGCTTAAATATGTC
GGCATGGTCTTAAACGAAGCGCTGCGCTTATGGCCAACTACGCCTGCGTTTTC
CCTATATGCAAAAGAAGATACGGTGCTTGGAGGAGAATATCCTTTAGAAAAA
GGCGACGAAGTAATGGTTCTGATTCCTCAGCTTCACCGTGATAAAACAGTCT
GGGGAGACGATGTGGAGGAGTTCCGTCCAGAGCGTTTTGAAAATCCAAGTGC
GATTCCGCAGCATGCGTTTAAACCGTTTGGAAACGGTCAGCGTGCGGCGCTT
GGTCAGCAGTTCGCTCTTCATGAAGCAACGCTGGTACTTGGTATGATGCTAAA
ACACTTTGACTTTGAAGATCATACAAACTACGAGCTCGATATTAAAGAAACT
TTAGGTTTAAAACCTAAAGGCTTTGTGGTAAAAGCAAAATCGAAAAAAATTC
CGCTTGGCGGTATTCCTTCACCTAGCACTGAACAGTCTGCTAAAAAAGTACGC
AAAAAGGCAGAAAACGCTCATAATACGCCGCTGCTTGTGCTATACGGTTCAA
ATATGGGTACCGCTGAAGGAACGGCGCGTGATTTAGCAGATATTGCAATGAG
CAAAGGATTTGCACCGCAGGTCGCAACGCTTGATTCACACGCCGGAAATCTT
CCGCGCGAAGGAGCTGTATTAATTGTAACGGCGTCTTATAACGGTCATCCGC
CTGATAACGCAAAGCAATTTGTCGACTGGTTAGACCAAGCGTCTGCTGATGA

AGTAAAAGGCGTTCGCTACTCCGTATTTGGATGCGGCGATAAAGACTGGGCT
ACTACGTATCAAAAAGTGCCTGCTTTTATCGATGAAACGCTTGCCGCTAAAG
GGGCAGAAAACATCGCTGACCGCGGTGAAGCAGATGCAAGCGACGACTTTG
AAGGCACATATGAAGAATGGCGTGAACATATGTGGAGTGACGTAGCAGCCTA
CTTTAACCTCGACATTGAAAACAGTGAAGATAATAAACCTACTCTTTCACTTC
AATTTGTCGACAGCGCCGCGGATATGCCGCTTGCGAAAATGCACGGTGCGTT
TTCAACGAACGTCGTAGCAAGCAAAGAACTTCAGCAGCCAGGCAGTGCACG
AAGCACGCGACATCTTGAAATTGAACTTCCAAAAGAAGCTTCTTATCAAGAA
GGAGATCATTTAGGTGTTATTCCTCGCACTATGAAGGAATAG

*4.2.10. uAmD8** Protein Sequence:
TIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQ

RLIKEACDESRFDKSLSQMLKFLRDFLGDGLATSWTHEKNWKKAHNILLPSFSQ

QAMKGYHAMMVDIAVQLVQKWERLNADEHIEVSEDMTRLTLDTIGLCGFNYRF

NSFYRDQPHPFIISMVRALDEVMNKLQRANPDDPAYDENKRQFQEDIKVVNDLV

DKIIADRKAGGEQSDDLLTQMLNGKDPETGEPLDDVNIRYQIITFLVAGHEGTSG

LLSFALYFLVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRL

WPTTPAFSLYAKEDTVLGGEYPLEKGDEVMVLIPQLHRDKTVWGDDVEEFRPER

FENPSAIPQHAFKPFGNGQRAALGQQFALHEATLVLGMMLKHFDFEDHTNYELD

IKETLGLKPKGFVVKAKSKKIPLGGIPSPSTEQSAKKVRKKAENAHNTPLLVLYG

SNMGTAEGTARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNGHPP

DNAKQFVDWLDQASADEVKGVRYSVFGCGDKDWATTYQKVPAFIDETLAAKG

AENIADRGEADASDDFEGTYEEWREHMWSDVAAYFNLDIENSEDNKPTLSLQFV

DSAADMPLAKMHGAFSTNVVASKELQQPGSARSTRHLEIELPKEASYQEGDHLG

VIPRTMKE

B.4.3 Product Quantitation and TTNs

B.4.3.1 Analytical Scale Reaction Setup

Single colonies from *E. coli* BL21(DE3) cells, transformed with plasmids containing P411 variants were picked from LB_{amp} agar plates using sterile toothpicks and grown in 5 mL of LB_{amp} in 10-mL sterile tubes at 37 °C (at 220 rpm and 80% humidity) for 12–18 hours. The preculture (500 µL) was used to inoculate 50 mL of HB_{amp} medium in 125-mL conical flasks. The expression culture plate was then incubated at 37 °C (at 220 rpm and 80% humidity) for 2.5 hours (OD_{600} = 80–100) and then chilled on ice for 30 minutes. After that, 50 mL IPTG (0.5 mM final concentration) and 50 mL ALA (1.0 mM final concentration) were added, and the proteins were expressed at 22 °C (at 220 rpm) for 20–24 h. Cells were then pelleted at 20 °C and 4,000 g for 5 min. The cells were resuspended with M9-N buffer (pH 8.2) containing 20 mM glucose (OD_{600} = 30). In 2-mL screw cap vials, 780 mL of the resuspended cells were charged and set up in triplicates. A stock solution (20 mL) of the mixture of hydrocarbon (200 mM dissolved in ethanol) and *N*-(pivaloyloxy)acetamide (400 mM dissolved in ethanol) were added to each of the vials. The vials were shaken at 700 rpm under ambient conditions for 12–16 hours. The remainder of the cell suspension was used for protein quantitation with the hemechrome assay (section 1.3).

Work-up for LC-MS Analysis: Ethanol (800 mL/ vial) was added, the resulting suspension was mixed by vortexing and then centrifuged (14,000 *g* for 10 min) to remove the cell debris. The supernatant (400 mL) was transferred into a 2-mL screw cap vial for LC-MS

analysis. Product formation was quantified by LC-MS based on the calibration curve of the corresponding racemic reference compound.

Work-up for GC-FID Analysis: To the reaction mixture was added 1:1 EtOAc/Hexane (350 mL/ vial) containing bibenzyl (0.5 mM dissolved in 1:1 EtOAc/Hexane), the resulting suspension was mixed by vortexing, and then centrifuged (14,000 *g* for 10 min) to remove the cell debris. The supernatant (200 mL) was transferred into a 2-mL screw cap vial with an insert for GC-FID analysis. Product formation was quantified by GC-FID based on the calibration curve of the corresponding racemic reference compound.

B.4.3.2 Calibration Curves for Standard Products and Reaction Product Quantitation

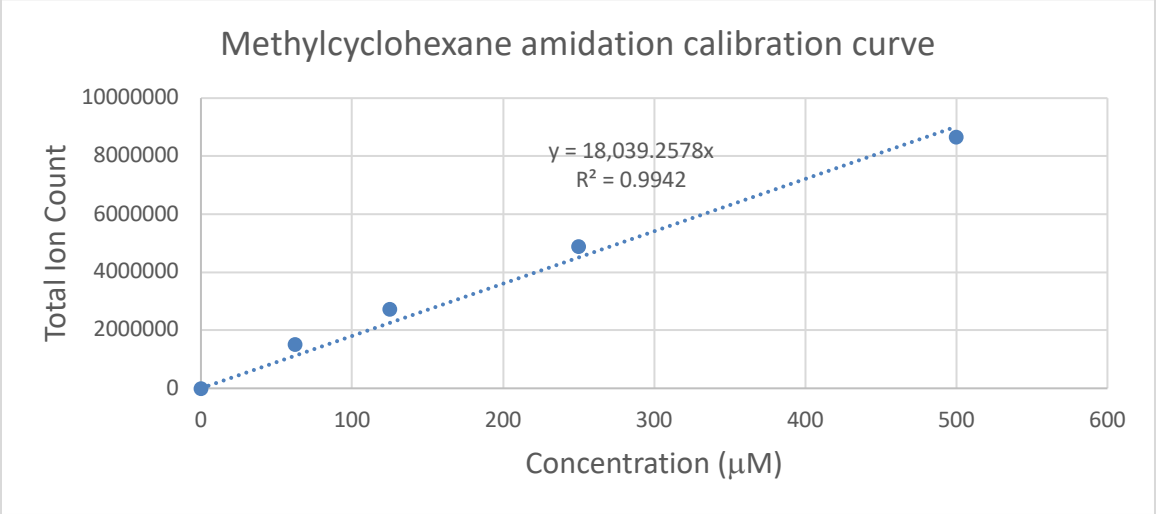
Enzymatic reactions on analytic scale were performed following the general procedure described in section 4.3.1. Product formation was quantified by HPLC-MS or GC-FID based on the calibration curve of the corresponding reference compound. TTN is defined as the concentration of product divided by the concentration of heme protein measured by the hemochrome assay as described in section 1.3. Calibration curves of commercially available or synthesized reference compounds were created for the determination of yield and TTN. Harvested cells were resuspended with M9-N buffer (pH 8.2) containing glucose (20 mM) to OD_{600} = 30. Then, 780 mL of the resuspended cells were aliquoted to 2-mL screw cap vials. Product standard stocks (20 μL, 400–50 mM stock in ethanol) were added. The mixture was vortexed and then analyzed.

For LC-MS: The ion counts at selected ion monitoring mode of product were measured. For all the analyses, water and acetonitrile containing 0.1% acetic acid were used as eluents for a Kromasil® C18 column. The methods used 30–95% acetonitrile (0.3–2.0 minutes) and 95–30% (2.0–2.5 minutes) with 3-μL sample injections. The flow rate was 1.5 mL/minute, and the column was maintained at 22 °C.

For GC-FID: The total area of the racemic product peak was measured using chiral GC-FID. For all analysis, CP-Chirasil-Dex-CP (Agilent) was used as the chiral stationary phase. The method used was: start at 50 °C, ramp at 10 °C/minute to 130 °C and hold for 22 minutes, finally ramp at 35 °C/minute to 200 °C and hold for 10 minutes, using 5-μL sample injections.

B.4.3.3 Yields and TTN for Methylcyclohexane Reaction across *uAmD* Lineage (Figure 3-2B)

Based on total ion count in 156 m/z single ion mode for standard product (1-*N*-acetylamino-2-methylcyclohexane), obtained from LCMS.



Variant	Pdt TIC	[Pdt] / mM	Avg. [Pdt] / mM	SD [Pdt] / mM	[PC] / mM	TTN	Avg TTN	SD TTN
uAmD0	875994.1	48.5	50.7	2.6	4.91	9.9	9.9	0.6
	903386.3	50.1			4.91	10.2		
	931676.3	51.6			4.91	10.5		
	936101.1	51.9			4.91	10.6		
	893523.4	49.5			4.91	10.1		
	900372.9	49.9			4.91	10.2		
	690565.6	47.5			5.39	8.8		
	747106.6	51.4			5.39	9.5		
	759974.6	52.3			5.39	9.7		
	814611.3	56.0			5.39	10.4		
	675896.1	46.5			5.39	8.6		
	777515.6	53.5			5.39	9.9		
	1166537.8	64.7			3.33	19.4		
	1199392.5	66.5			3.33	19.9		
	1077291.6	59.7			3.33	17.9		

	1062266.6	58.9			3.33	17.7		
	1227549.8	68.0			3.33	20.4		
uAmD1	1182044.4	65.5	60.3	7.4	3.33	19.7	18.4	2.1
	896748.2	61.7			3.20	19.2		
	931868.3	64.1			3.20	19.9		
	622940.2	42.8			3.20	13.4		
	709330.3	48.9			3.20	15.2		
	911749.1	62.7			3.20	19.5		
	867860.6	59.7			3.20	18.6		
	1385376	76.8			2.70	28.4		
	1449265.5	80.3			2.70	29.7		
	1464323.4	81.2			2.70	30.0		
	1498926.3	83.1			2.70	30.7		
	1369136.4	75.9			2.70	28.1		
uAmD2	1396578.9	77.4	75.8	6.4	2.70	28.6	25.5	4.3
	1142318.8	78.6			3.34	23.5		
	1057241	72.7			3.34	21.7		
	863492.4	59.4			3.34	17.8		
	1004425.1	69.1			3.34	20.6		
	1111092.5	76.4			3.34	22.8		
	1144866.3	78.7			3.34	23.5		
	1499272.4	83.1			3.50	23.7		
	1512753.5	83.9			3.50	23.9		
	1501109.5	83.2			3.50	23.7		
	1564313.6	86.7			3.50	24.7		
	1570826.8	87.1			3.50	24.9		
uAmD3	1526438.1	84.6	86.8	3.4	3.50	24.2	28.3	4.4
	1332466.8	91.6			2.74	33.4		
	1313349.4	90.3			2.74	32.9		
	1219479.4	83.9			2.74	30.6		
	1307359.6	89.9			2.74	32.8		
	1237441.5	85.1			2.74	31.0		
	1344169.4	92.5			2.74	33.7		
	1560261.1	86.5			3.29	26.3		
	1584269.6	87.8			3.29	26.7		
	1758217	97.5			3.29	29.6		
	1773558.6	98.3			3.29	29.9		

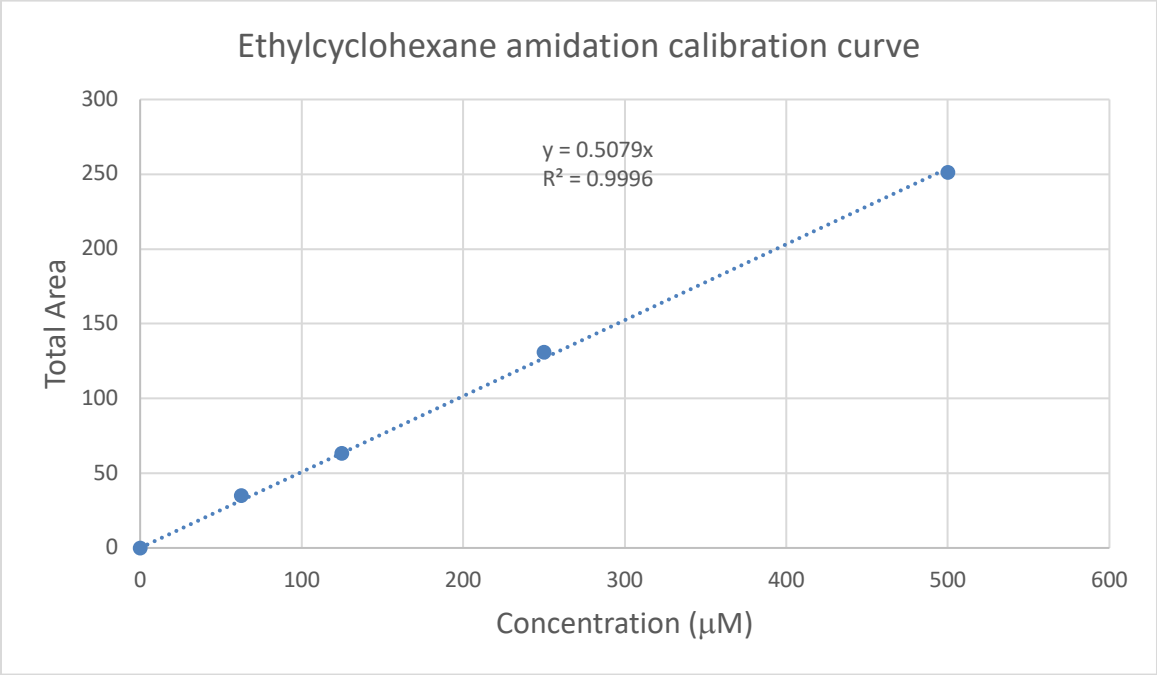
	1703646.4	94.4			3.29	28.7		
uAmD4	1696347.5	94.0	94.7	4.4	3.29	28.6	34.5	6.6
	1370079.9	94.2			2.36	39.8		
	1413068.1	97.2			2.36	41.0		
	1341893.3	92.3			2.36	38.9		
	1483566.8	102.0			2.36	43.1		
	1366913.4	94.0			2.36	39.7		
	1423752.3	97.9			2.36	41.3		
	2272560.8	125.9			4.79	26.3		
	2279263.5	126.3			4.79	26.4		
	2414186.8	133.8			4.79	27.9		
uAmD5	2376446.8	131.7			4.79	27.5		
	2342605	129.9			4.79	27.1		
	2345873.8	130.0	137.1	8.9	4.79	27.1	27.8	1.1
	2080576	143.1			5.08	28.2		
	1983013.4	136.4			5.08	26.8		
	2202116.3	151.5			5.08	29.8		
	2188682.5	150.5			5.08	29.6		
	2085827.5	143.5			5.08	28.2		
	2073632.9	142.6			5.08	28.1		
	3826004.3	212.1			3.87	54.7		
	3932513.3	217.9			3.87	56.3		
	4043732.3	224.2			3.87	57.8		
	4000667.8	221.8			3.87	57.2		
	4216126	233.7			3.87	60.3		
uAmD6	3996289.3	221.5	227.5	8.0	3.87	57.2	53.5	4.1
	3381888.3	232.6			4.67	49.7		
	3463729.8	238.2			4.67	50.9		
	3366713.5	231.6			4.67	49.5		
	3436943.3	236.4			4.67	50.5		
	3306902.5	227.5			4.67	48.6		
	3380968.5	232.5			4.67	49.7		
	6424544	356.1			2.89	123.1		
	6365837.5	352.9			2.89	121.9		
	6668496.5	369.7			2.89	127.7		
	6452037.5	357.7			2.89	123.6		
	6574276.5	364.4			2.89	125.9		

uAmD7	6426624	356.2	369.6	16.6	2.89	123.1	98.9	26.5
	5513669.5	379.2			5.15	73.6		
	5233626	359.9			5.15	69.8		
	5706874	392.5			5.15	76.1		
uAMD7	5219220.5	358.9			5.15	69.6		
	5874961.5	404.1			5.15	78.4		
	5582787	383.9			5.15	74.5		
uAmD8	7243449	401.5	411.6	14.8	3.71	108.2	102.3	8.4
	7358974	407.9			3.71	109.9		
	7431819	411.9			3.71	111.0		
	7360940.5	408.0			3.71	109.9		
	7383612.5	409.3			3.71	110.3		
	7283403	403.7			3.71	108.8		
	6086471	418.6			4.38	95.5		
	6397953.5	440.1			4.38	100.3		
	5815946	400.0			4.38	91.2		
	5758805.5	396.1			4.38	90.3		
	6412815	441.1			4.38	100.6		
	5833333	401.2			4.38	91.5		
uAmD9	7991622	442.9	452.9	15.3	3.46	127.8	121.6	8.3
	8047741.5	446.1			3.46	128.7		
	8095354	448.7			3.46	129.4		
	8089430.5	448.4			3.46	129.4		
	8115218	449.8			3.46	129.8		
	8002133	443.6			3.46	127.9		
	6572102	452.0			4.01	112.5		
	6420101.5	441.6			4.01	109.9		
	6555119	450.9			4.01	112.3		
	7074436.5	486.6			4.01	121.1		
	6427025.5	442.1			4.01	110.0		
	7021783	482.9			4.01	120.2		

Notes: Pdt TIC = product total ion count; [Pdt] = product concentration in reaction; Avg. [Pdt] = average product concentration; SD [Pdt] = standard deviation of product concentration; [PC] = protein concentration in reaction; TTN = total turnover number; Avg. TTN = average total turnover number; SD TTN = standard deviation of total turnover number.

B.4.3.4 Yields and TTNs for Substrates in Figure 3-3A

*Ethylcyclohexane (**10**)*
Based on total area for standard product (1-*N*-acetylamino-2-ethylcyclohexane), obtained from GC-FID.

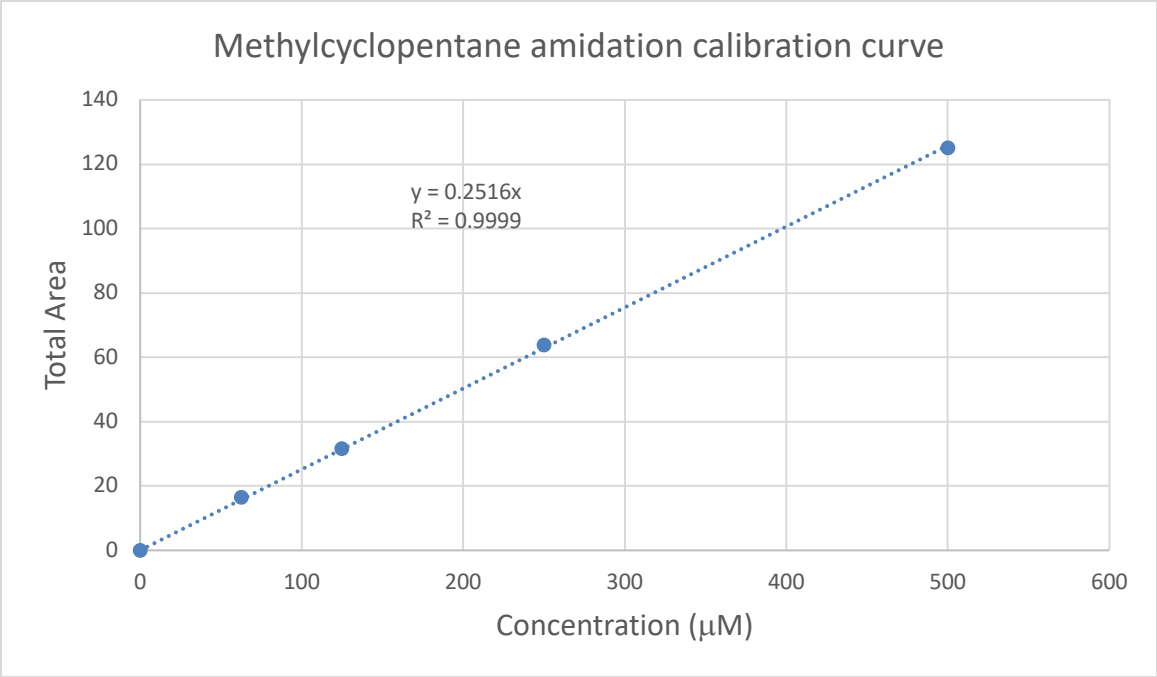


Variant	Pdt Area	[Pdt] / mM	Avg. [Pdt] / mM	SD [Pdt] / mM	[PC] / mM	TTN	Avg TTN	SD TTN
uAmD9	614.6	1214.1	1218.8	6.6	6.25	194.2	194.9	1.0
	619.3	1223.4			6.25	195.7		

Notes: Pdt Area = total product area; [Pdt] = product concentration in reaction; Avg. [Pdt] = average product concentration; SD [Pdt] = standard deviation of product concentration; [PC] = protein concentration in reaction; TTN = total turnover number; Avg. TTN = average total turnover number; SD TTN = standard deviation of total turnover number.

*Methylcyclopentane (**12**)*

Based on total area for standard product (1-*N*-acetylamino-2-methylcyclopentane), obtained from GC-FID.

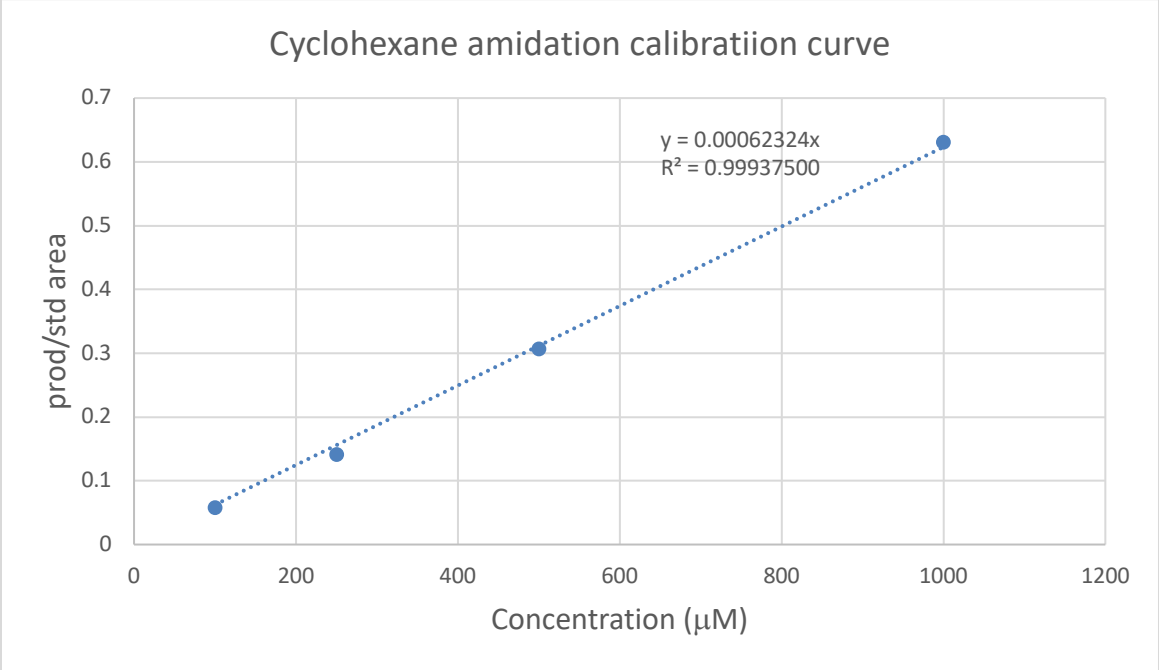


Variant	Pdt Area	[Pdt] / mM	Avg. [Pdt] / mM	SD [Pdt] / mM	[PC] / mM	TTN	Avg TTN	SD TTN
uAmD6	204	831.6	823.7	22.7	7.08	117.4	106.5	12.9
	193.8	790.1			8.53	92.6		
	204.5	833.7			7.08	117.7		
	205.9	839.4			8.53	98.4		

Notes: Pdt Area = total product area; [Pdt] = product concentration in reaction; Avg. [Pdt] = average product concentration; SD [Pdt} = standard deviation of product concentration; [PC] = protein concentration in reaction; TTN = total turnover number; Avg. TTN = average total turnover number; SD TTN = standard deviation of total turnover number.

*Cyclohexane (**14**)*

Based on GC-FID areas of standard product (*N*-acetylaminocyclohexane) and internal standard (bibenzyl)



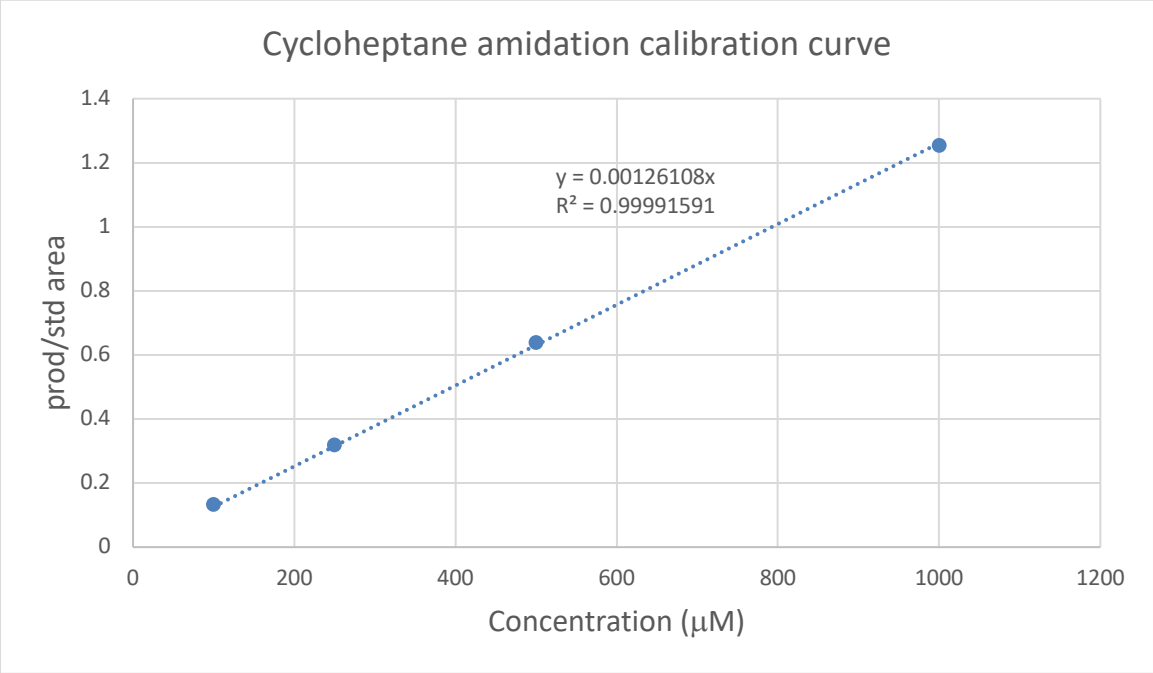
Variant	Pdt	IS	Pdt/IS	Pdt [uM]	PC [uM]	TTN	Avg TTN	SD TTN
	15.45	243.00	0.06	101.99	6.10	16.72		
*uAMD8**[7]	17.60	238.91	0.07	118.21	6.40	18.47	17.60	1.23

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,

[7] Here the '*' is part of the variants name. uAmD8* is between uAmD8 and uAmD9. During evolution from uAmD8 to uAmD9, we found a variant that had higher activity but lower selectivity (i.e. uAmD8*). We decided to use this variant to test substrate scope.

*Cycloheptane (**16**)*

Based on GC-FID areas of standard product (*N*-acetylaminocycloheptane) and internal standard (bibenzyl)



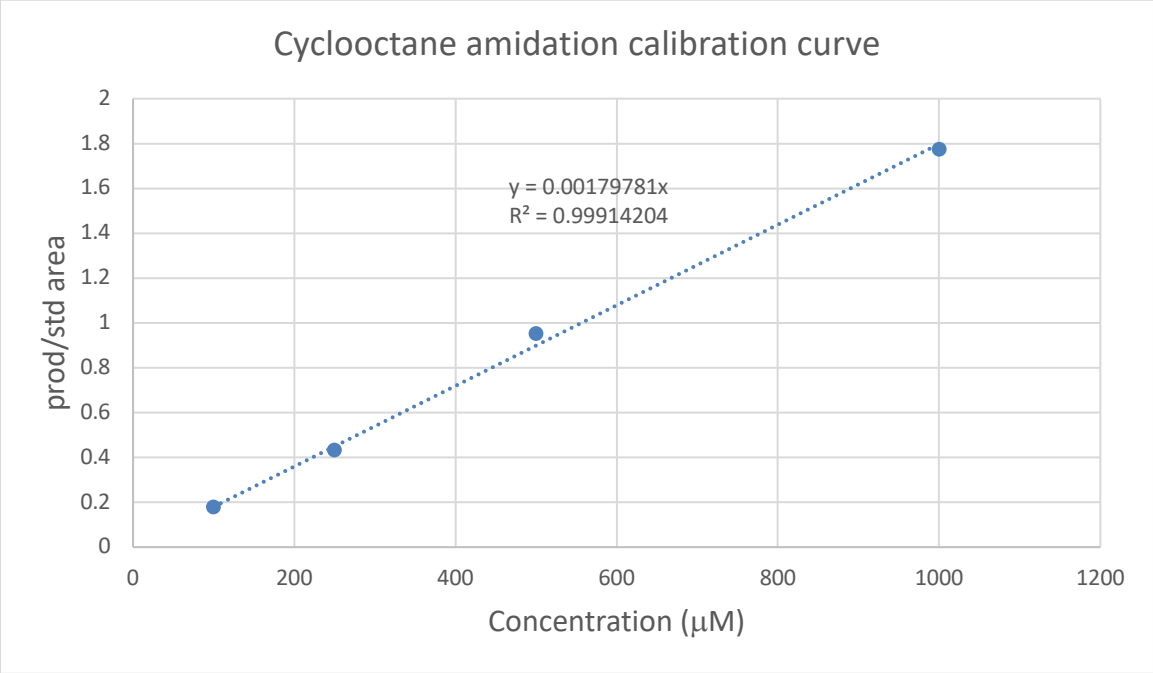
Variant	Pdt	IS	Pdt/IS	Pdt [uM]	PC [uM]	TTN	Avg TTN	SD TTN
uAMD8[8]*	43.00	241.08	0.18	141.45	6.20	22.81	24.10	1.81
	51.00	245.17	0.21	164.97	6.50	25.38		

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,

[8] Here the '*' is part of the variants name. uAmD8* is between uAmD8 and uAmD9. During evolution from uAmD8 to uAmD9, we found a variant that had higher activity but lower selectivity (i.e. uAmD8*). We decided to use this variant to test substrate scope.

*Cyclooctane (**18**)*

Based on GC-FID areas of standard product (*N*-acetylaminocyclooctane) and internal standard (bibenzyl)



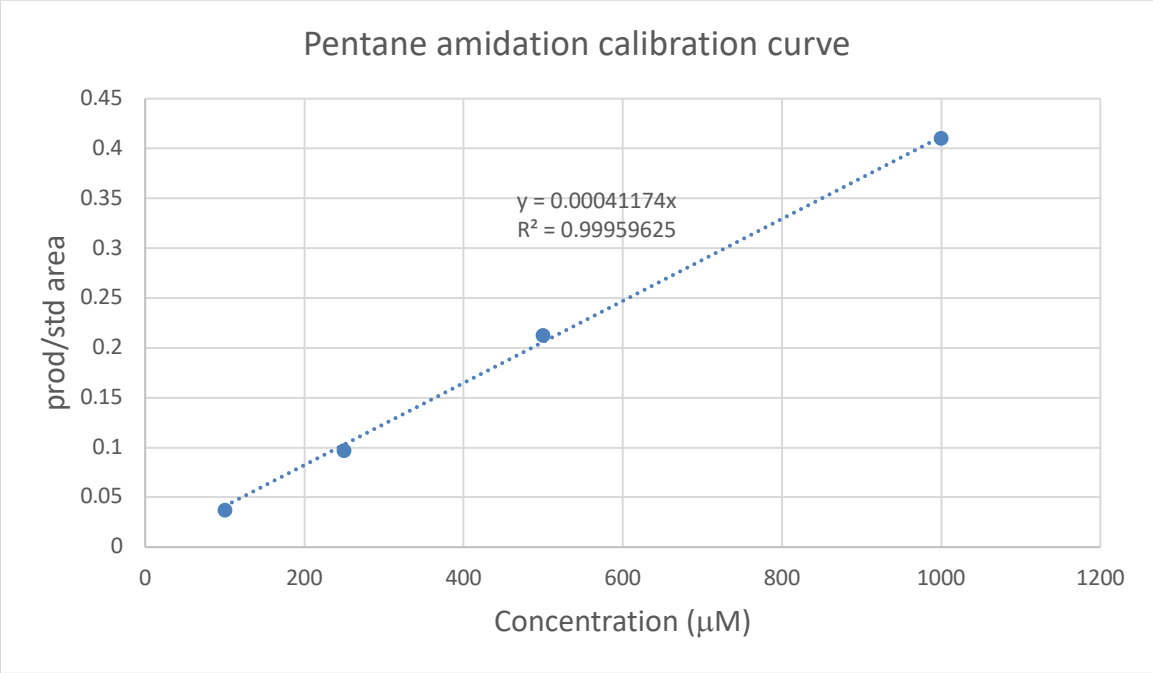
Variant	Pdt	IS	Pdt/IS	Pdt [uM]	PC [uM]	TTN	Avg TTN	SD TTN
	24.71	237.61	0.10	57.87	6.20	9.33		
uAMD8[9]*	28.86	240.35	0.12	66.81	6.50	10.28	9.81	0.67

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,

[9] Here the '*' is part of the variants name. uAmD8* is between uAmD8 and uAmD9. During evolution from uAmD8 to uAmD9, we found a variant that had higher activity but lower selectivity (i.e. uAmD8*). We decided to use this variant to test substrate scope.

*Pentane (**19**)*

Based on total GC-FID areas of standard product (3-*N*-acetylaminopentane) and internal standard (bibenzyl)



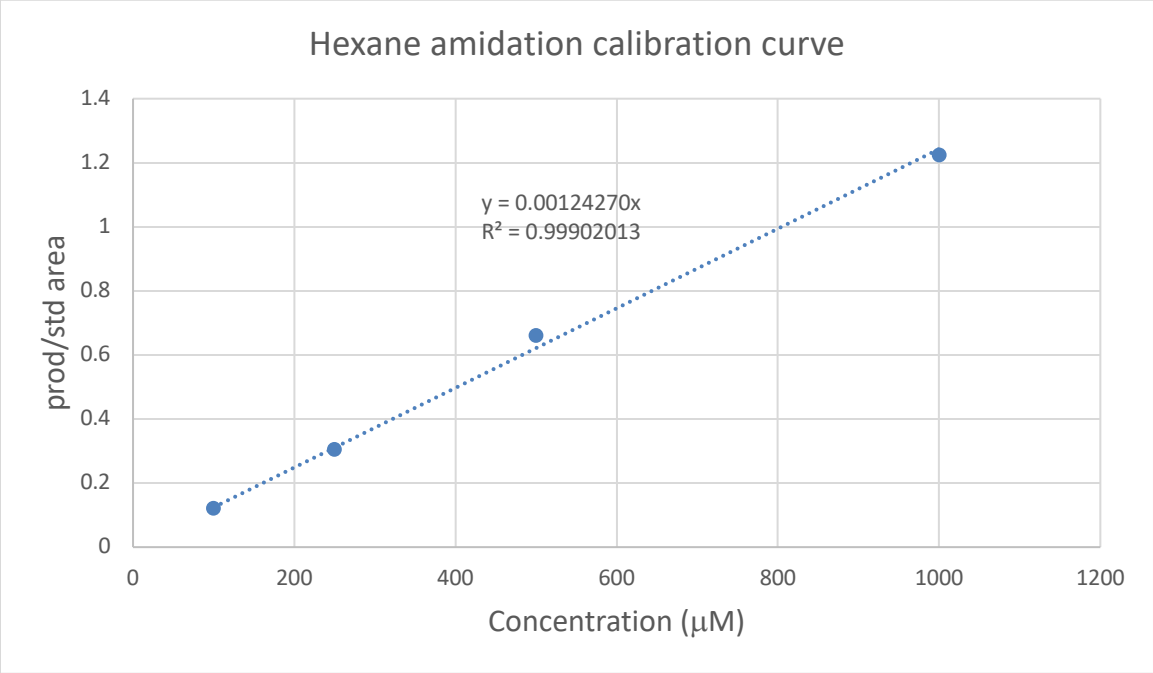
Variant	Pdt	IS	Pdt/IS	Pdt [uM]	PC [uM]	TTN	Avg TTN	SD TTN
uAMD8[1]0*	16.63	211.42	0.08	191.10	6.20	30.82	33.09	2.27
	20.10	212.46	0.09	229.84	6.50	35.36		

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,

[10] Here the '*' is part of the variants name. uAmD8* is between uAmD8 and uAmD9. During evolution from uAmD8 to uAmD9, we found a variant that had higher activity but lower selectivity (i.e. uAmD8*). We decided to use this variant to test substrate scope.

*Hexane (**20**)*

Based on total GC-FID areas of standard product (3-*N*-acetylaminohexane) and internal standard (bibenzyl)

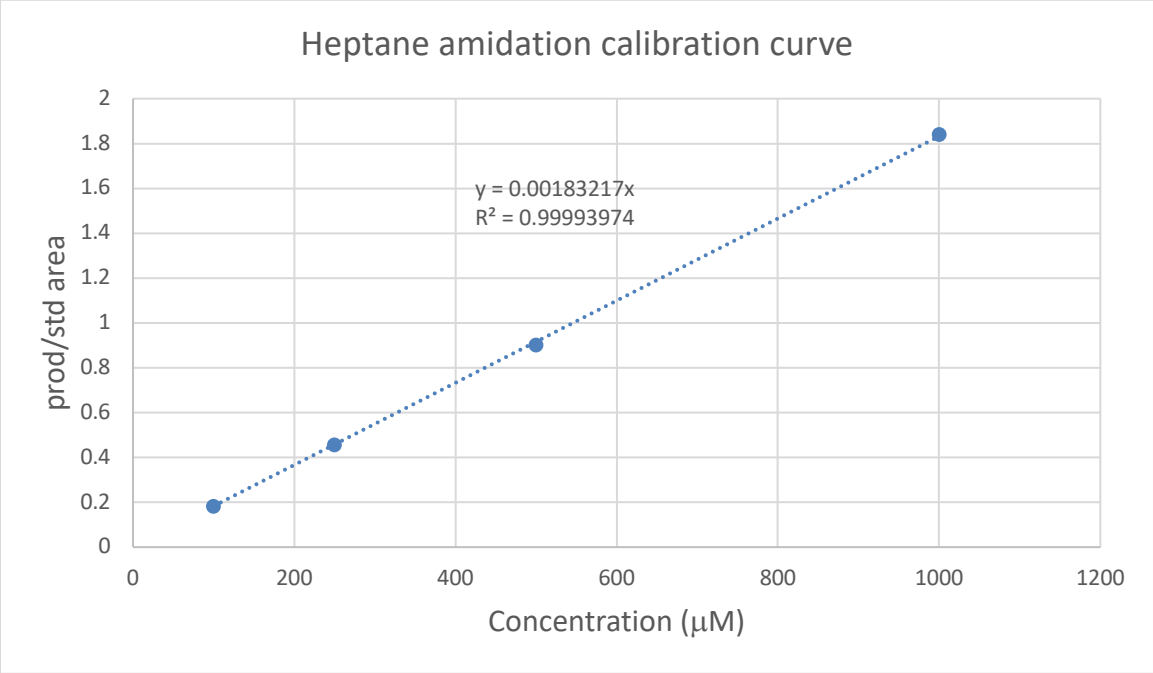


Variant	Pdt	IS	Pdt/IS	Pdt [uM]	PC [uM]	TTN	Avg TTN	SD TTN
	60.86	211.84	0.29	231.31	6.20	37.31		
uAMD6	67.21	215.00	0.31	251.70	6.50	38.72	38.02	0.71

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,

*Heptane (**21**)*

Based on total GC-FID areas of standard product (2-*N*-acetylaminoheptane) and internal standard (bibenzyl)

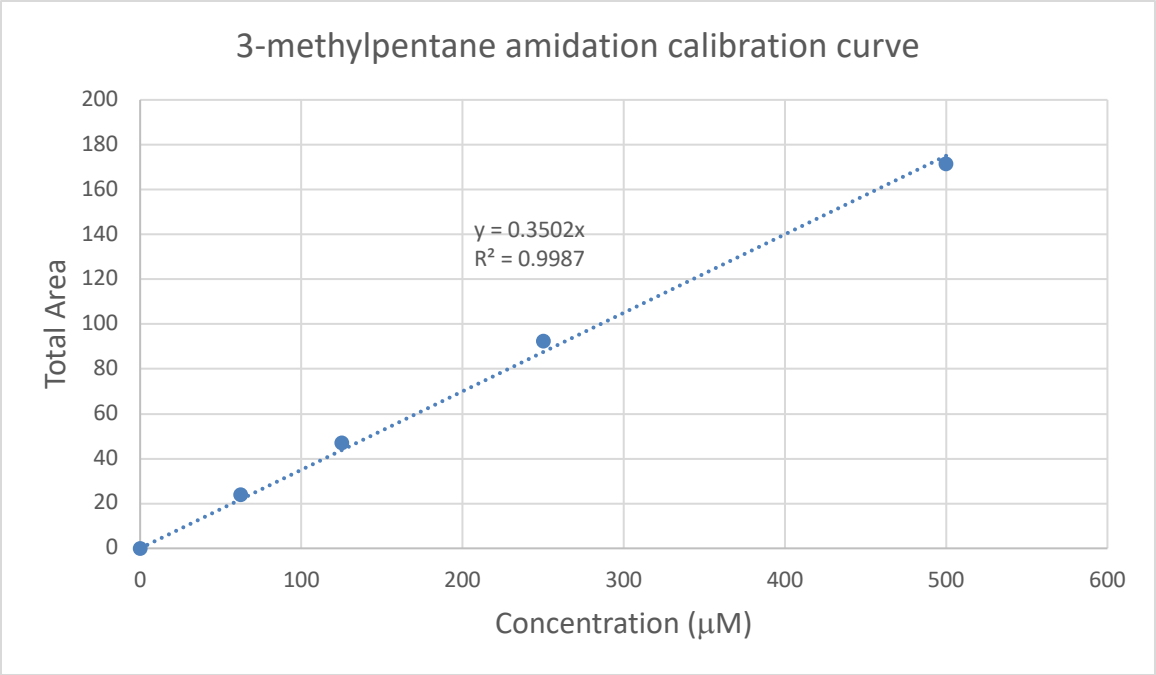


Variant	Pdt	IS	Pdt/IS	Pdt [uM]	PC [uM]	TTN	Avg TTN	SD TTN
uAMD7-T438S	99.23	216.00	0.46	250.77	6.20	40.45	44.44	3.99
	126.31	219.00	0.58	314.82	6.50	48.43		

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

*3-methylpentane (**22**)*

Based on total area for standard product (3-*N*-acetylamino-3-methylpentane), obtained from GC-FID.



Variant	Pdt Area	[Pdt] / mM	Avg. [Pdt] / mM	SD [Pdt] / mM	[PC] / mM	TTN	Avg TTN	SD TTN
uAmD6	161.9	472.0	485.1	13.9	7.12	66.3	67.1	0.8
	171.3	499.4			7.33	68.1		
	162.7	474.3			7.12	66.6		
	169.7	494.8			7.33	67.5		

Notes: Pdt Area = total product area; [Pdt] = product concentration in reaction; Avg. [Pdt] = average product concentration; SD [Pdt} = standard deviation of product concentration; [PC] = protein concentration in reaction; TTN = total turnover number; Avg. TTN = average total turnover number; SD TTN = standard deviation of total turnover number.

B.5. Synthesis and stereochemical assignment of enantiomerically-enriched amide standards

Commercial amines (Enamine) available as a mixture of diastereomers were acetylated with *Candida antarctica lipase B (CalB lipase acrylic resin, Sigma),* leading to the selective acetylation of one of the four enantiomers. The product amide was recrystallized in hexanes/ethyl acetate to provide crystals suitable for X-ray diffraction, which allowed establishment of absolute configuration of the CalB derived amides.

General Procedure for CalB acetylation (adapted from literature)[8]:

In a 10-mL round-bottom flask equipped with a stir bar, 1 mmol of the diastereomeric amines, 2 mmol of ethyl acetate, and 80 mg of CalB were dissolved in 4 mL of anhydrous toluene. The reaction mixture was stirred at 200 rpm at 50 ºC for 12–24 hrs. The reaction mixture was then diluted with water and ethyl acetate. The organic layer was separated, and the aqueous layer was extracted with ethyl acetate (15 mL X 2). The combined organic layers were dried over sodium sulfate, concentrated, and purified by column chromatography (EtOAc) to give the enantioenriched amide in 20–25% yield as a white solid. This material was recrystallized by vapor diffusion (EtOAc/Hexanes) to give crystals suitable for X-ray diffraction.

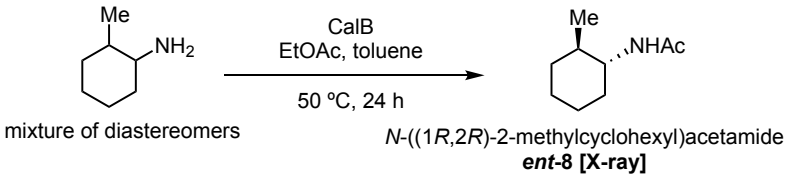
Upon assignment of the CalB derived enantiomer, the GC peaks for the diastereomeric amides can be partially assigned. Thus, the absolute configuration of the *uPA* or *uAMD* derived products can be assigned.

Chiral GC-FID Methods used to resolve the isomeric amides were:

Method 1: start at 50 °C, ramp at 10 °C/minute to 130 °C and hold for 22 minutes. Ramp at 70 °C/minute to 200 °C and hold for 5 minutes.

Method 2: start at 130 °C, ramp at 5 °C/minute to 200 °C and hold for 5 minutes.

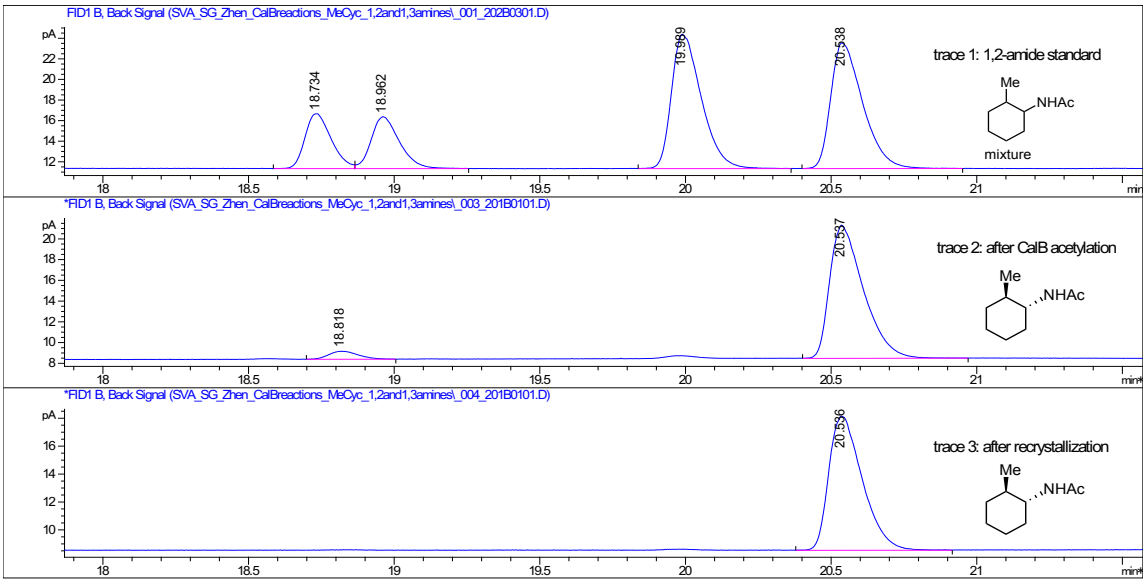
B.5.1. Acetylation of 2-methyl-aminocylohexane



mixture of diastereomers

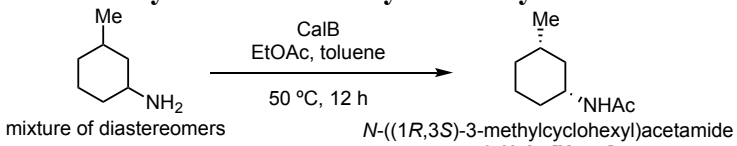
N-((1*R*,2*R*)-2-methylcyclohexyl)acetamide
***ent*-8 [X-ray]**

Chiral GC-FID Chromatogram for standard amides and CalB derived product (Method 1)



For the four diastereomeric peaks of the standard (trace 1), the following assignments can be made:
peak 1 and peak 2 (18.73, 18.96 min) – *cis* isomers.
peak 3 (19.98 min) – (*S,S*) - trans isomer
peak 4 (20.53 min) – (*R,R*) - trans isomer; CalB derived product, X-ray structure
Note that the CalB derived isomer (**ent-8**) is the enantiomer of the product obtained from *uAMD* amidation (**8**).

B.5.2. Acetylation of 3-methyl-aminocylohexane

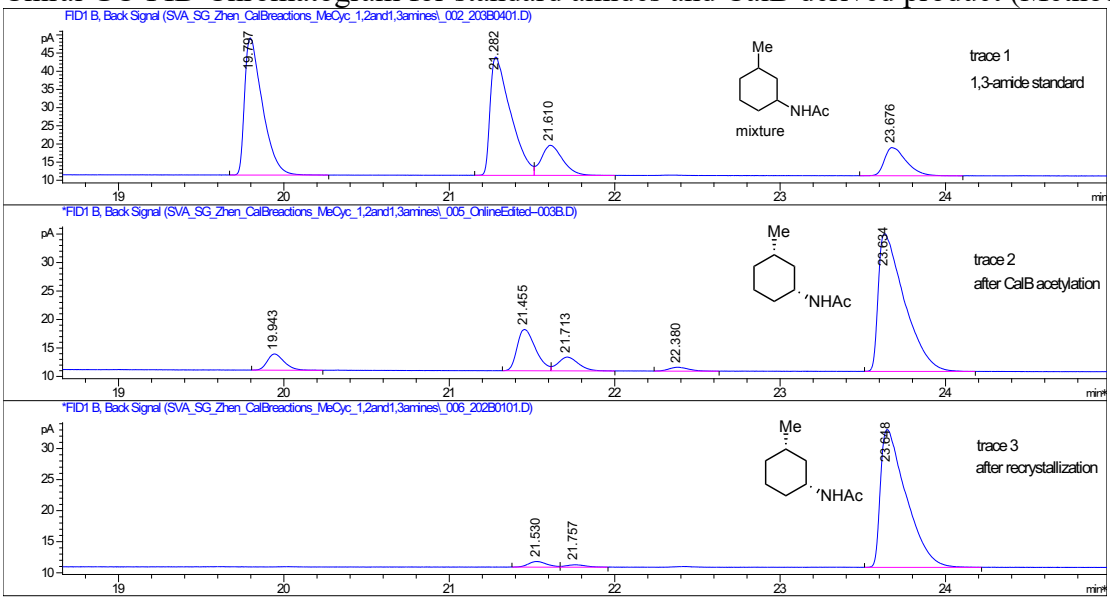


mixture of diastereomers

N-((1*R*,3*S*)-3-methylcyclohexyl)acetamide

6-*N-Ac* [X-ray]

Chiral GC-FID Chromatogram for standard amides and CalB derived product (Method 1)



For the four diastereomeric peaks of the standard (trace 1), the following assignments can be made:

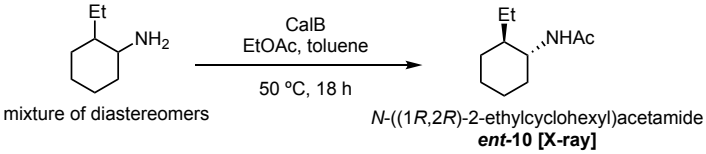
peak 1 and peak 2 (19.79, 21.28 min) – *trans* isomers.

peak 3 (21.61 min) – (*S,R*) - *cis* isomer

peak 4 (23.67 min) – (*R,S*) - *cis* isomer; CalB derived product, X-ray structure

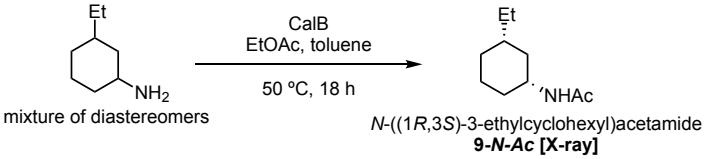
Note that the CalB derived isomer (**6-*N-Ac***) matches with the enantiomer of the amide derived from *uPA* amination (**6**).

B.5.3. Acetylation of 2-ethyl-aminocylohexane and 3-ethyl-aminocylohexane



mixture of diastereomers

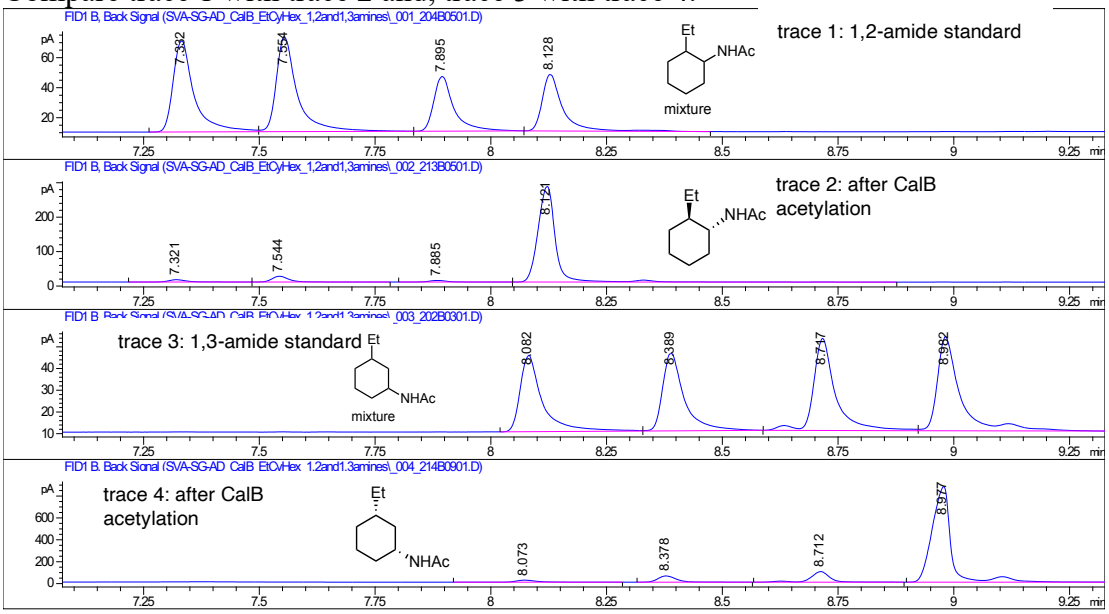
N-((1*R*,2*R*)-2-ethylcyclohexyl)acetamide
***ent*-10 [X-ray]**



mixture of diastereomers

N-((1*R*,3*S*)-3-ethylcyclohexyl)acetamide
9-*N-Ac* [X-ray]

Chiral GC-FID Chromatogram for standard amides and CalB derived product (Method 2). Compare trace 1 with trace 2 and, trace 3 with trace 4.



For the four diastereomeric peaks of the 1,2-amide standard (trace 1), the following assignments can be made:
peak 1 and peak 2 (7.33, 7.55 min) – *cis* isomers.
peak 3 (7.89 min) – (*S,R*) - *trans* isomer
peak 4 (8.12 min) – (*R,S*) - *trans* isomer; CalB derived product, X-ray structure
For the four diastereomeric peaks of the 1,3-amide standard (trace 3), the following assignments can be made:
peak 1 and peak 2 (8.08, 8.38 min) – *trans* isomers.
peak 3 (8.71 min) – (*S,R*) - *cis* isomer
peak 4 (8.97 min) – (*R,S*) - *cis* isomer; CalB derived product, X-ray structure
Note that the CalB derived isomer (**ent-10**) is the enantiomer of the product obtained from *uAMD* amidation (**10**).
Note that the CalB derived isomer (**9-*N-Ac***) matches with the enantiomer of the amide derived from *uPA* amination (**9**).

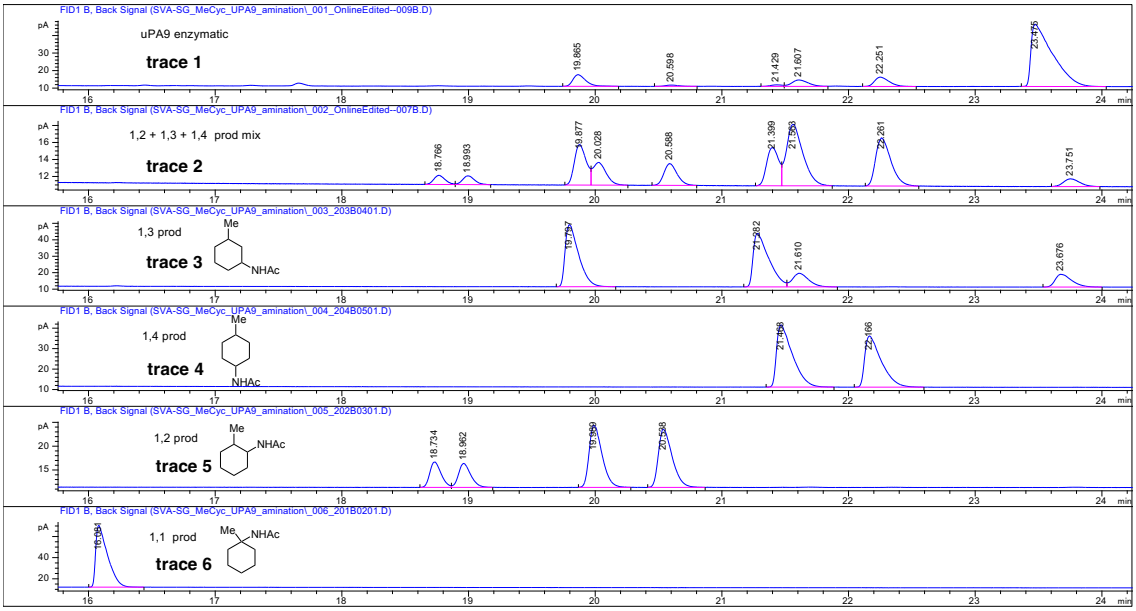
B.6. Selectivity Details for Enzymatic Amination of Methylcyclohexane (Ancillary Data for Fig. 2A)

Site-, diastereo-, and enantio- selectivity was determined by derivatizing the enzymatic product with acetyl chloride, followed by resolution of the resultant acetamides (**7**) by chiral GC equipped with an FID detector.

The whole-cell amination reaction with *uPA9* was set up with 20 ml of OD_{600} = 30 cells, as described previously for analytical scale reactions. After 16 hours, 20 mL of ethanol were added to the reaction mixtures. The mixture was vortexed and centrifuged at 14,000*g* for 5 minutes to precipitate coagulated proteins and cell debris. To an HPLC vial with an insert, 200 µL of the supernatants were transferred and analyzed by reverse-phase HPLC-MS. The remaining supernatants were combined and transferred to a 50-mL vial and ethanol was evaporated. To the remainder were added 20 mL ethyl acetate, 1 mL of acetyl chloride, and 1 g of potassium bicarbonate. After shaking this mixture overnight, the organic layer was separated and concentrated. The residue was redissolved in 1 mL of hexane and ethyl acetate mixture (1:1) and subjected to chiral GC-FID to determine product isomer distribution.

Stacked GC traces of UPA9 enzymatic reaction (derivatized) (trace 1), standard products (traces 3-6) and a mixture of standard products (trace 2):

Method: start at 50 °C, ramp at 10 °C/minute to 130 °C and hold for 22 minutes. Ramp at 70 °C/minute to 200 °C and hold for 5 minutes.



The acetamide of the primary amine

(*R*)-*N*-(1-(4-methoxyphenyl)ethyl)acetamide (**20**) [SVA-1-37]

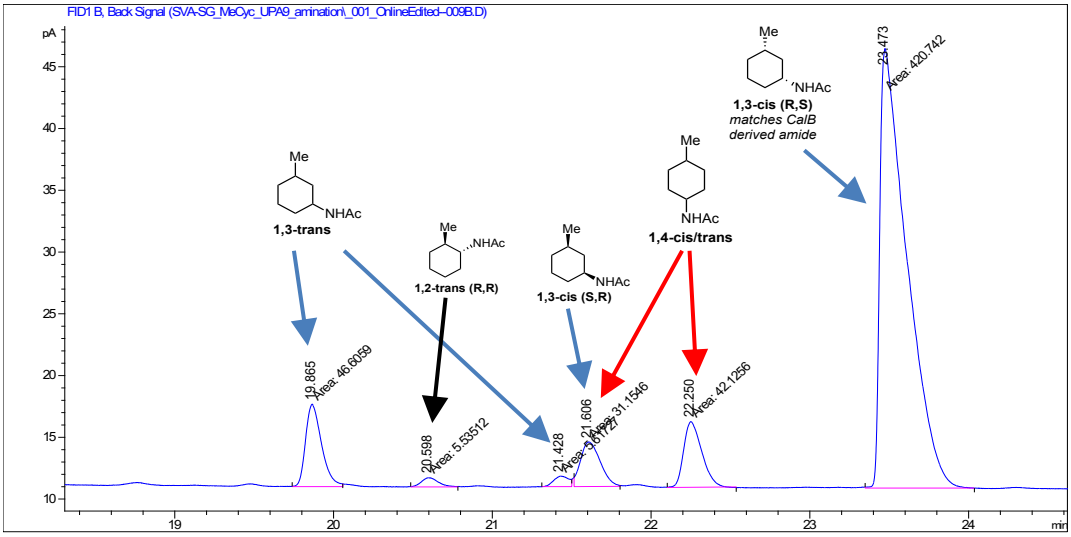
Notes:

a. The 1,3-isomer at 23.47 min is the major product from the enzymatic reaction (trace 1). It matches the peak for the CalB derived amide **6-*N*-*Ac*** (see section 5.2), thus permitting assignment of absolute configuration. A slight change in retention time as compared to the standard time (23.67 min, trace 3) is seen because of difference in concentration (confirmed by co-injection; not shown).

b. The standard products consist of a mixture of diastereomers and thus give multiple peaks (peaks with same areas belong to the same diastereomer)

c. One of the 1,3-isomer coelutes with one of the 1,4-isomer (21.61 min). These cannot be unambiguously resolved.

Zoomed in Chromatogram of the enzymatic reaction along with peak assignments: Peaks are assigned based on the absolute configurations determined by CalB acetylations (see previous section)



Entry	Time	Area	Area%	assignment
1	19.865	46.6	8.446	1,3-trans, enantiomer 1
2	20.598	5.5	1.003	1,2-trans (R,R)
3	21.428	5.6	1.018	1,3-trans, enantiomer 2
4	21.606	31.2	5.646	mixture of 1,4 and 1,3-cis (S, R) isomers
5	22.25	42.1	7.634	1,4-isomer
6	23.473	420.7	76.252	1,3-cis (R,S)

Site selectivity for 1,3 functionalization = (8.4 + 1.0 + 76.3)/100 = **85.7 %**

d.r. for 1,3 (cis:trans) = (420.7) : (46.6 + 5.6) = 420.7: 52.2 = **8.0 : 1**

e.r for 1,3 cis product = 420.7 : 31.2 = **93:7**

e.r. for 1,3 trans product = 46.6 : 5.6 = **89:11**

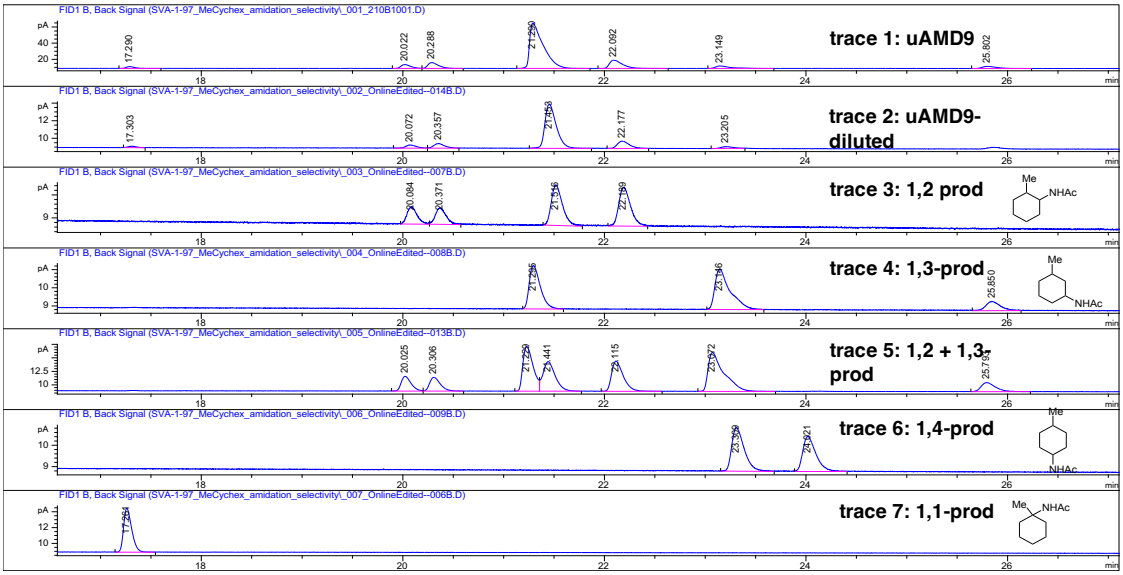
Note: because two of the minor product isomers overlap (at 21.60 min), the selectivity calculations are made conservatively and result in an underestimate. The actual selectivity is expected to be slightly higher.

B.7. Selectivity Details for Enzymatic Amidation of Methylcyclohexane (Ancillary Data for Fig. 2B)

Site- diastereo- and enantio- selectivity was determined by resolution of the acetamides (**7**) by chiral GC equipped with an FID detector. Briefly, a whole-cell amidation reaction with a reaction volume of 800 mL was setup as described previously (section 4.3.1). After 6 h, 350 mL of ethyl acetate/ hexanes (1:1) solution was added, the mixture was vigorously shaken and centrifuged (14000 rpm, 8 min, RT). The organic layer was analyzed by GC to determine isomeric distribution.

Stacked GC traces of uAMD9 enzymatic product (trace 1), diluted enzymatic product solution (trace 2), standard products (traces 3–7) and a mixture of standard products (trace 5):

Method: start at 50 °C, ramp at 10 °C/minute to 130 °C and hold for 22 minutes. Ramp at 70 °C/minute to 200 °C and hold for 5 minutes.

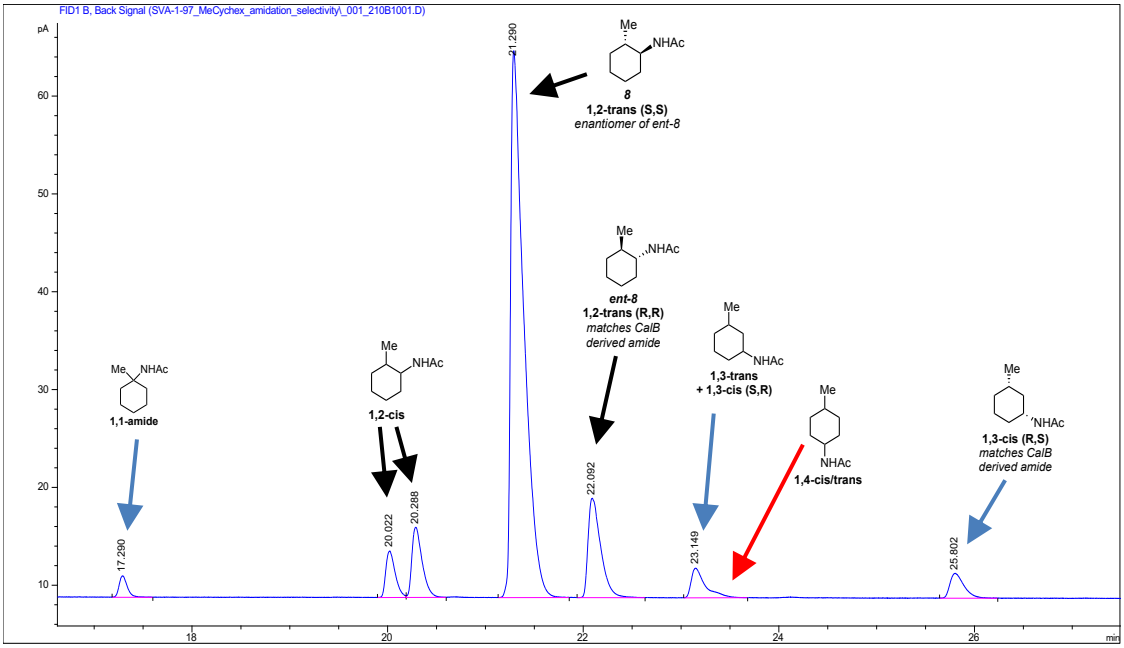


Notes:

a. The peak at 22.1 (traces 1-3,5) matches the peak for the CalB derived amide **ent-8** (see section 5.1), with a determined absolute configuration. The uAMD8 derived major product **8** (1,2-isomer at 21.29 min, trace 1; 21.45 min, trace 2) is the enantiomer of **ent-8**, thus permitting assignment of absolute configuration.

b. The standard products consist of a mixture of diastereomers and thus give multiple peaks (peaks with same areas belong to the same diastereomer)

Zoomed in Chromatogram of the enzymatic reaction along with peak assignments: Peaks are assigned based on standard products and the absolute configurations determined by CalB acetylations (see previous section)



Entry	Time	Area	Area%	assignment
1	17.290	12.8	1.65	1,1-isomer
2	20.022	31.8	4.12	1,2-cis, enantiomer 1
3	20.288	52.6	6.82	1,2-cis, enantiomer 2
4	21.290	526.7	68.31	**8;** 1,2-trans (S,S)
5	22.092	91.4	11.85	**ent-8;** 1,2-trans (R,R)
6	23.149	30.6	3.96	1,3-trans + 1,3-cis (S,R) + 1,4-isomer (formation uncertain)
7	25.802	25.3	3.27	1,3-cis (R,S)

Site selectivity (ss) for 1,3 functionaliation = (4.1 + 6.8 + 68.3 + 11.8)/100 = **91%**

dr for 1,2 (trans/cis) = (526.7 + 91.4) : (31.8 + 52.6) = 618.1: 84.4 = **7.3 : 1**

er for 1,2 trans product (**8**) = 526.7 : 91.4 = **85:15**

Identical selectivity was observed for *uAMD7* and *uAMD9.*

For *uAMD6:*

ss = 89%

dr for 1,2(trans/cis) = 5:1

er for 1,2 trans product (**8**) = 75:25

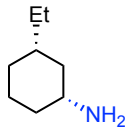
For *uAMD5* and *uAMD4:*

ss = 76%

dr for 1,2(trans/cis) = 6:1

er for 1,2 trans product (**8**) = 65:35

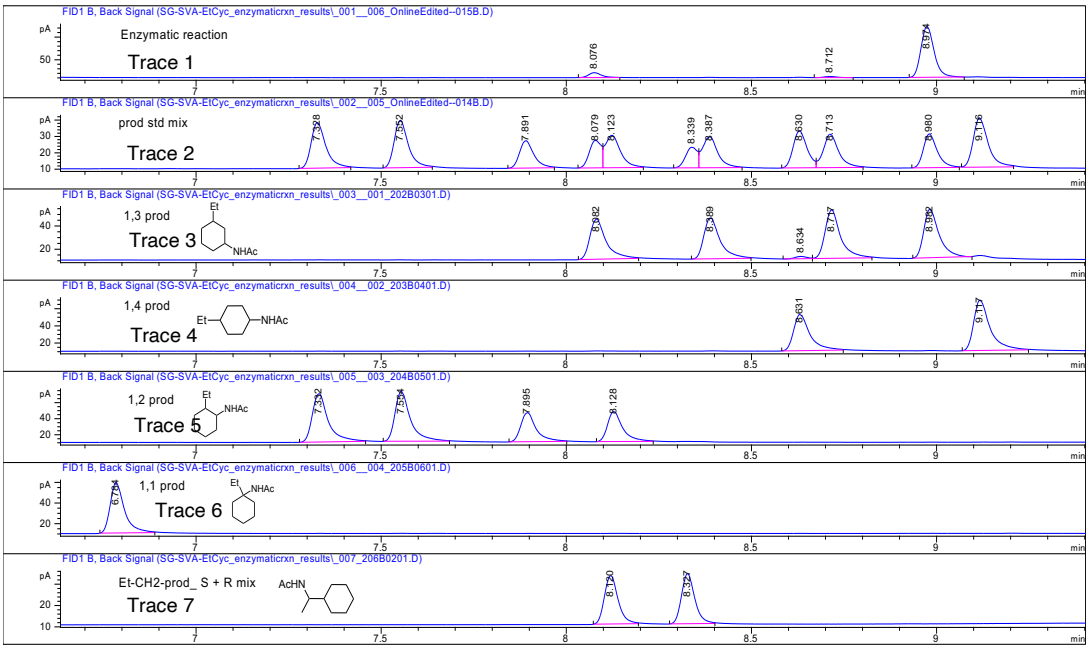
B.8. Selectivity Data for Substrates in Fig. 3A



9, 135 TTN
> 99% *ss*, 10.5:1 *dr*
98:2 *er*

Stacked GC traces of UPA9 enzymatic reaction (derivatized) (trace 1), standard products (traces 3–7) and a mixture of standard products (trace 2):

Method: start at 130 °C, ramp at 5 °C/minute to 200 °C and hold for 5 minutes.



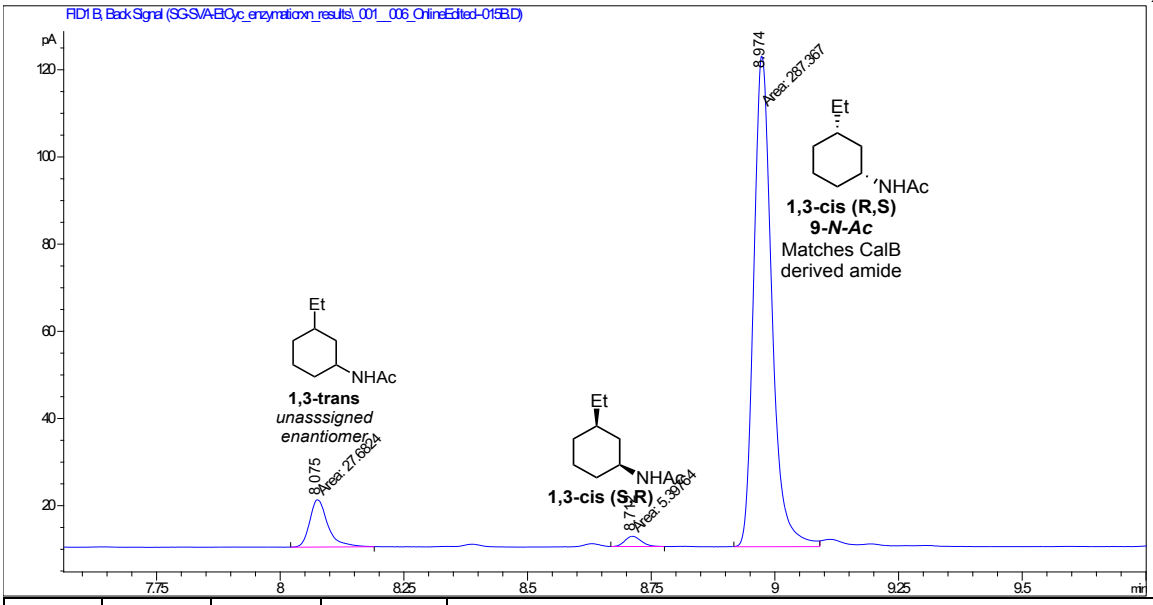
Notes:

a. The peak of the major product at 8.97 (traces 1–3) matches the peak for the CalB derived amide **9-*N-Ac,*** with a determined absolute configuration (see section 5.3).

b. The standard products consist of a mixture of diastereomers and thus give multiple peaks (peaks with same areas belong to the same diastereomer)

Zoomed in Chromatogram of the enzymatic reaction along with peak assignments:

Peaks are assigned based on standard products and the absolute configurations determined by CalB acetylations (see previous section)



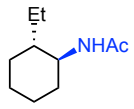
Entry	Time	Area	Area%	assignment
1	8.075	27.7	8.64	1,3-trans, enantiomer 1
2	8.712	5.4	1.68	1,3-cis (S,R)
3	8.974	287.4	89.68	**9-*N*-*Ac*,** 1,3-cis (R,S). Matches CalB derived amide.

ss > 99%; only 1,3-isomer detected.

dr = (287.4 + 5.4)/27.7 = 10.5:1

er = 98:2

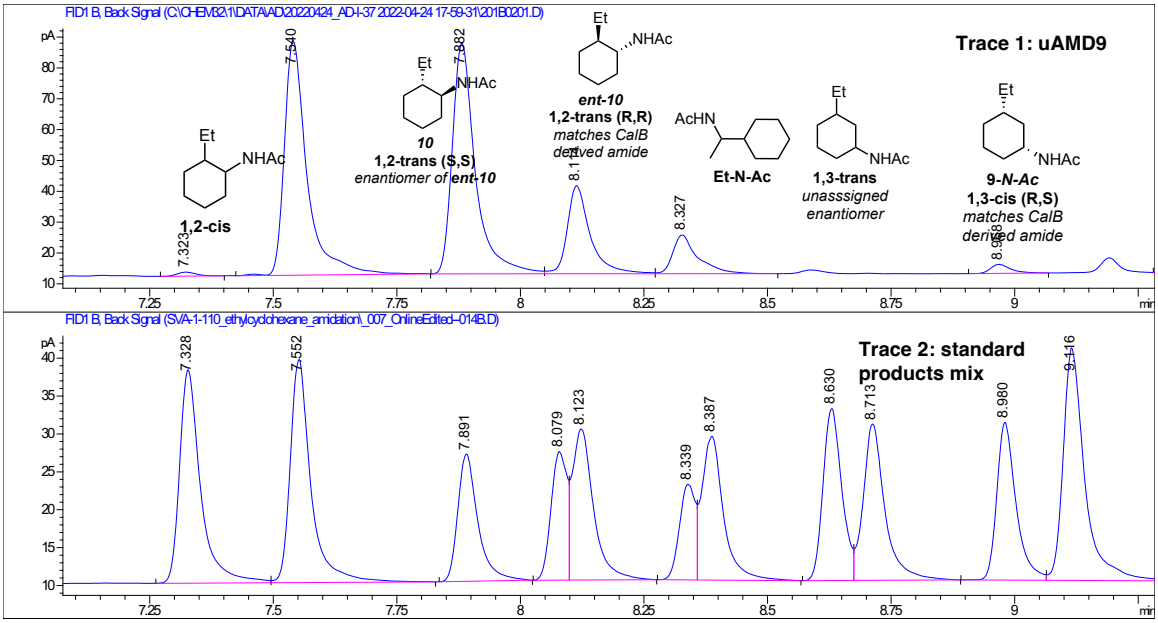
Note that *er* for the minor diastereomer (1,3-trans) is also high (estimated >98:2).



10, 205 TTN[b]
78% *ss*, 1:1 *dr*
cis diast. - 98:2 *er*
trans diast. (S,S)- 75:25 er

Stacked GC traces of uAMD9 enzymatic reaction (trace 1), standard products mix (trace2; for site specific isomer standards, see traces for ***9*** *on previous pages)*

Method: start at 130 °C, ramp at 5 °C/minute to 200 °C and hold for 5 minutes.



Entry	Time	Area	Area%	assignment
1	7.324	3	0.474	1,2-cis, enantiomer 1
2	7.54	240.3	38.368	1,2-cis, enantiomer t2
3	7.883	249	39.747	**10,** 1,2 – trans (S,S)
4	8.113	86.5	13.812	***ent*-10** 1,2 – trans (R,R) trans (matches CalB derived amide) + Et-N-Ac
5	8.327	38.8	6.195	Et-N-Ac + 1,3- trans, enantiomer 1
6	8.97	8.8	1.403	**9-*N*-Ac** 1,3 – cis (R,S) (matches CalB derived amide)

total area = 626.4

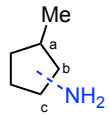
ss (for 1,2) = (3+240.3+249)/ 626.4 = 78%

dr (1,2-cis: 1,2-trans) = (3 + 240.3): 249 = 1:1

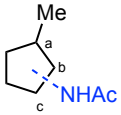
er (1,2-cis product) = (240.3: 3) = 98:2

er (1,2-trans product) = 249: 86.5 = 75:25

Note: because some of the minor product isomers overlap (at 8.11 and 8.32 min), the selectivity calculations are made conservatively and result in an underestimate. The actual selectivity may be slightly higher.



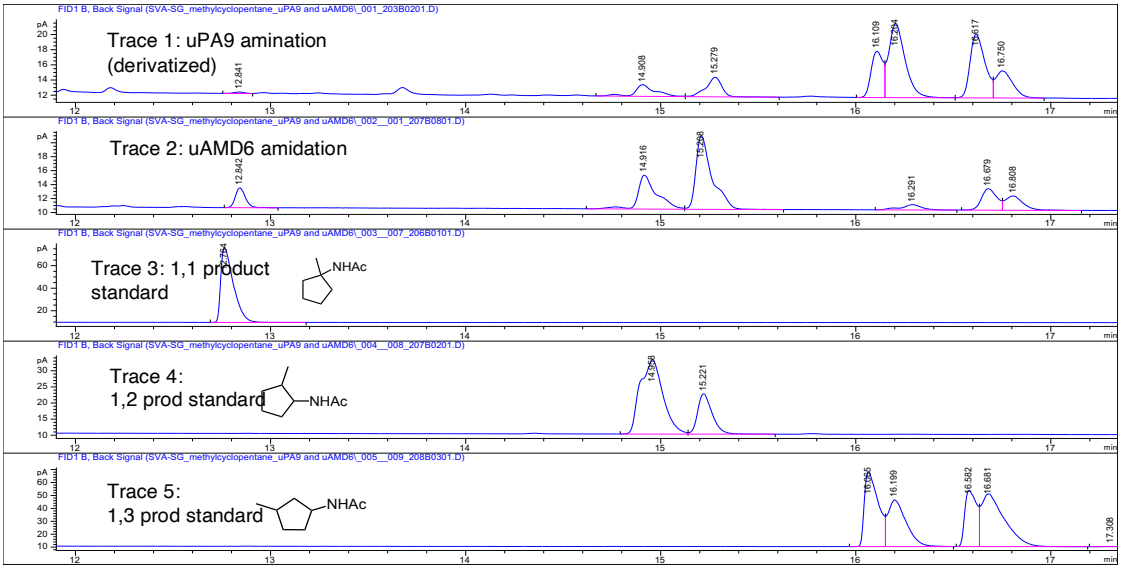
11, 105 TTN[a]
ss for b:c = 1:6



12, 110 TTN[c]
ss for a:b:c
= 1:8:3

Stacked GC traces of uPA9 enzymatic reaction, derivatized (trace 1), uAMD6 enzymatic reaction (trace 2) and site-specific standard products (traces 3–5)

Method: start at 50 °C, ramp at 10 °C/minute to 130 °C and hold for 22 minutes. Ramp at 70 °C/minute to 200 °C and hold for 5 minutes.



Note: while site specific isomers are well resolved, individual diastereomers of site isomers cannot be resolved.

With *uPA9*

Entry	Time	Area	assignment
1	12.84	-	1,1 isomer
2	14.9-15.3	22.4	1,2 isomer
3	16.1-16.8	147.6	1,3 isomer

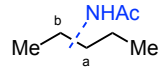
ss for 1,2 isomer: 1,3 isomer = 22.4:147.6 = 1:6

With *uAMD6*

Entry	Time	Area	assignment

1	12.84	10	1,1 isomer
2	14.9-15.3	86.7	1,2 isomer
3	16.1-16.8	33.6	1,3 isomer

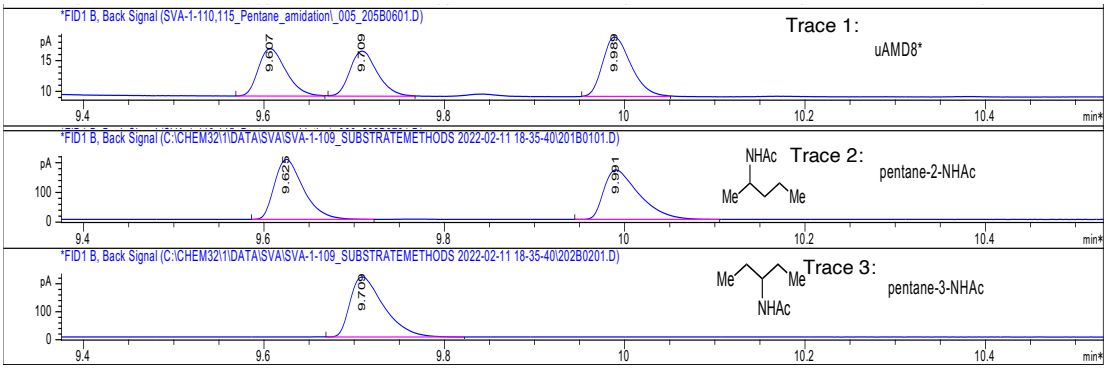
ss for 1,1 isomer: 1,2 isomer: 1,3 isomer = 10: 86.7:33.6 = 1:8:3



19, 35 TTN[d]
ss for a:b = 1:2.4
er (for b) = 56:44

Stacked GC traces for uAMD8 enzymatic reaction (trace 1) and site-specific standard products (traces 2,3)*

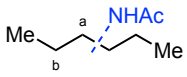
Method: start at 50 °C, ramp at 10 °C/minute to 130 °C and hold for 5 minutes. Ramp at 70 °C/minute to 200 °C and hold for 5 minutes.



Entry	Time	Area	Area%	assignment
1	9.607	16.9	31.086	2-NHAc enantiomer 1
2	9.709	15.8	29.142	3-NHAc
3	9.989	21.6	39.772	2-NHAc enantiomer 2

ss = (16.9+21.6):15.8 = 38.5:15.8 = **2.4:1**

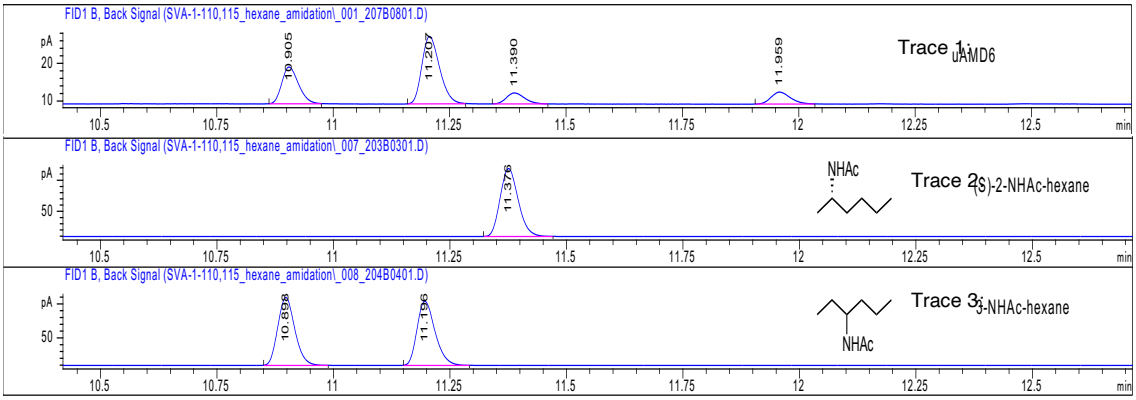
er for 2-NHAc =21.6: 16.9 = 56:44



20, 40 TTN[c]
ss for a:b = 4:1
er (for a) = *65:35*
er (for b) = *54:46*

Stacked GC traces for uAMD6 enzymatic reaction (trace 1) and site-specific standard products (traces 2,3)

Method: start at 50 °C, ramp at 10 °C/minute to 130 °C and hold for 5 minutes. Ramp at 70 °C/minute to 200 °C and hold for 5 minutes.

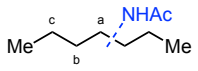


Entry	Time	Area	Area%	assignment
1	10.905	25.3	27.729	3-NHAc enantiomer 1
2	11.207	48	52.696	3-NHAc enantiomer 2
3	11.39	8.2	8.956	2-NHAc (S)
4	11.959	9.7	10.620	2-NHAc (R)

ss for 3-isomer : 2-isomer = (25.3 + 48): (8.2 + 9.7) = 73.3: 17.9 = 4:1

er for 3-isomer = 48:25.3 = 65:35

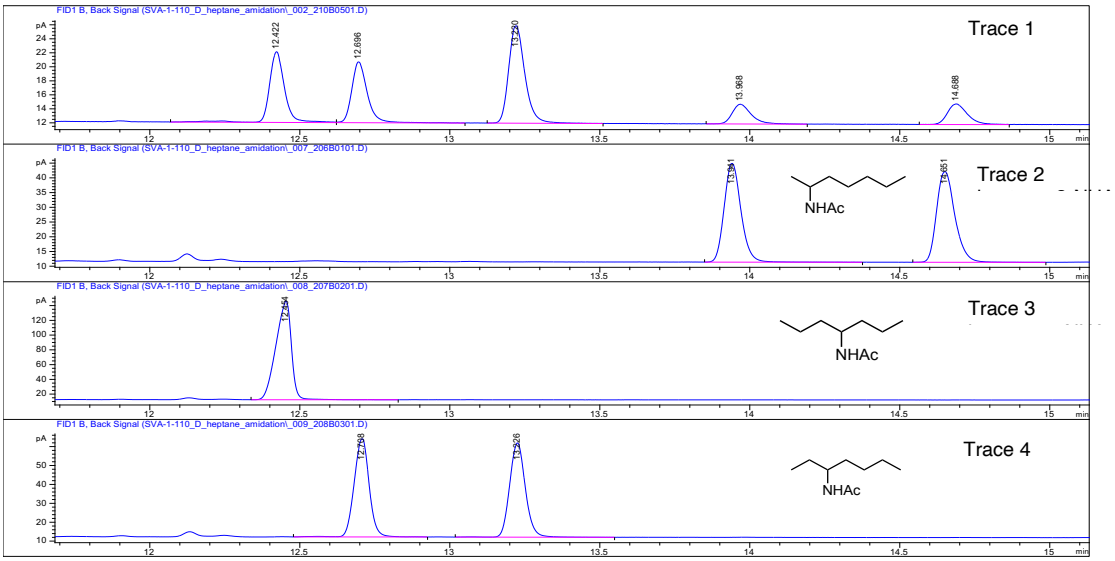
er for 2-isomer = 9.7: 8.2 = 54:46



21, 45 TTN[e]
ss for a:b:c = 1.3: 3.2:1
er (for b) = *63:37*
er (for b) = *53:47*

Stacked GC traces for uAMD7-T438S enzymatic reaction (trace 1) and site-specific standard products (traces 2-4)

Method: start at 50 °C, ramp at 10 °C/minute to 130 °C and hold for 10 minutes. Ramp at 70 °C/minute to 200 °C and hold for 5 minutes.

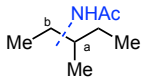


Entry	Time	Area	Area%	assignment
	12.422	33.7	23.707	heptane-4-NHAc
	12.696	30.4	21.364	heptane-3-NHAc enantiomer 1
	13.22	52.3	36.779	heptane-3-NHAc enantiomer 2
	13.969	12	8.427	heptane-2-NHAc enantiomer 1
	14.689	13.8	9.722	heptane-2-NHAc enantiomer 2

ss; 4-isomer : 3-isomer : 2-isomer = 33.7: (30.4 +52.3): (12 + 13.8) = 1.3: 3.2: 1

er for 3-isomer: 63:37

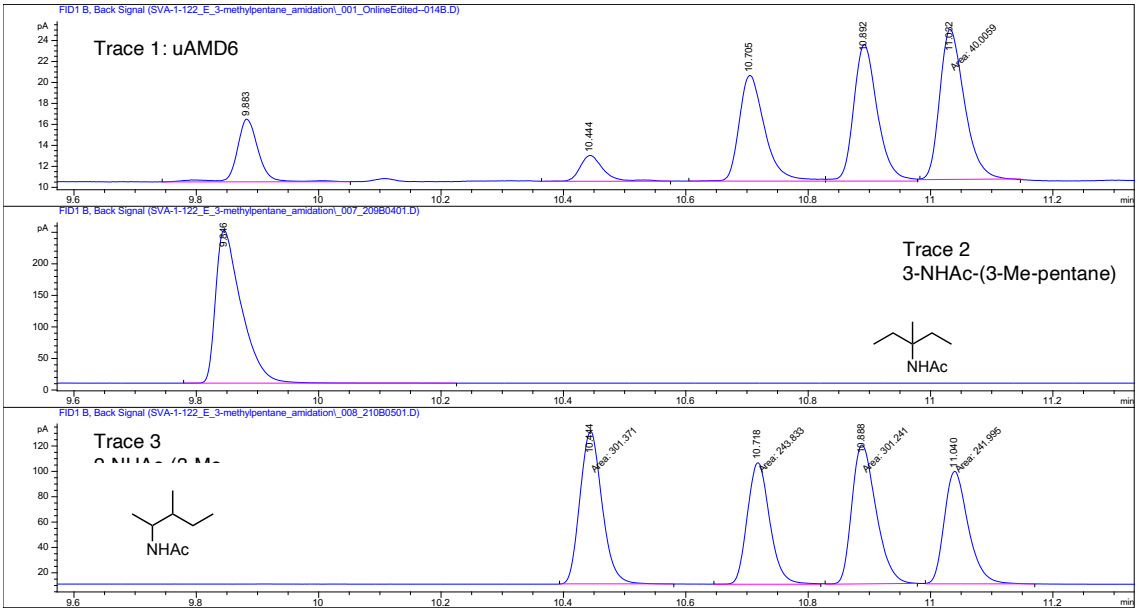
er for 2-isomer: 53:47



22, 70 TTN[c]
ss for a:b = 1:8
dr (for b) = 1:1.7
er (minor diast.) = 85:15
er (major diast.) = 58:42

Stacked GC traces for uAMD6 enzymatic reaction (trace 1) and site-specific standard products (traces 2,3).

Method: start at 50 °C, ramp at 10 °C/minute to 130 °C and hold for 10 minutes. Ramp at 70 °C/minute to 200 °C and hold for 5 minutes.



Entry	Time	Area	Area%	assignment
1	9.883	13.1	10.8	3-NHAc-(3-Me-pentane)
2	10.444	5.9	4.9	2-NHAc-(3-Me-pentane) diastereomer 1, enantiomer 1
3	10.705	28.1	23.2	2-NHAc-(3-Me-pentane) diastereomer 2, enantiomer 1
4	10.892	33.8	28.0	2-NHAc-(3-Me-pentane) diastereomer 1, enantiomer 2
5	11.032	40	33.1	2-NHAc-(3-Me-pentane) diastereomer 2, enantiomer 2

ss for 3-isomer : 2-isomer = 10.8: 89.2 = 1:8

dr for for 2-isomer, diastereomer 1: diastereomer 2 = (33.8 + 5.9): (28.1 + 40.1) = 39.7: 68.1 = 1:1.7

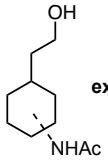
er (diastereomer 1) =33.8: 5.9 = 85:15,

er (diastereomer 2) = 40.1: 28.1 = 58:42

B.9. Initial Activity Substrate Scope

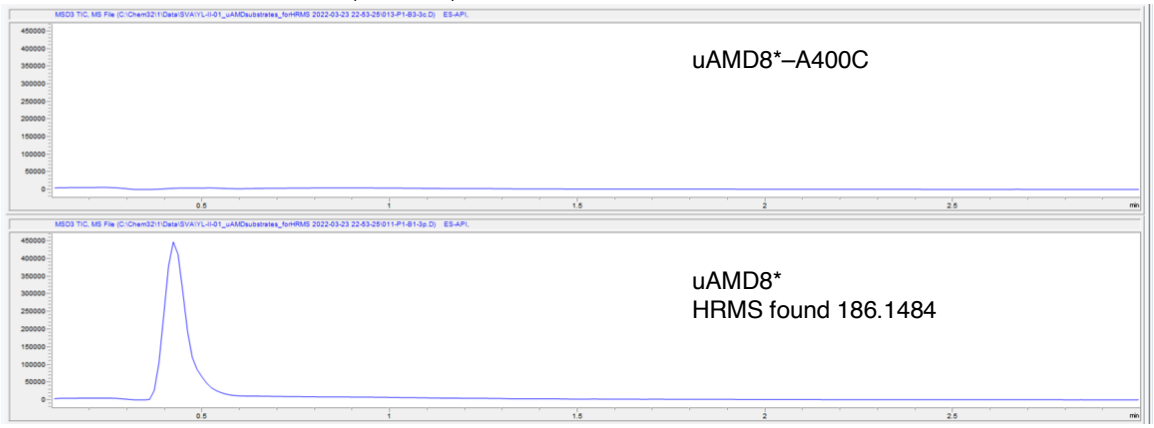
For each substrate, an analytical scale, whole-cell reaction for amination (anaerobic) or amidation (aerobic) was set up with *uPA9* or *uAMD8** respectively as described in sections 3.3.1 and 4.3.1. In each case, reactions were also set up with *uAMD8*-A400C.* This variant, with an axial Cys mutation is inactive for nitrene transfer reactions and acts as a negative control. As described in sections 3.3.1 and 4.3.1, the reactions were worked up by adding ethanol, centrifuged, and the supernatant analyzed by HPLC-MS for detection of the [M+H] product ion. The sample was also analyzed by HRMS to obtain the exact mass for the [M+H] analyte. For all substrates in Fig. 3B, product was detected only in the reactions with *uPA9* and *uAMD8**.

23

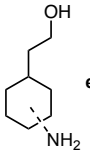


exact mass [M+H]:
186.1494

LCMS trace for SIM detection (M+1,186)

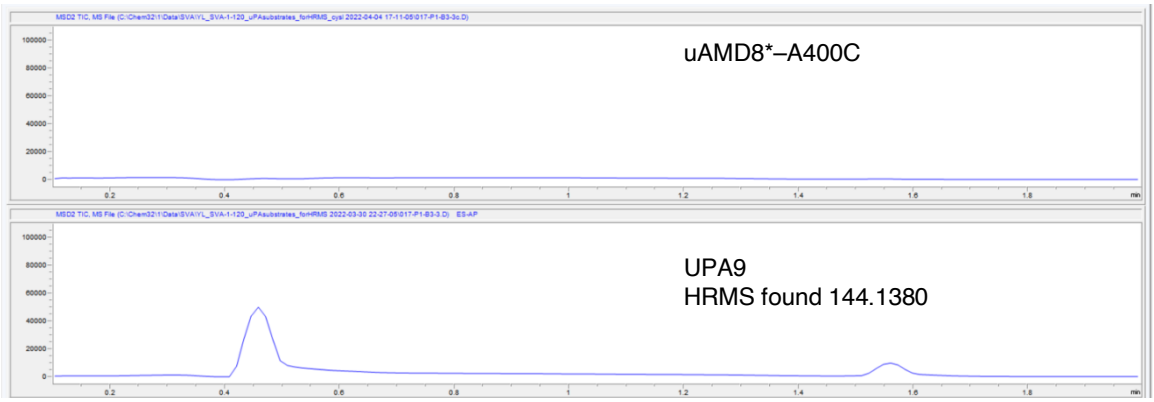


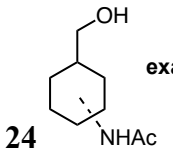
23



exact mass [M+H]:
144.1388

LCMS trace for SIM detection (M+1,144)

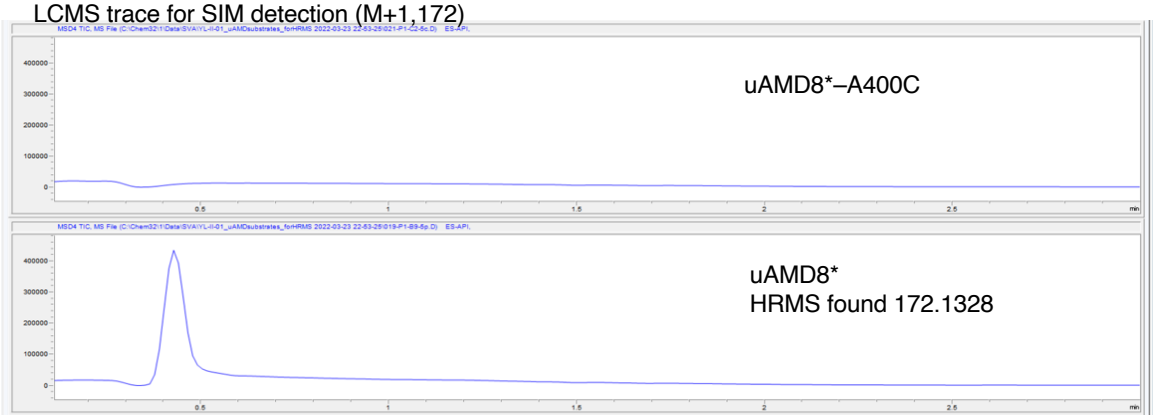


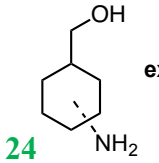


24

exact mass [M+H]:
172.1338

LCMS trace for SIM detection (M+1,172)

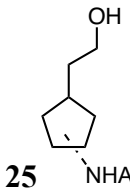




24

exact mass [M+H]:
130.1232

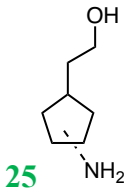
LCMS trace for SIM detection (M+1,130)



exact mass [M+H]:
172.1338

25

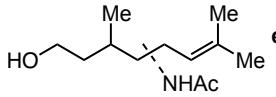
LCMS trace for SIM detection (M+1,172)



exact mass [M+H]:
130.1232

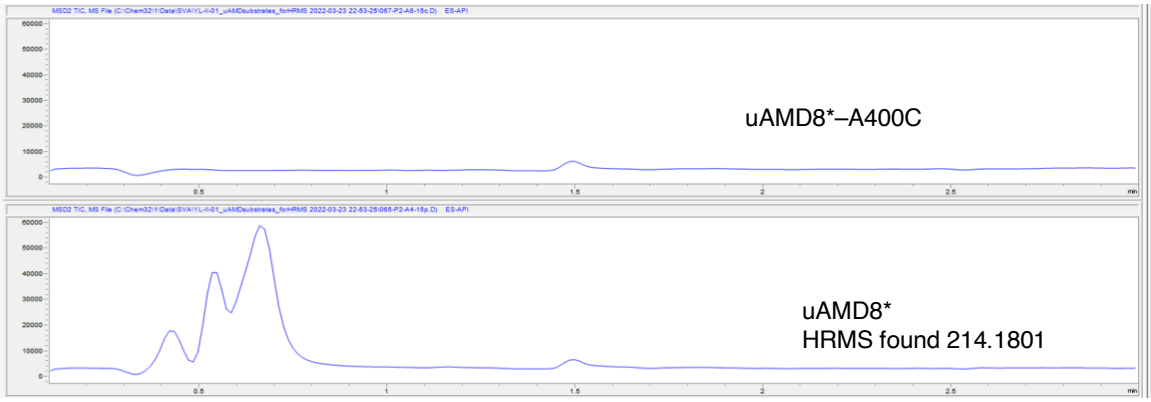
25

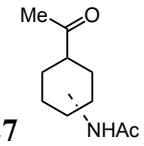
LCMS trace for SIM detection (M+1,130)



26 **exact mass [M+H]: 214.1807**

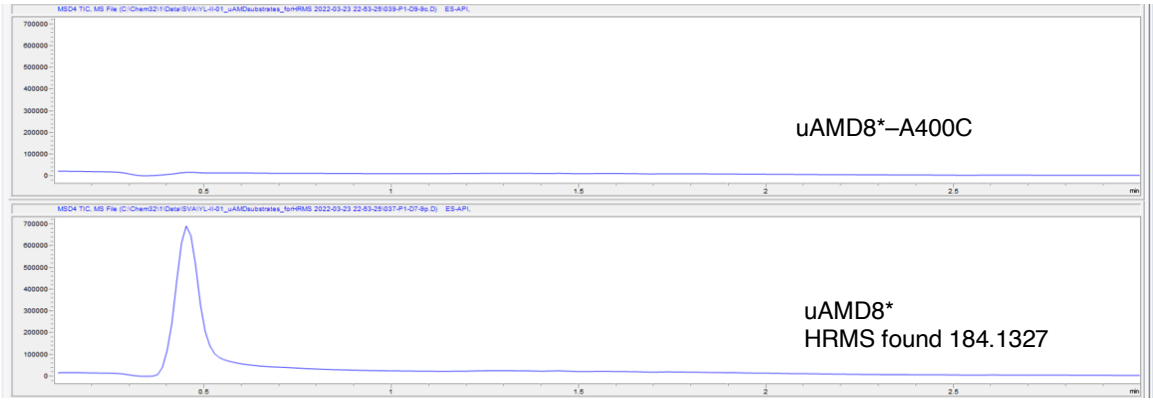
LCMS trace for SIM detection (M+1,214)

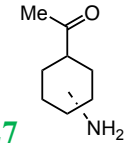




27 **exact mass [M+H]: 184.1338**

LCMS trace for SIM detection (M+1,184)

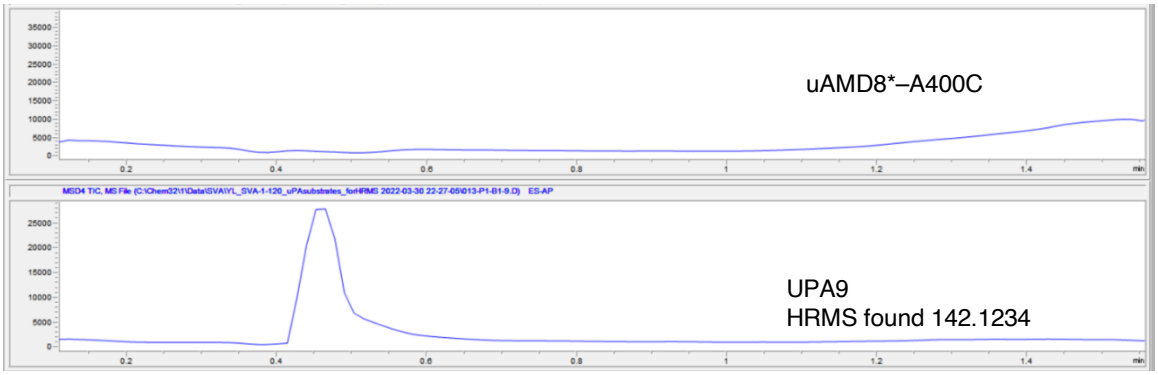


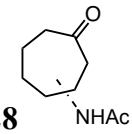


exact mass [M+H]:
142.1232

27

LCMS trace for SIM detection (M+1,142)

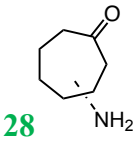




exact mass [M+H]:
170.1181

28

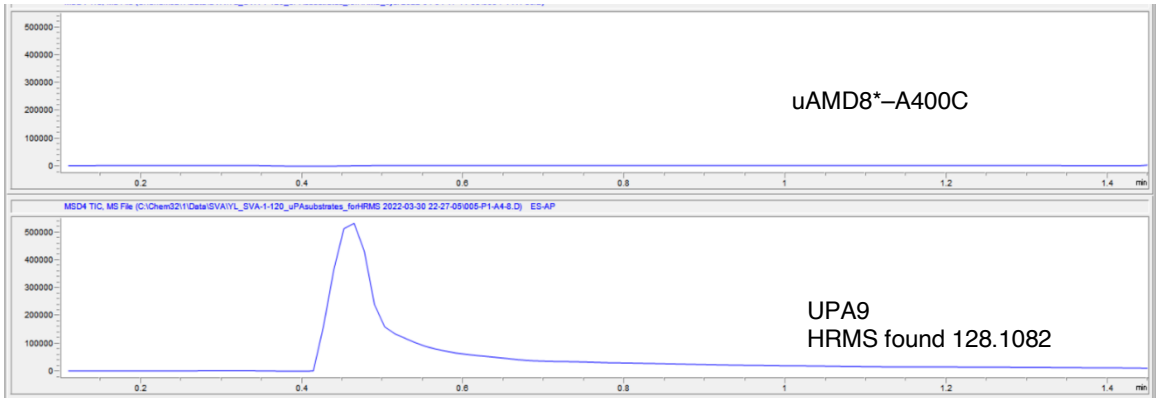
LCMS trace for SIM detection (M+1,170)

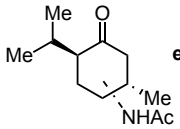


exact mass [M+H]: 128.1075

28

LCMS trace for SIM detection (M+1,128)

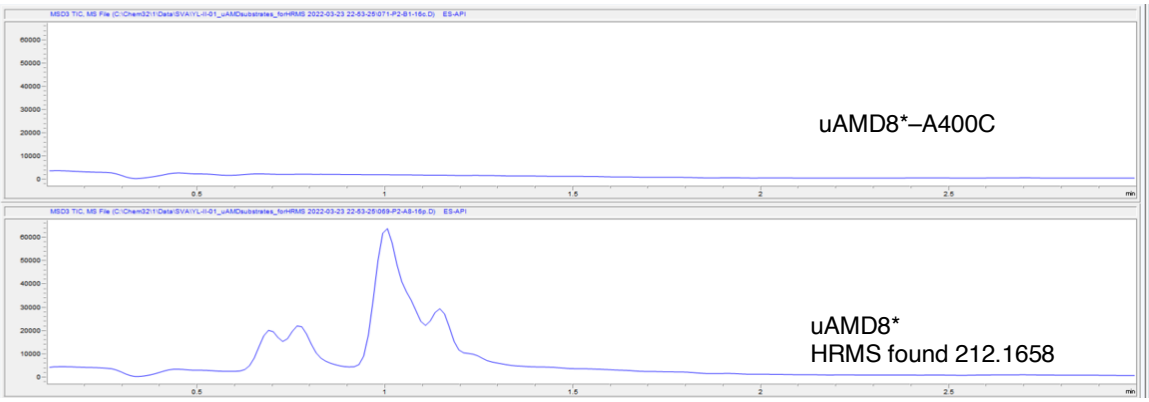


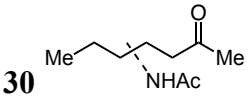


exact mass [M+H]: 212.1651

29

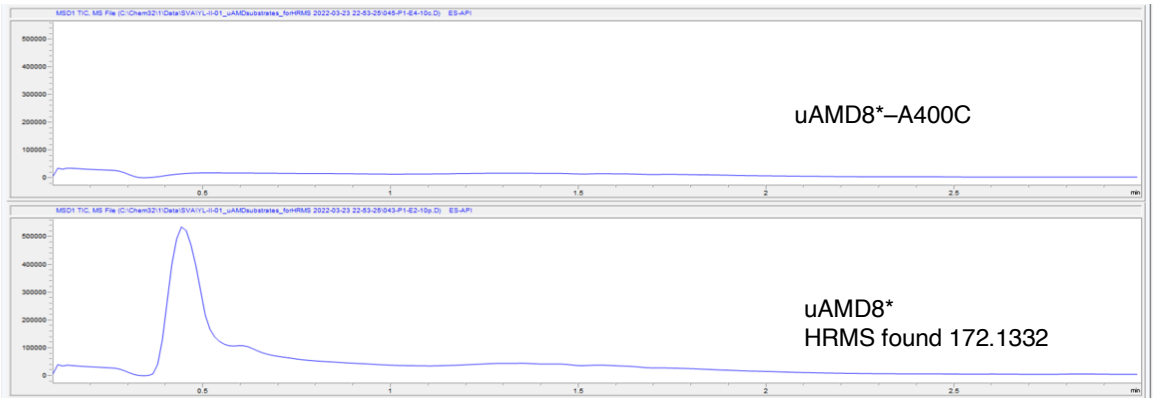
LCMS trace for SIM detection (M+1,212)

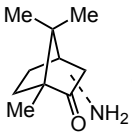




30 **exact mass [M+H]: 172.1338**

LCMS trace for SIM detection (M+1,172)

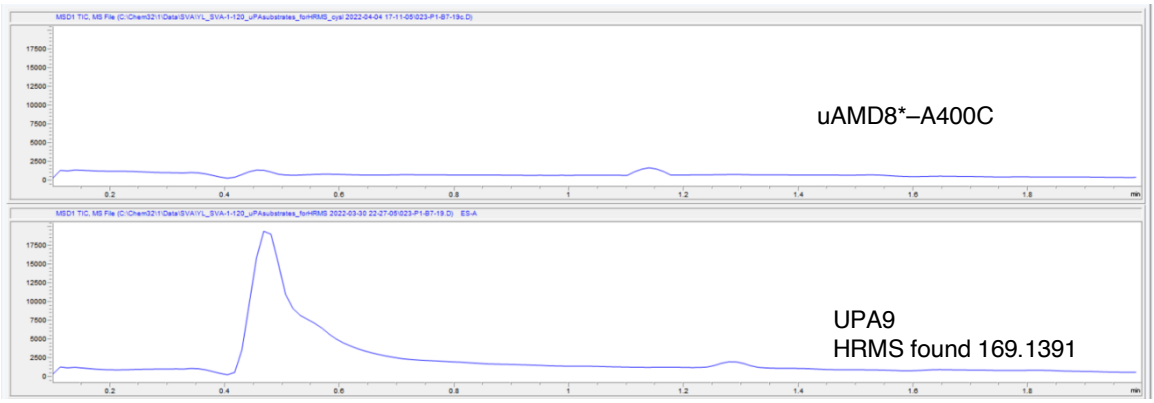


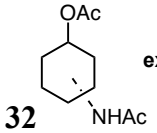


exact mass [M+H]: 168.1388

31

LCMS trace for SIM detection (M+1,168)

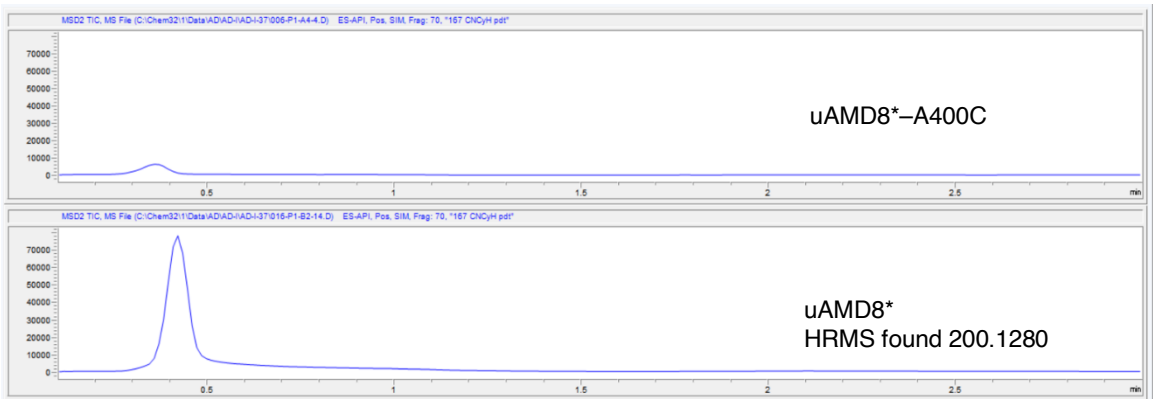


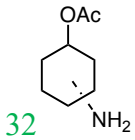


32

exact mass [M+H]: 200.1287

LCMS trace for SIM detection (M+1,200)

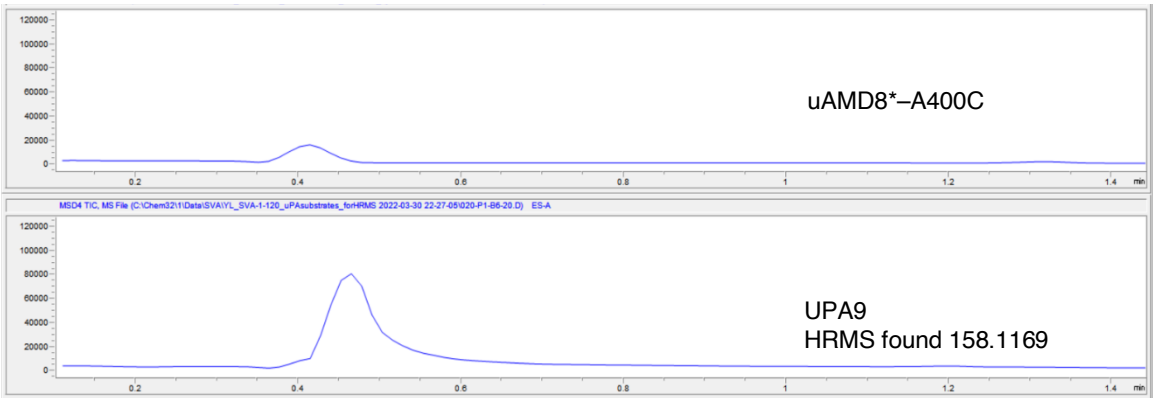


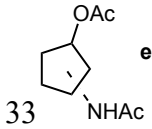


32

exact mass [M+H]: 158.1181

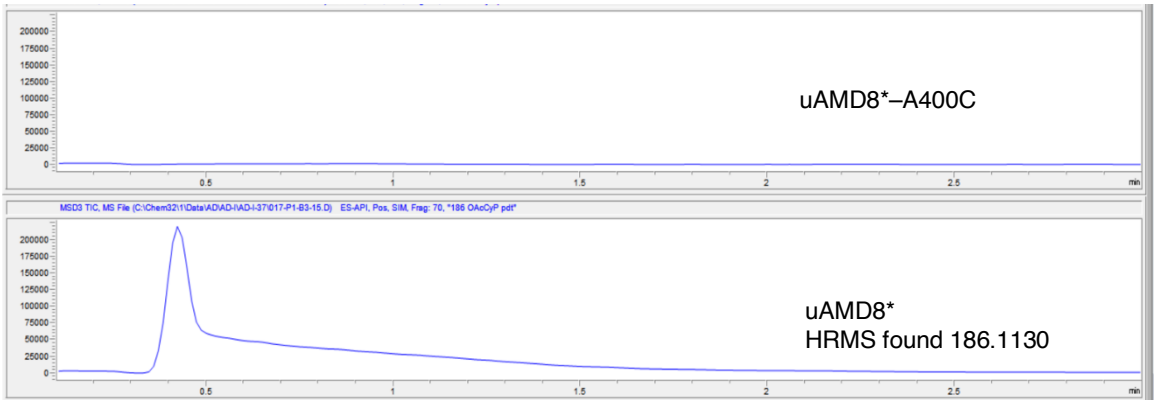
LCMS trace for SIM detection (M+1,158)

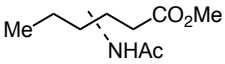




33 **exact mass [M+H]: 186.1128**

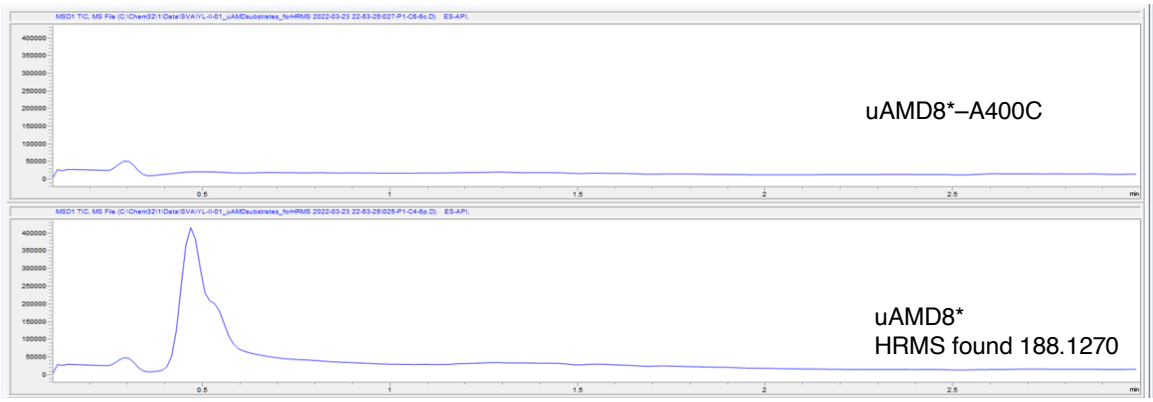
LCMS trace for SIM detection (M+1,186)

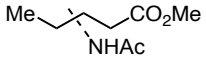




34 **exact mass [M+H]: 188.1287**

LCMS trace for SIM detection (M+1,188)

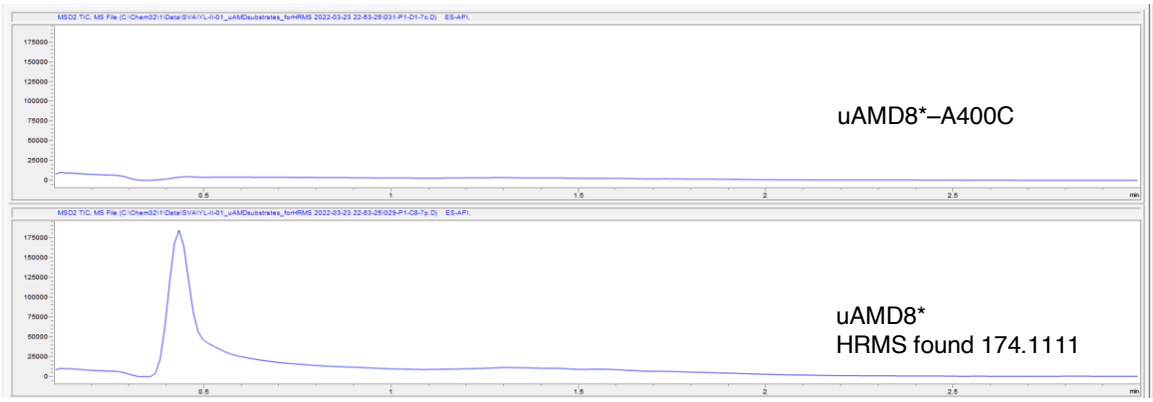


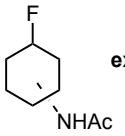


exact mass [M+H]: 174.1130

35

LCMS trace for SIM detection (M+1,174)

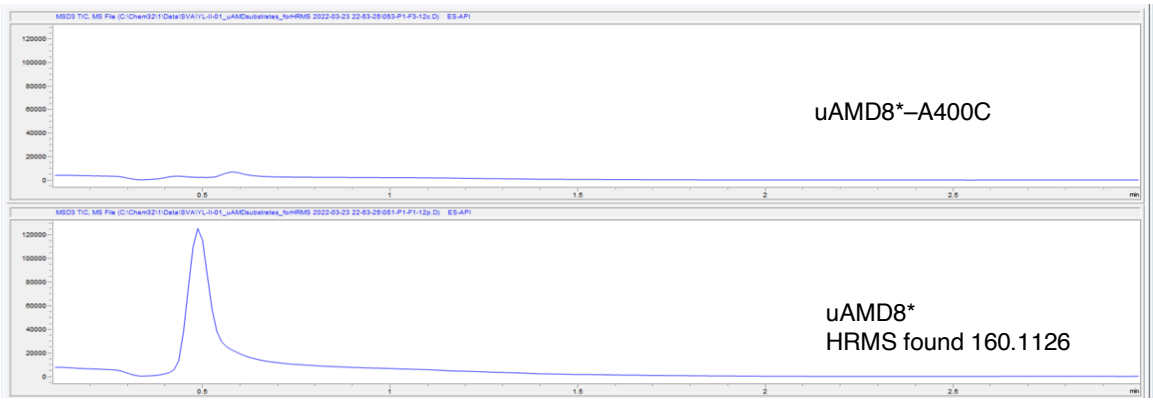


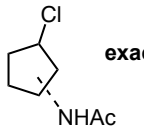


exact mass [M+H]: 160.1138

36

LCMS trace for SIM detection (M+1,160)

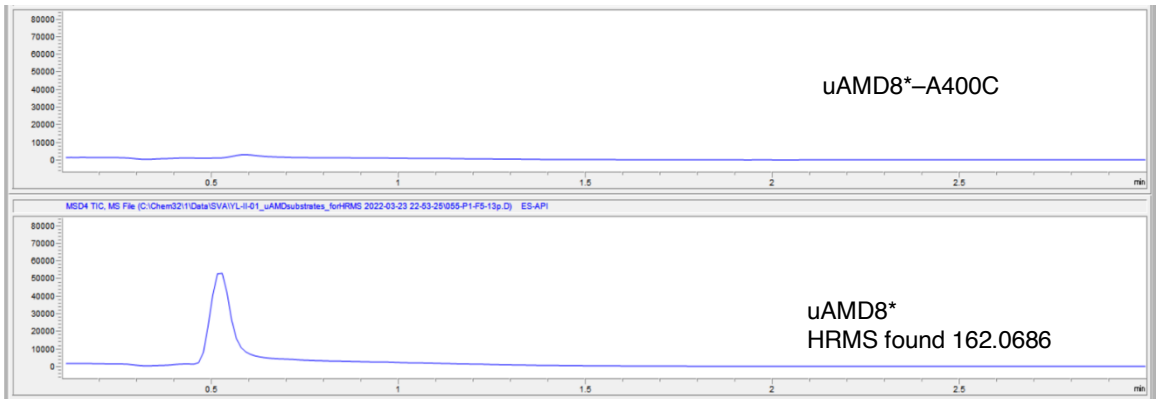


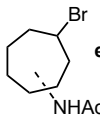


exact mass [M+H]:
162.0690

37

LCMS trace for SIM detection (M+1,162)

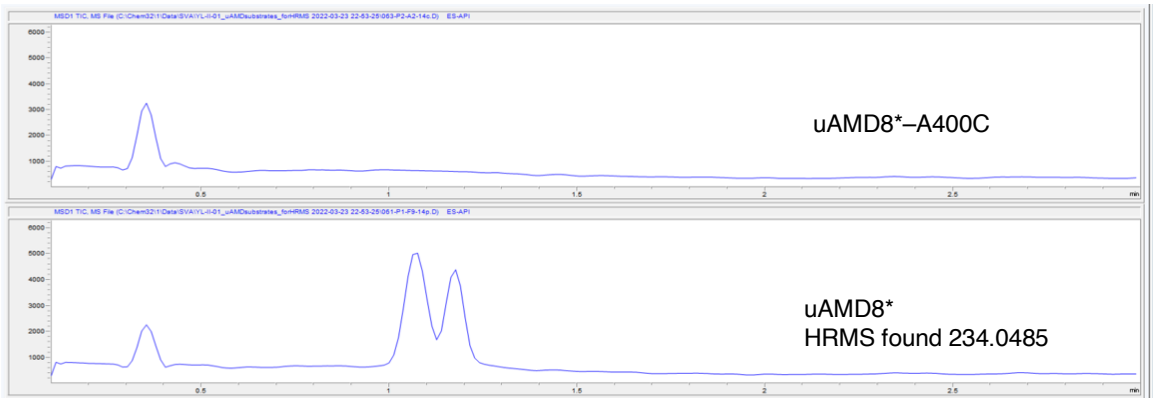


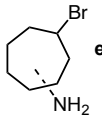


exact mass [M+H]:
234.0494

38

LCMS trace for SIM detection (M+1,234)

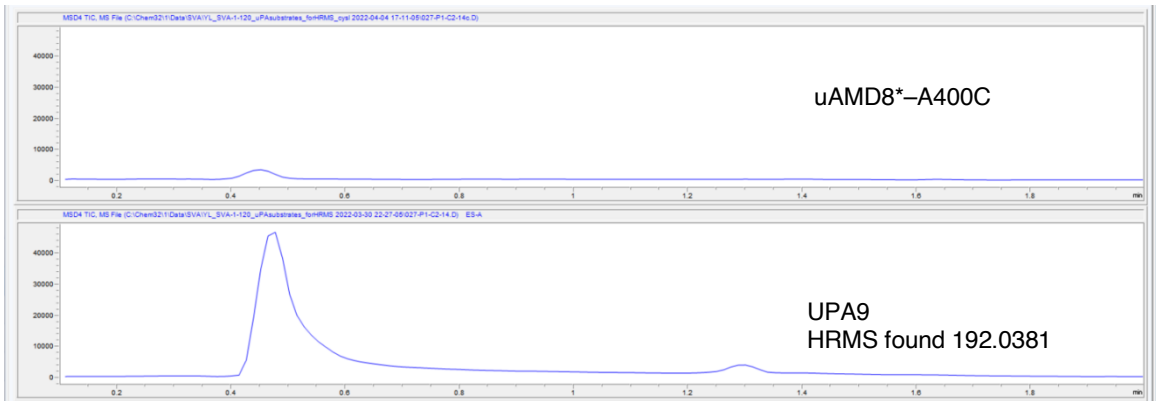


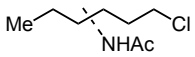


exact mass [M+H]: 192.0388

38

LCMS trace for SIM detection (M+1,192)

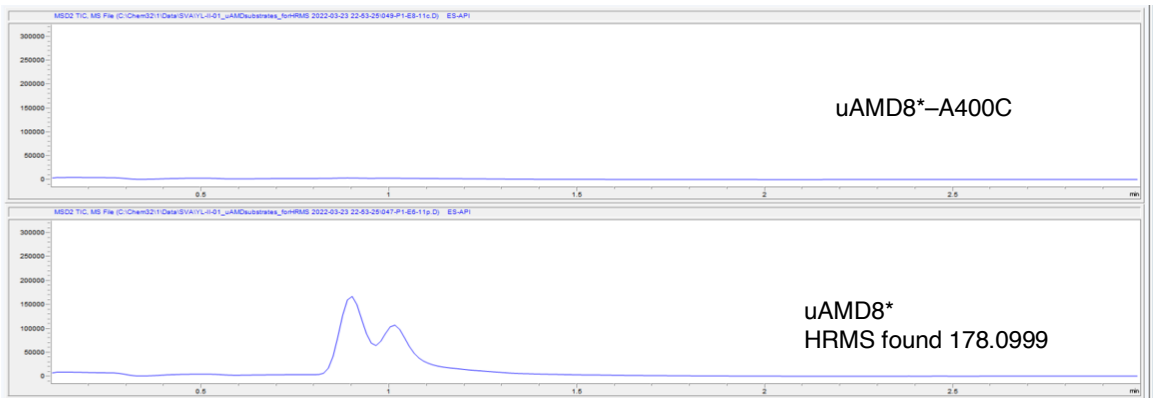


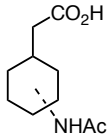


exact mass [M+H]: 178.1005

39

LCMS trace for SIM detection (M+1,178)

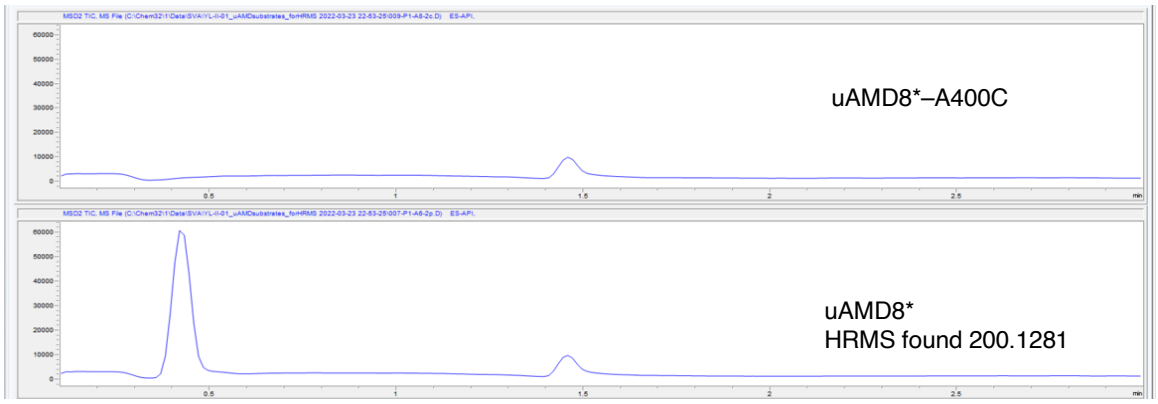


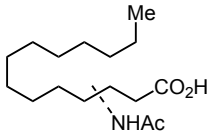


exact mass [M+H]: 200.1287

40

LCMS trace for SIM detection (M+1,200)

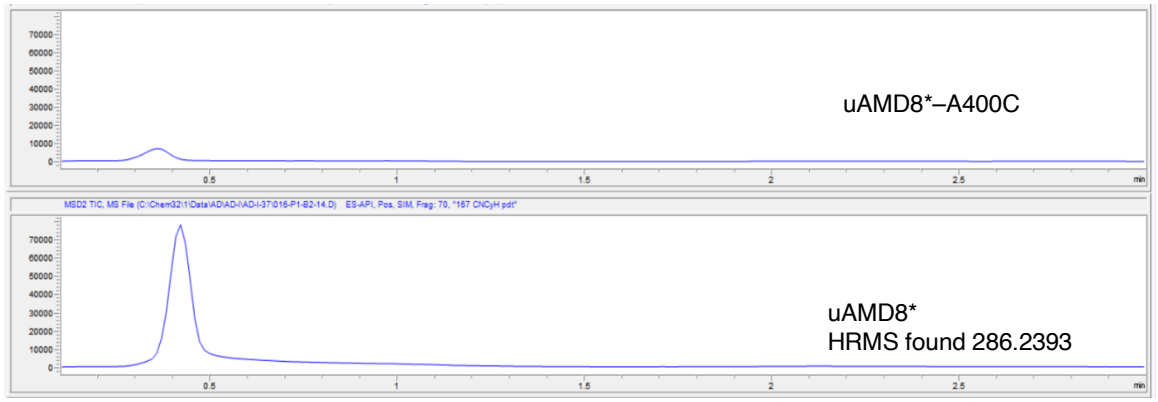


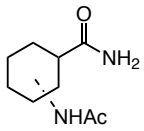


exact mass [M+H]: 286.2382

41

LCMS trace for SIM detection (M+1,286)

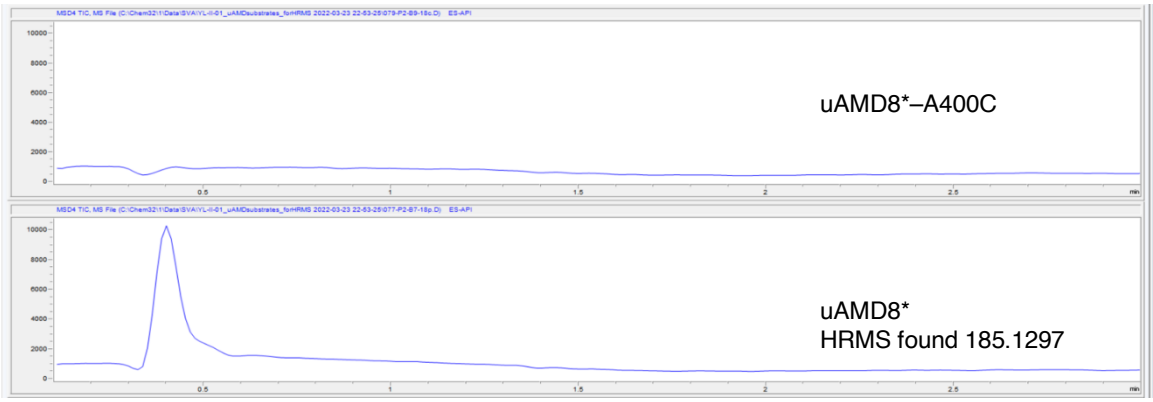


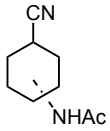


exact mass [M+H]:
185.1290

42

LCMS trace for SIM detection (M+1,185)

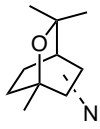




exact mass [M+H]:
167.1184

43

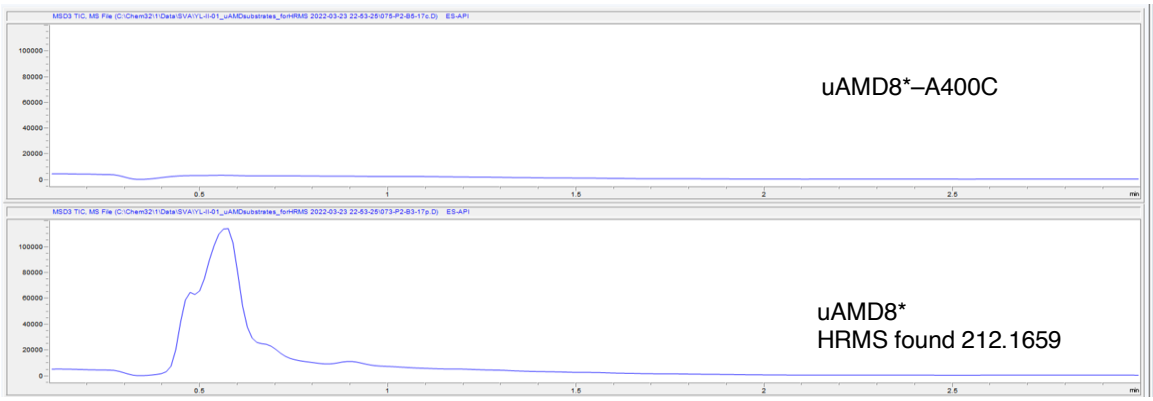
LCMS trace for SIM detection (M+1,167)



exact mass [M+H]: 212.1651

44

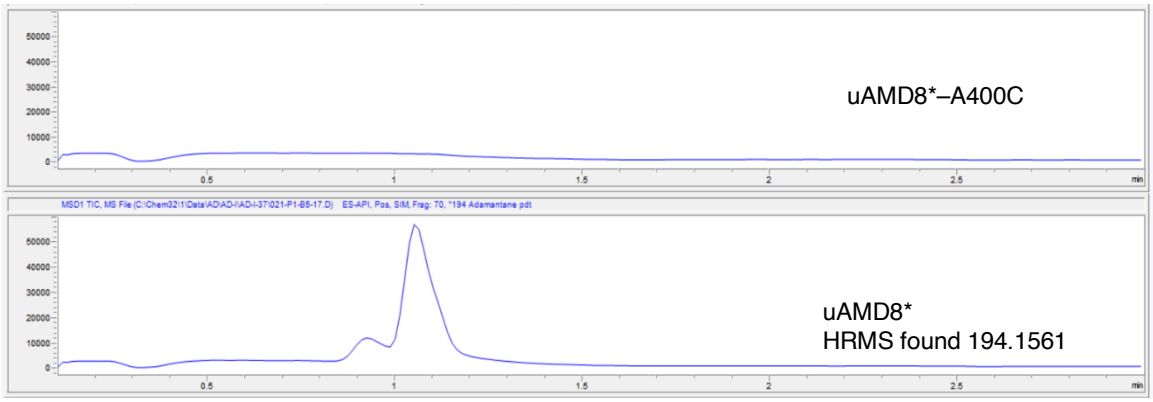
LCMS trace for SIM detection (M+1,212)

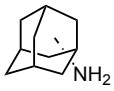


exact mass [M+H]: 194.1545

45

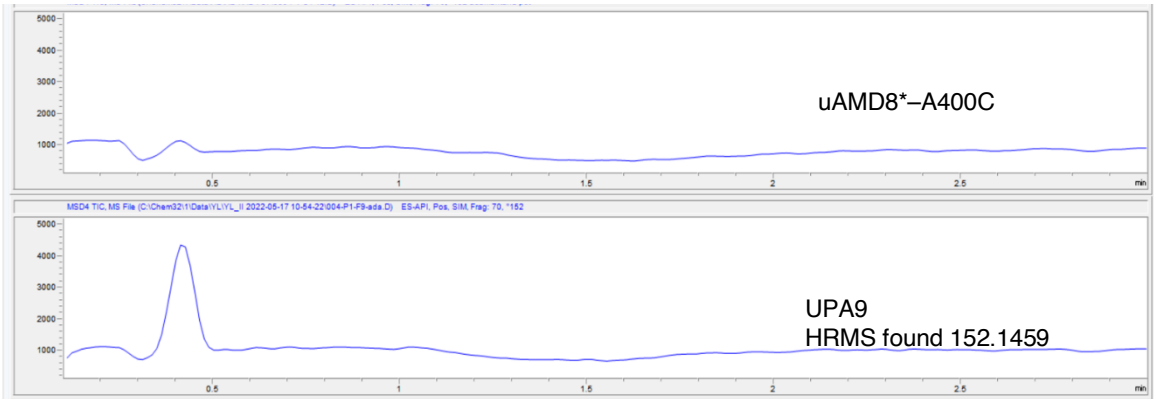
LCMS trace for SIM detection (M+1,194)

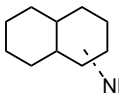




exact mass [M+H]: 152.1439

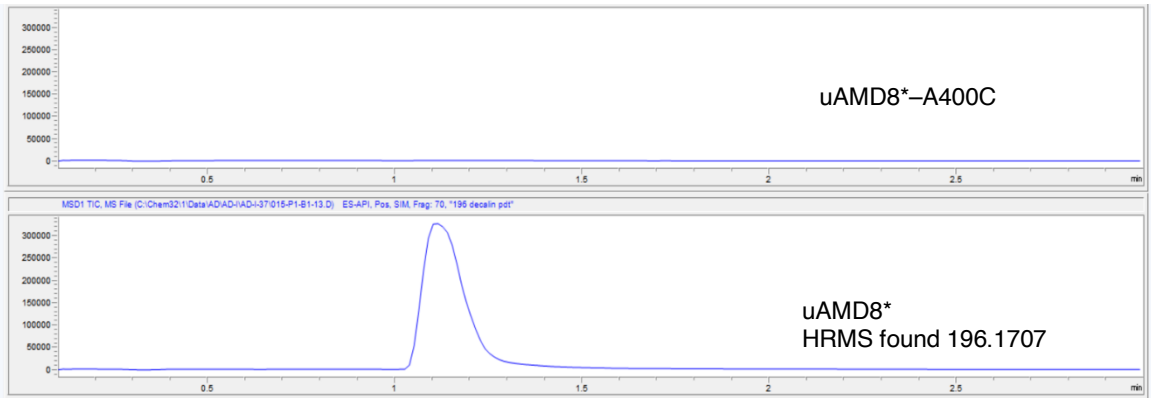
45 LCMS trace for SIM detection (M+1,152)

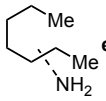




exact mass [M+H]: 196.1701

46 LCMS trace for SIM detection (M+1,196)

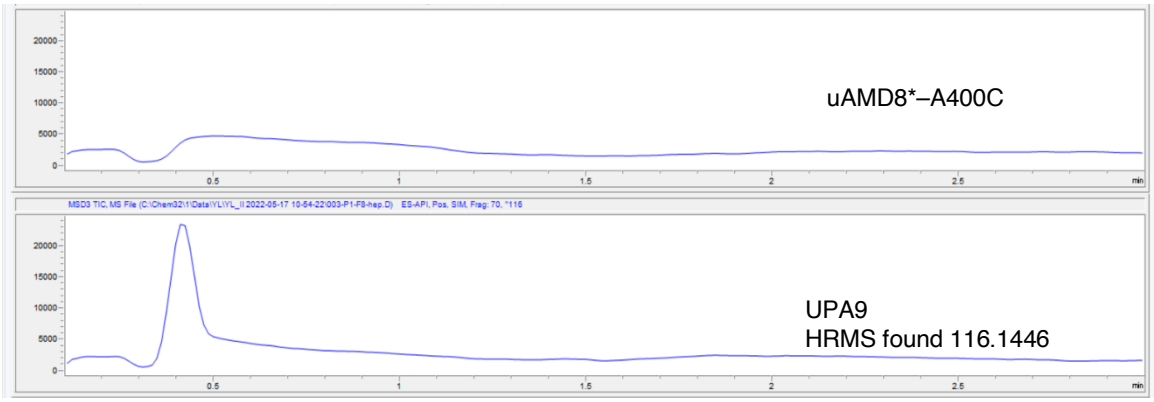




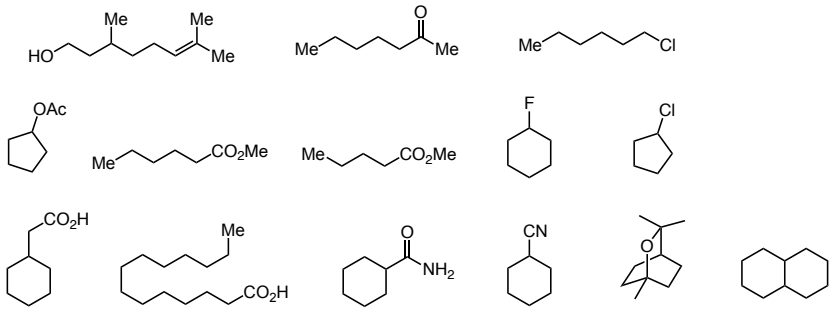
exact mass [M+H]: 116.1439

47

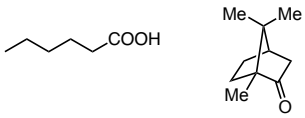
LCMS trace for SIM detection (M+1,116)



The following compounds failed to show any product peak with *uPA9* (incompetent for amination):

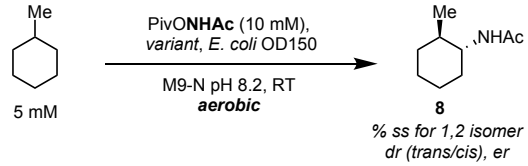


The following compounds failed to show any product peak with *uAMD8** (incompetent for amidation):



B.10. Miscellaneous Experiments

B.10.1. Effect on Amidation Selectivity by Mutations at Positions 74 and 438.



variant	74	438	%ss	dr	er
uAMD7-T438A	M	A	83	12	84
uAMD7-T438G	M	G	84	14	89
uAMD7-T438S	M	S	84	14	86
uAMD6-Q74A, V252E, T438S	A	S	86	11	83
uAMD6-Q74L, T438C	L	S	86	12	82
uAMD6-Q74L, T438C	L	C	86	8	83
uAMD6-Q74A, T438S	A	G	88	11	81
uAMD6-T438C	Q	C	89	6	80
uAMD6	Q	T	90	5	75
uAMD6-Q74L	L	T	90	6	80
uAMD7	M	T	91	8	85
uAMD6-V252T, T438S	Q	S	92	9	80
uAMD6-Q74A, V252E	A	T	93	6	80

B.10.2. Kinetic Isotope Effect Studies for Amidation of Cyclohexane/ d_{12}–Cyclohexane with *uAMD8

Competitive KIE

From a freshly streaked plate of *E. cloni* ® cells with the *uAMD8**[1] plasmid, three colonies were inoculated into three separate, 4-mL LB media tubes for overnight cultures at 37 ºC. These were utilized for triplicate experiments. Details for only one are described henceforth. Next day (~ after 16 hrs), 1 mL of the starter culture was transferred to 50 mL HB media (supplemented with Ampicillin and Glucose Mix) in a 150-mL Erlenmeyer flask and grown with vigorous shaking at 37 ºC for ~150 min, by which time the cell optical density (OD_{600}) was ~1.0. The flask was chilled on ice for 45 min. ALA (1 mM) and IPTG (0.5 mM) were supplemented, and the flask shaken at 22 ºC (110 rpm) for 20 hrs. The culture was centrifuged (3000*g*, 6 min), and the cell pellets were resuspended in 2 mL of M9-N buffer (pH 8.2) to give a suspension with OD_{600}~140. All manipulations were done on ice. A stock solution of cyclohexane, 200 mM, d_{12}–cyclohexane, 200 mM and **2**, 400 mM in isopropanol was prepared (substrate mix). A quench solution (QS) was made consisting of 1.8 mg AcNHPh (external standard) dissolved in 100 mL EtOH. The QS (200 mM) was transferred to l.5-mL labeled Eppendorf tubes that would be used as timepoints for reaction monitoring. The KIE reaction was setup as follows. To 1.47 mL of the cell suspension in a 2-mL GC vial, 37 mL of substrate mix were added at RT (designated as t=0), the vial capped and vigorously shaken. The vial was then shaken at RT (600 rpm). At

[1] Here the '*' is part of the variants name. uAmD8* is between uAmD8 and uAmD9. During evolution from uAmD8 to uAmD9, we found a variant that had higher activity but lower selectivity (i.e. uAmD8*). We decided to use this variant to test substrate scope.

regular intervals, 200 mL of the reaction mixture were added to the QS, the Eppendorf tube was vortexed and centrifuged (20000*g*, 5 min) and the supernatant analyzed by LCMS (4 mL injections) to quantify product formation based on single ion mass peak (M+1) for the non-deuterated and deuterated product.

Sample	Time (sec)	Dprod|Area	Hprod|Area	std|Area	H/D area	**Av H/D**	**sdev H/D**
A1	120	15614	153559	340424	9.83	**9.77**	**0.41**
A2	240	31463	311359	308763	9.90		
A3	360	45373	436472	300490	9.62		
A4	600	69979	676561	295403	9.67		
A5	1080	106966	1012180	296504	9.46		
A6	2160	141755	1351300	285555	9.53		
A7	4200	172446	1686070	299921	9.78		
A8	9120	208071	1982290	332809	9.53		
A9	34800	193188	1876590	330723	9.71		
B1	125	38011	359122	300243	9.45		
B2	250	72658	693164	287427	9.54		
B3	372	100253	959826	281898	9.57		
B4	610	134824	1240550	277635	9.20		
B5	1095	172491	1522490	251278	8.83		
B6	2175	190723	1780750	281797	9.34		
B7	4215	204808	2038880	291587	9.96		
B8	9135	238609	2276830	327910	9.54		
B9	34815	236967	2320640	339724	9.79		
C1	135	24239	236829	338319	9.77		
C2	260	45661	457996	342660	10.03		
C3	380	67106	650111	316086	9.69		
C4	620	97362	977082	311786	10.04		
C5	1105	118229	1230840	287797	10.41		
C6	2190	165801	1660790	302569	10.02		
C7	4230	191005	1916840	312615	10.04		
C8	9150	222960	2382270	349406	10.68		
C9	34830	252408	2721530	376160	10.78		

[A1-A9, B1-B9 and C1-C9 are the triplicate experiments. **Thus, competitive KIE is calculated as 9.77**]

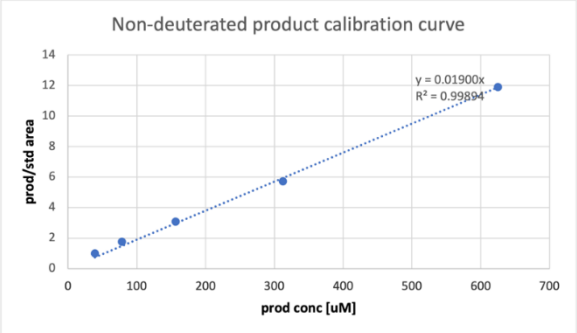
Non-Competitive KIE

A procedure similar to the competitive experiment was followed. Triplicate experiments (experiments 1-3) starting from three different single colonies were performed. Two separate substrate mixtures (SM) were prepared: SM1– cyclohexane, 200 mM, **2**, 400 mM and SM2 – d_{12}-cyclohexane, 200 mM, **2**, 400 mM.

The 2-mL M9 cell suspension (OD~140) was divided into two parts to set up separate reactions with SM1 and SM2 simultaneously. The procedure for SM1 is described henceforth. To 975 mL of the cell suspension in a 2-mL GC vial, 25 mL of substrate mix were added at RT (designated as t=0), the vial capped and vigorously shaken. The vial was then shaken at RT (600 rpm). At regular intervals, 200 mL of the reaction mixture were added to the QS, the Eppendorf tube was vortexed and centrifuged (20000*g*, 5 min), and the supernatant analyzed by LCMS (4 mL injections) to quantify product formation based on single ion mass peak (M+1) for product. A calibration curve with the non-deuterated product was used as a reference to calculate product concentrations. Similarly, analysis of the reaction with SM2 provided a time-course for the deuterated substrate. A ratio of initial rates for the two reactions provides the KIE.

Product calibration curve:

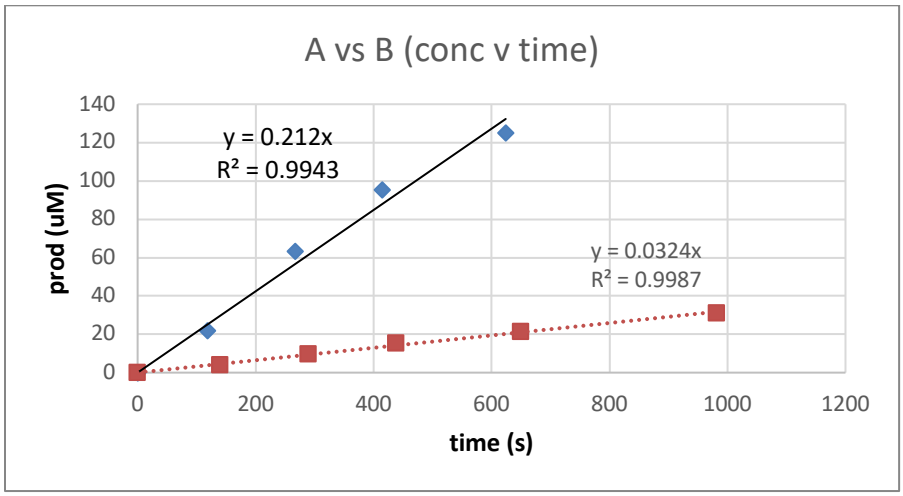
conc (uM)	H prod area	std area	prod/std area
39.0625	568110	572249	0.99276713
78.125	940332	531793	1.76822937
156.25	1493680	485063	3.07935258
312.5	2433890	424635	5.73172254
625	3870650	325345	11.8970631
1250	5145310	245541	20.9549933
2500	7739960	160968	48.0838428



experiment 1:

A: cyclohexane, B: d_{12}-cyclohexane

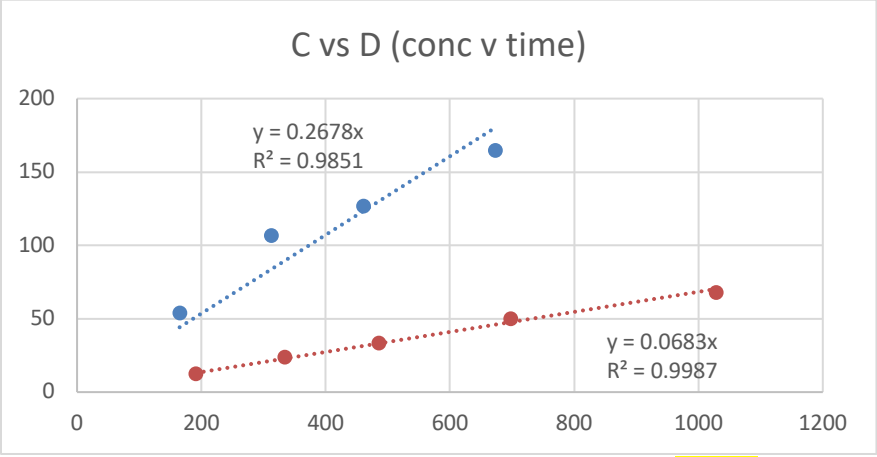
SAMPLE	Time (s)	PRODAREA	STD AREA	prod/std	prod(uM)
A0	**0**				**0**
A1	119	193201	459926	0.42006975	22.1089342
A2	267	535584	444787	1.20413591	63.3755743
A3	415	822859	453274	1.81536775	95.5456709
A4	624	1070740	450883	2.37476241	124.987495
A5	903	1360120	462087	2.9434284	154.917284
A6	1383	1632530	497945	3.27853478	172.554462
A7	2427	1823540	468015	3.89632811	205.0699
A8	5838	1883180	505282	3.7269881	196.157268
A9	15690	2376250	512340	4.63803334	244.107018
B0	0				0
B1	140	34737	473416	0.07337521	3.86185333
B2	290	86218	468740	0.18393566	9.68082407
B3	439	139483	478511	0.29149382	15.3417801
B4	650	199876	494790	0.40396128	21.2611198
B5	982	301339	509712	0.59119464	31.1155071
B6	1410	409515	509481	0.80378856	42.3046611
B7	2455	545415	501790	1.08693876	57.2073031
B8	5885	641452	538991	1.19009779	62.636726
B9	15727	861653	606297	1.42117312	74.7985853



KIE from initial rates of the curves = 0.211/0.0324 = **6.54**

Experiment 2:
C: cyclohexane, D: d_{12}-cyclohexane

SAMPLE	Time (s)	PRODAREA	STD AREA	prod/std	prod(uM)
C0	0				0
C1	165	528980	516243	1.02467249	53.930131
C2	313	954835	471239	2.02622236	106.643282
C3	461	1175110	488219	2.40693213	126.680639
C4	673	1496870	477398	3.13547606	165.025056
C5	932	1699120	469172	3.62152899	190.606789
C6	1440	1992890	470799	4.2329954	222.789231
C7	2486	2193960	481557	4.55597157	239.787977
C8	5920	2283890	499877	4.56890395	240.468629
C9	15763	2572640	520488	4.94274604	260.144528
D0	0				0
D1	191	121621	504918	0.24087278	12.6775145
D2	334	226710	498751	0.45455548	23.9239726
D3	486	331068	519897	0.63679537	33.5155455
D4	698	472311	496071	0.95210363	50.1107174
D5	1028	667972	515506	1.29575989	68.1978892
D6	1469	801772	509738	1.57291	82.7847371
D7	2520	986491	525225	1.87822552	98.8539749
D8	5958	1159000	541760	2.13932369	112.595983
D9	15796	1368940	590380	2.31874386	122.039151

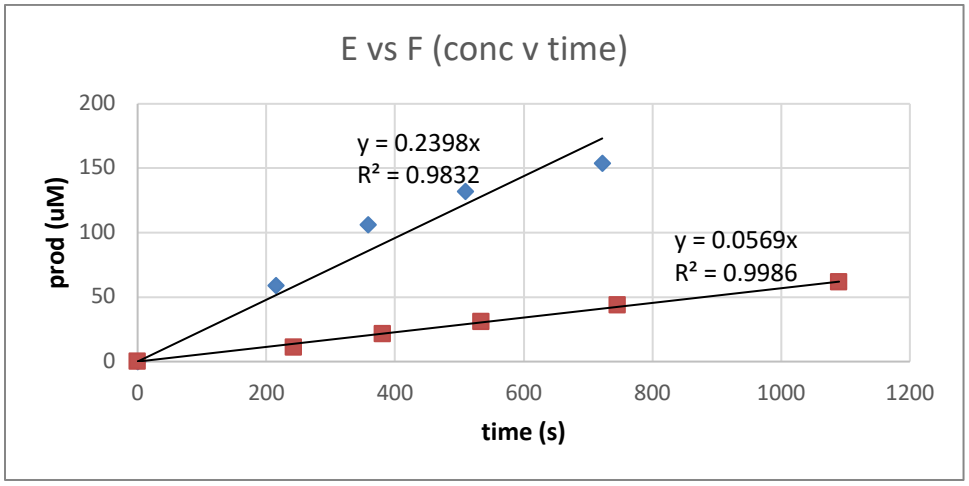


KIE from initial rates of the curves = 0.2678/0.0683 = **3.90**

Experiment 3:

E: cyclohexane, F: d_{12}-cyclohexane

SAMPLE	Time (s)	PRODAREA	STD AREA	prod/std	Time	prod(uM)
E0	0				0	0
E1	215	556271	496707	1.11991778	215	58.943041
E2	358	951749	472267	2.01527737	358	106.06723
E3	509	1219040	485683	2.50994991	509	132.102627
E4	722	1422630	486724	2.92286799	722	153.835157
E5	1058	1695110	495765	3.41918046	1058	179.956866
E6	1499	1805960	496569	3.63687624	1499	191.414539
E7	2557	1948400	548690	3.5510033	2557	186.89491
E8	5985	2313480	510096	4.53538158	5985	238.704293
E9	15830	2629970	538538	4.88353654	15830	257.028239
F0	0				0	0
F1	243	105928	506184	0.20926778	243	11.0140935
F2	381	204107	500354	0.40792519	381	21.4697468
F3	534	302150	519932	0.58113369	534	30.5859835
F4	746	435885	523491	0.83265042	746	43.8237062
F5	1090	599705	511560	1.17230628	1090	61.7003305
F6	1529	776861	522011	1.4882081	1529	78.3267423
F7	2591	946448	543038	1.74287619	2591	91.7303257
F8	6025	1079010	549171	1.96479785	6025	103.410413
F9	15888	1231810	624247	1.9732734	15888	103.856495



KIE from initial rates of the curves = 0.2398/0.0659 = **4.20**

Thus, KIE values of 6.54, 3.90 and 4.20 are calculated from experiments 1–3, giving an average value of 4.8 ± 1.4

B.10.3. Synthesis of Standard Compounds

In all cases, commercially available amines (Sigma, Enamine, Ambeed, Combi Blocks) were utilized as amine standards without further purification. These commercially available amines were acetylated with acetyl chloride to give the corresponding *N*-acetyl amides in 40-80% yield as white solids or colorless oil, for use as amide standards.[1] Nitrene precursors **1** and **2** were synthesized according to literature procedures.[1,2]

AcNHOPiv (2)
^{1}H NMR (δ, 23 °C, $CDCl_3$): 9.29 (NH, br, 1H), 2.00 (s, 3H), 1.27 (s, 9H). ^{13}C NMR ((δ, 23 °C, $CDCl_3$): 176.8, 168.4, 38.4, 27.2, 27.1, 19.8. *The other peaks correspond to residual EtOAc. Spectral data are well-matched with those available in the literature.[1]

1-*N*-acetylamino-2-methylcyclohexane (7[1,2])
^{1}H NMR (δ, 23 °C, $CDCl_3$): 5.46 (NH, br, 0.4H), 5.24 (NH, br, 1H), 4.08–4.02 (m, 0.4H), 3.52–3.45 (m, 1H), 2.00 (s, 1H), 1.99 (s, 3H), 1.97–1.91 (m, 1H), 1.89–1.45 (m, 8H), 1.39–1.28 (m, 2H), 1.24–0.97 (m, 5H), 0.92 (d, J = 8 Hz, 3H), 0.86 (d, J = 8 Hz, 1H). *The fractional integrals represent the minor diastereomer, as the standard product is a 3:1 diastereomeric mixture.

1-*N*-acetylamino-2-methylcyclohexane (ent-8) from calB
The product was synthesized based on the procedure described in section 5.1. ^{1}H NMR (δ, 23 °C, $CDCl_3$): 5.20 (NH, br, 1H), 3.52–3.45 (m, 1H), 1.98 (s, 3H), 1.97–1.93 (m, 1H), 1.76–1.64 (m, 4H), 1.39–1.29 (m, 1H), 1.23–1.14 (m, 2H), 1.12–1.01 (m, 2H), 0.93 (d, J = 4.8 Hz, 3H).

1-*N*-acetylamino-2-ethylcyclohexane (10[1,2])
^{1}H NMR (δ, 23 °C, $CDCl_3$): 6.07 (NH, br, 1H), 5.88 (NH, br, 1H), 4.20–4.14 (m, 1H), 3.61–3.52 (m, 1H), 2.05 (d, J = 3.2 Hz, 6H), 1.94–1.83 (m, 2H), 1.75–0.96 (m, 20H), 0.89–0.83 (m, 6H). The standard product is 1:1 diastereomeric mixture.

1-*N*-acetylamino-2-ethylcyclohexane (ent-10) from calB
The product was synthesized based on the procedure described in section 5.3. ^{1}H NMR (δ, 23 °C, $CDCl_3$): 5.83 (NH, br, 1H), 3.62–3.53 (m, 1H), 2.05 (s, 3H), 1.97–1.91(m, 1H), 1.89–1.85 (m, 1H), 1.74–1.68 (m, 2H), 1.60–1.51 (m, 1H), 1.38–1.24 (m, 2H), 1.19–1.06 (m, 4H), 1.05–0.97 (m, 1H), 0.86 (t, J = 7.2 Hz, 3H).

1-*N*-acetylamino-3-methylcyclohexane (7[1,3])

^{1}H NMR (δ, 23 °C, $CDCl_3$): 6.48 (NH, br, 1H), 4.16–4.11 (m, 1H), 3.81–3.71 (m, 0.3H), 2.10 (d, J = 0.8 Hz, 4H), 1.94–1.90 (m, 0.5H), 1.77–1.55 (m, 7H), 1.49–1.30 (m, 3H), 1.08–1.04 (m, 1H), 0.93–0.89 (m, 4H), 0.84–0.79 (m, 1H). *The fractional integrals represent the minor diastereomer, as the standard product is a 3:1 diastereomeric mixture. Spectral data match those reported in the literature.[9]

1-*N*-acetylamino-3-methylcyclohexane (6-*N*-*Ac*) from calB

The product was synthesized based on the procedure described in section 5.2. ^{1}H NMR (δ, 23 °C, $CDCl_3$): 5.91 (NH, br, 1H), 3.79–3.69 (m, 1H), 2.00 (s, 3H), 1.95–1.90 (m, 2H), 1.76–1.71 (m, 1H), 1.66–1.60 (m, 1H), 1.51–1.44 (m, 1H), 1.39–1.29 (m, 1H), 1.05–0.98 (m, 1H), 0.89 (d, J = 8.0 Hz, 3H), 0.83–0.71 (m, 2H). Spectral data match those reported in the literature.[9]

1-*N*-acetylamino-3-ethylcyclohexane (10[1,3])

^{1}H NMR (δ, 23 °C, $CDCl_3$): 5.69 (NH, br, 1H), 5.52 (NH, br, 1H), 4.14–4.09 (m, 1H), 3.77–3.70 (m, 1H), 2.04 (s, 1H), 2.00 (s, 3H), 1.98 (s, 3H), 1.78–1.55 (m, 7H), 1.39–1.20 (m, 10H), 1.06–0.94 (m, 2H), 0.89–0.84 (m, 6H), 0.80–0.66 (m, 2H). The standard product is 1:1 diastereomeric mixture.

1-*N*-acetylamino-3-ethylcyclohexane (9-*N*-*Ac*) from calB

The product was synthesized based on the procedure described in section 5.3. ^{1}H NMR (δ, 23 °C, $CDCl_3$): 5.84 (NH, br, 1H), 3.77–3.71 (m, 1H), 2.02 (s, 3H), 1.97–1.92 (m, 2H), 1.79–1.69 (m, 2H), 1.39–1.19 (m, 4H), 1.07–0.97 (m, 1H), 0.87 (t, J = 7.2 Hz, 3H), 0.81–0.69 (m, 2H).

1-*N*-acetylamino-2-methylcyclopentane (12[1,2])

^{1}H NMR (δ, 23 °C, $CDCl_3$): 5.45 (NH, br, 2H), 4.27–4.23 (m, 1H), 3.79–3.75 (m, 1H), 2.16–2.03 (m, 3H), 2.00 (d, J = 2.8 Hz, 6H), 1.97–1.79 (m, 3H), 1.73–1.56 (m, 6H), 1.48–1.41 (m, 1H), 1.36–1.20 (m, 3H), 1.02 (d, J = 6.8 Hz, 3H), 0.88 (d, J = 6.8 Hz, 3H).

1-*N*-acetylamino-3-methylcyclopentane (12[1,3])

^{1}H NMR (δ, 23 °C, $CDCl_3$): 5.69 (NH, br, 2H), 4.30–4.17 (m, 2H), 2.26–2.14 (m, 7H), 1.97 (s, 6H), 1.94–1.75 (m, 3H), 1.65–1.50 (m, 2H), 1.47–1.32 (m, 2H), 1.27–1.11 (m, 2H), 1.01 (d, J = 6.4 Hz, 3H), 0.99 (d, J = 6.8 Hz, 3H), 0.95–0.87 (m, 1H).

2-*N*-acetylaminopentane (19[b])

^{1}H NMR (δ, 23 °C, $CDCl_3$): 5.40 (NH, br, 1H), 4.00–3.93 (m, 1H), 1.95 (s, 3H), 1.42–1.29 (m, 4H), 1.10 (d, J = 6.4 Hz, 3H), 0.89 (t, J = 6.8 Hz, 3H). Spectral data match those reported in the literature.[10]

3-*N*-acetylaminopentane (19[a])

^{1}H NMR (δ, 23 °C, $CDCl_3$): 5.18 (NH, br, 1H), 3.83–3.74 (m, 1H), 1.99 (s, 3H), 1.59–1.49 (m, 2H), 1.41–1.30 (m, 2H), 0.89 (t, J = 7.2 Hz, 6H). Spectral data match those reported in the literature.[11]

2-*N*-acetylaminohexane (20^b)
^{1}H NMR (δ, 23 °C, $CDCl_3$): 5.47 (NH, br, 1H), 3.97–3.90 (m, 1H), 1.94 (s, 3H), 1.42–1.39 (m, 2H), 1.31–1.25 (m, 4H), 1.09 (d, J = 6.4 Hz, 3H), 0.87 (t, J = 8.0 Hz, 3H).

3-*N*-acetylaminohexane (20^a)
^{1}H NMR (δ, 23 °C, $CDCl_3$): 5.20 (NH, br, 1H), 3.90–3.83 (m, 1H), 1.99 (s, 3H), 1.55–1.26 (m, 6H), 0.92–0.87 (m, 6H).

3-*N*-acetylaminoheptane (21^b)
^{1}H NMR (δ, 23 °C, $CDCl_3$): 5.37 (NH, br, 1H), 3.85–3.80 (m, 1H), 1.98 (d, J = 3.2 Hz, 3H), 1.57–1.45 (m, 2H), 1.39–1.24 (m, 6H), 0.90–0.85 (m, 6H). Spectral data match those reported in the literature.[12]

4-*N*-acetylaminoheptane (21^a)
^{1}H NMR (δ, 23 °C, $CDCl_3$): 5.34 (NH, br, 1H), 3.96–3.90 (m, 1H), 1.99 (s, 3H), 1.48–1.24 (m, 8H), 0.90 (t, J = 7.2 Hz, 6H).

2-*N*-acetylamino-3-methylpentane (22^b)
^{1}H NMR (δ, 23 °C, $CDCl_3$): 5.45 (NH, br, 1H), 4.00–3.89 (m, 1H), 1.96 (s, 3H), 1.45–1.37 (m, 2H), 1.12–1.02 (m, 4H), 0.90–0.83 (m, 6H).

3-*N*-acetylamino-3-methylpentane (22^a)
^{1}H NMR (δ, 23 °C, $CDCl_3$): 5.17 (NH, br, 1H), 1.92 (s, 3H), 1.81–1.72 (m, 2H), 1.64–1.55 (m, 2H), 1.19 (s, 3H), 0.80 (t, J = 8.0 Hz, 6H).

1-*N*-acetylamino-1-methylcyclohexane (7^{1,1})
^{1}H NMR (δ, 23 °C, $CDCl_3$): 5.76 (NH, br, 1H), 2.02 (s, 3H), 2.00–1.96 (m, 2H), 1.52–1.40 (m, 7H), 1.36 (s, 3H), 1.31–1.29 (m, 1H). Spectral data match those reported in the literature.[13]

1-*N*-acetylamino-1-ethylcyclohexane (10^{1,1})
^{1}H NMR (δ, 23 °C, $CDCl_3$): 5.09 (NH, br, 1H), 2.01–1.97 (m, 2H), 1.96 (s, 3H), 1.77 (q, J = 7.2 Hz, 2H), 1.54–1.30 (m, 8H), 0.78 (t, J = 7.2 Hz, 3H). Spectral data match those reported in the literature.[14]

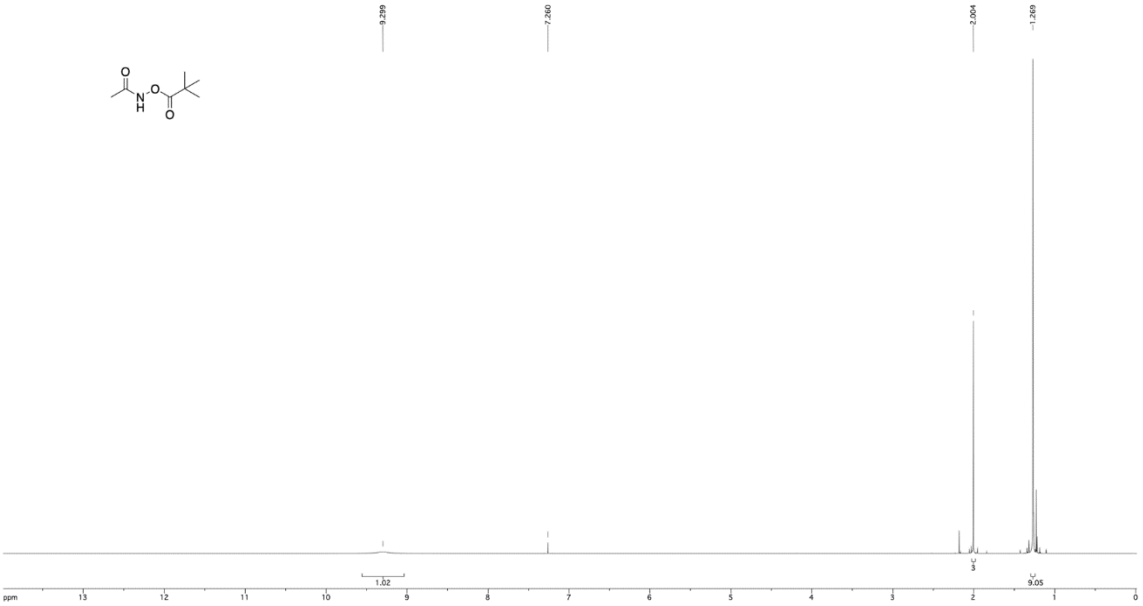
1-*N*-acetylamino-4-methylcyclohexane (7^{1,4})
^{1}H NMR (δ, 23 °C, $CDCl_3$): 6.24 (NH, br, 1H), 6.15 (NH, br, 1H), 4.00–3.98 (m, 1H), 3.70–3.66 (m, 1H), 2.04 (d, J = 3.6 Hz, 6H), 1.95–1.92 (m, 2H), 1.72–1.55 (m, 10H), 1.36–1.26 (m, 1H), 1.21–1.00 (m, 6H), 0.92 (d, J = 6.4 Hz, 3H), 0.88 (d, J = 6.4 Hz, 3H). The

standard product is 1:1 diastereomeric mixture. Spectral data match those reported in the literature.[15]

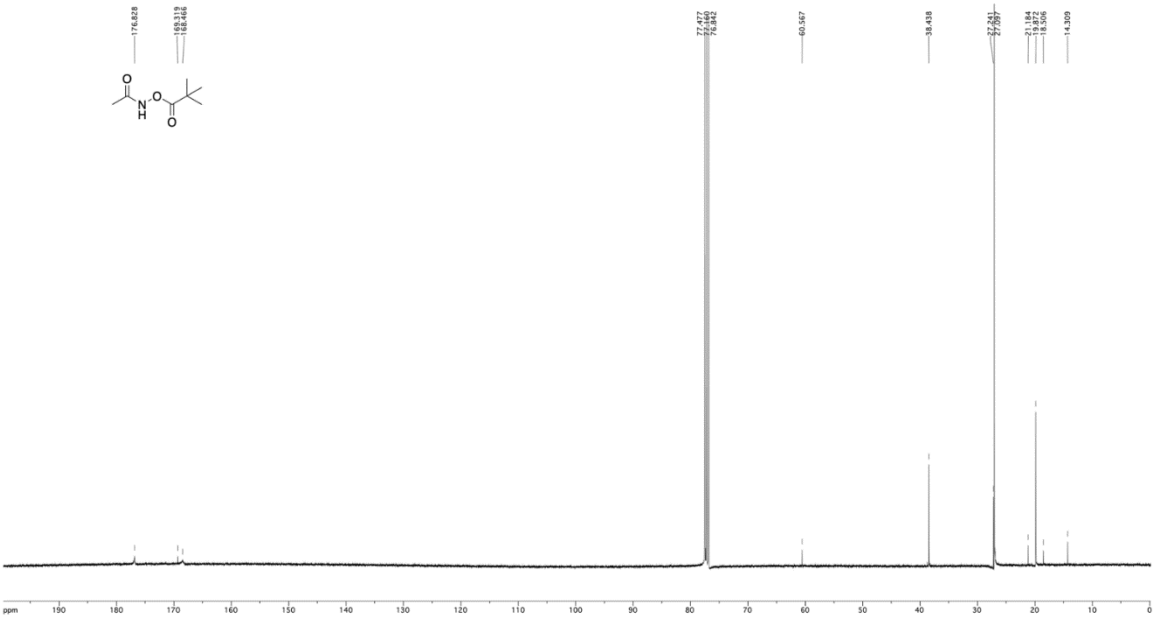
1-*N*-acetylamino-4-ethylcyclohexane ($10^{1,4}$)
^{1}H NMR (δ, 23 °C, $CDCl_3$): 5.65 (NH, br, 1H), 5.44 (NH, br, 1H), 4.02–3.98 (m, 1H), 3.70–3.65 (m, 1H), 1.95 (d, J = 11.6 Hz, 7H), 1.77–1.73 (m, 2H), 1.63–1.54 (m, 5H), 1.28–0.96 (m, 12H), 0.88–0.83 (m, 6H). The standard product is 1:1 diastereomeric mixture.

***N*-(cyclohexylmethyl)acetamide (7^{Me})**
^{1}H NMR (δ, 23 °C, $CDCl_3$): 5.57 (NH, br, 1H), 3.09 (d, J = 6.4 Hz, 2H), 1.99 (s, 3H), 1.74–1.64 (m, 6H), 1.50–1.40 (m, 1H), 1.28–1.09 (m, 3H), 0.97–0.87 (m, 2H). Spectral data match those reported in the literature. Spectral data match those reported in the literature.[16]

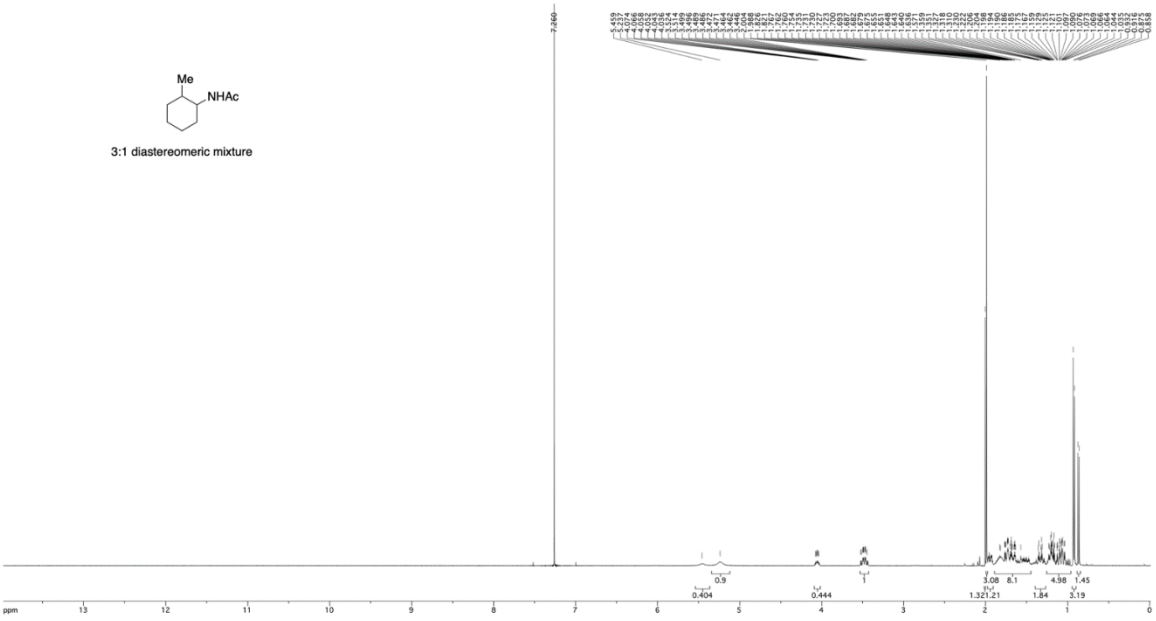
B.10.4. NMR Spectra of Standard Compounds



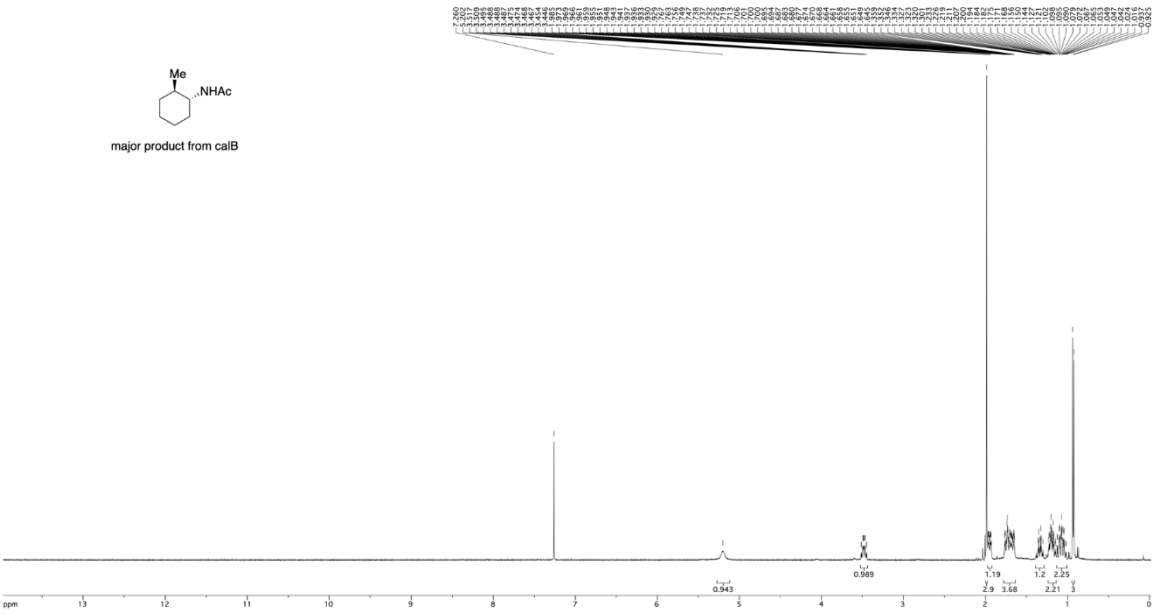
^{1}H-NMR spectrum of nitrene precursor **2** (AcNHOPiv) obtained in $CDCl_3$ at 23 °C.



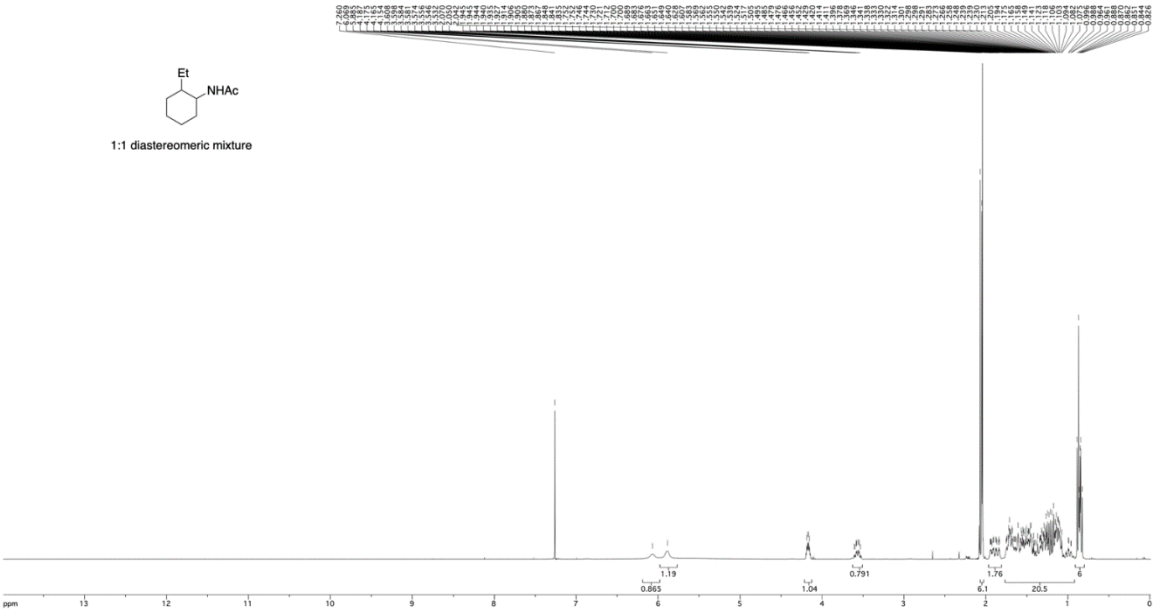
^{13}C-NMR spectrum of nitrene precursor **2** (AcNHOPiv) obtained in $CDCl_3$ at 23 °C.



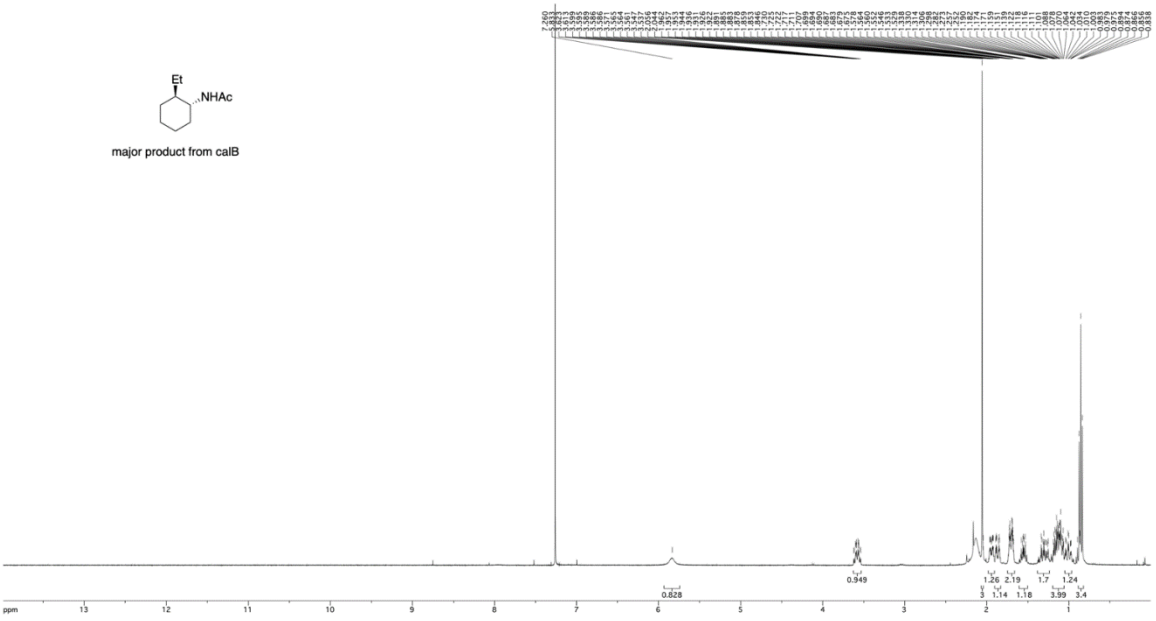
^{1}H-NMR spectrum of the diastereomeric mixture of 1-*N*-acetylamino-2-methylcyclohexane **7**[1,2] obtained in $CDCl_3$ at 23 °C.



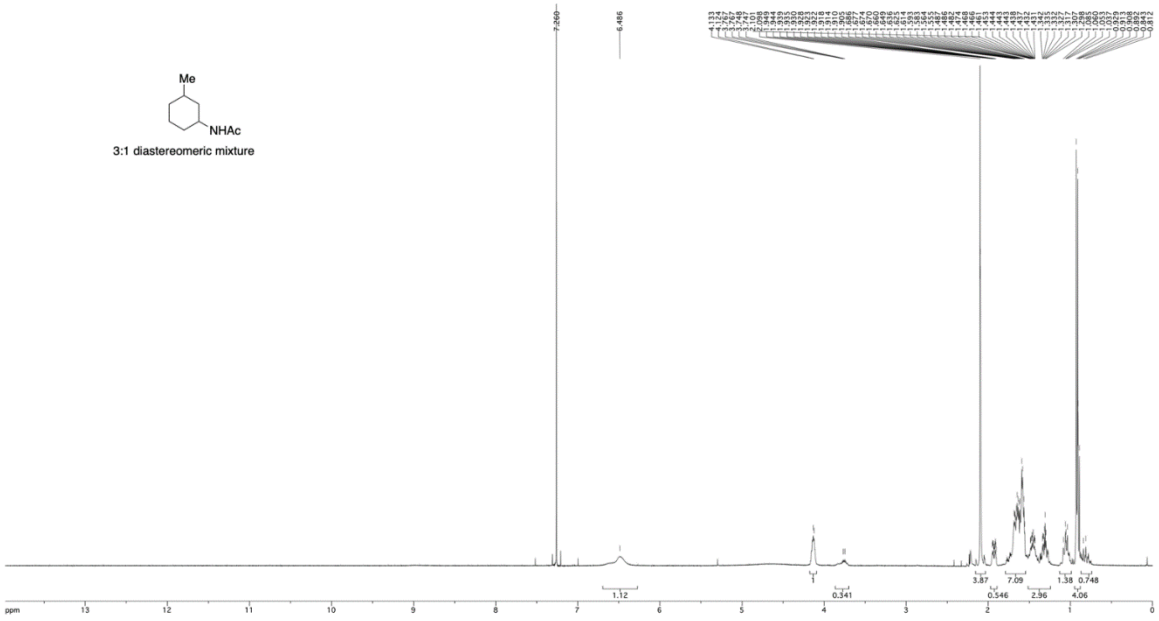
^{1}H-NMR spectrum of 1-*N*-acetylamino-2-methylcyclohexane **ent-8** synthesized using calB, obtained in $CDCl_3$ at 23 °C.



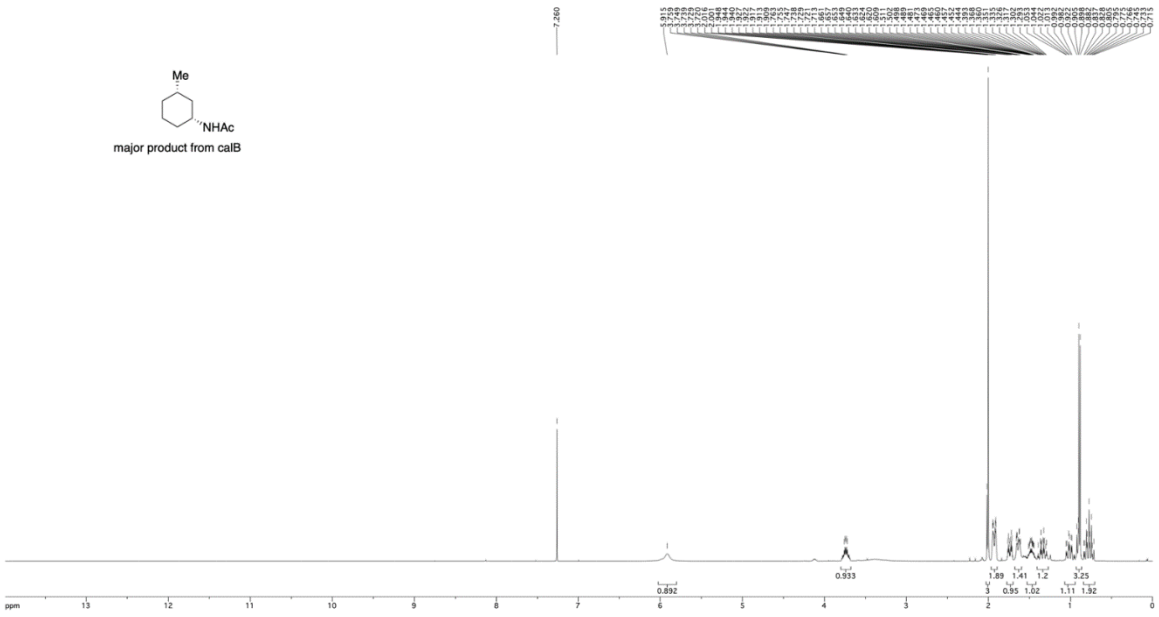
^{1}H-NMR spectrum of the diastereomeric mixture of 1-*N*-acetylamino-2-ethylcyclohexane **10**[1,2] obtained in $CDCl_3$ at 23 °C.



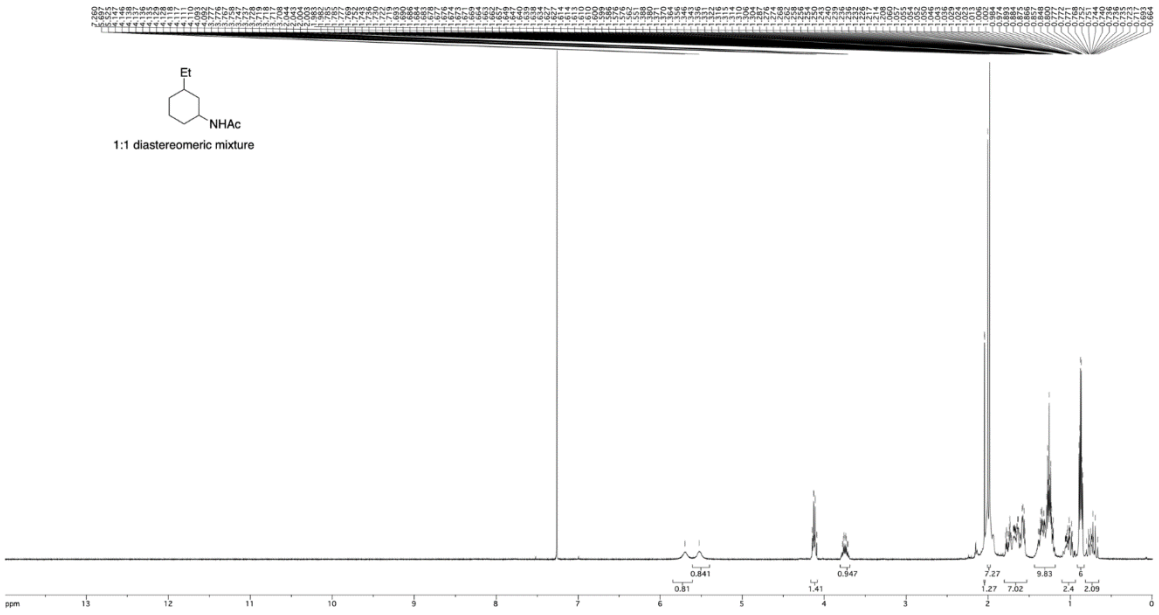
^{1}H-NMR spectrum of 1-*N*-acetylamino-2-ethylcyclohexane **ent-10** synthesized using calB, obtained in $CDCl_3$ at 23 °C.



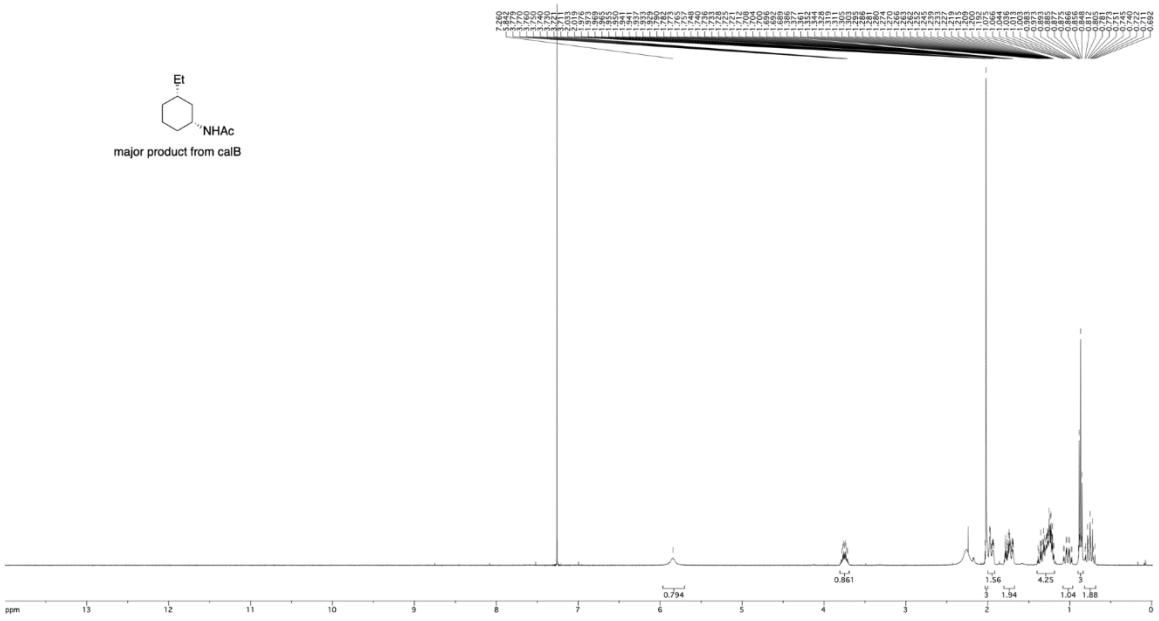
^{1}H-NMR spectrum of the diastereomeric mixture of 1-*N*-acetylamino-3-methylcyclohexane **7**^{1,3} obtained in $CDCl_3$ at 23 °C.



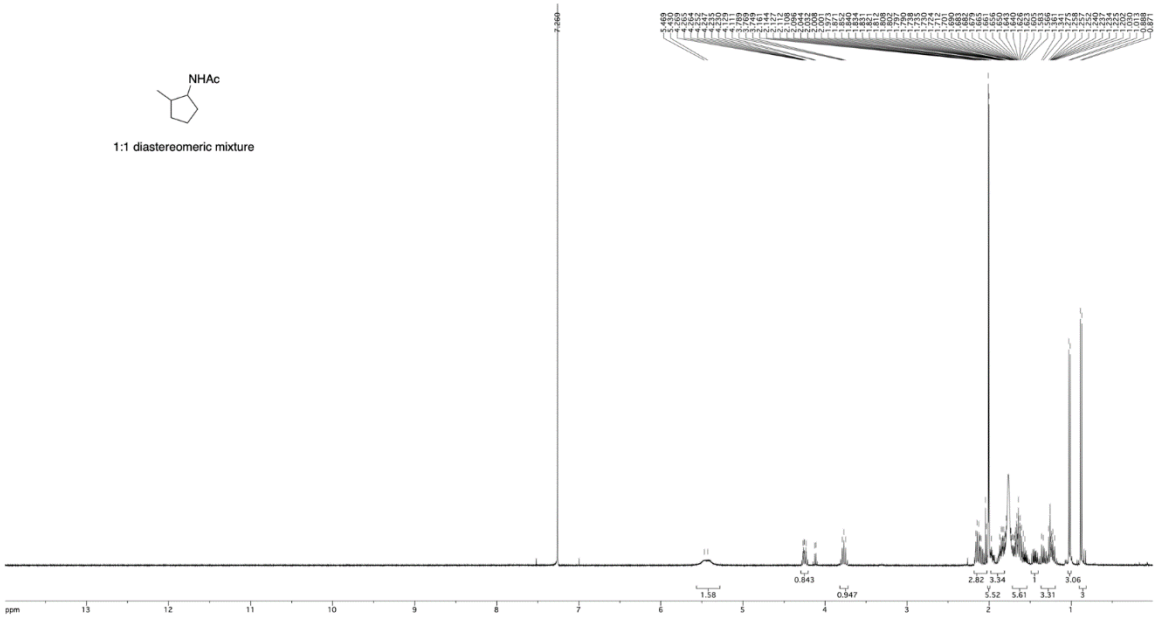
^{1}H-NMR spectrum of 1-*N*-acetylamino-3-methylcyclohexane **6-*N-Ac*** synthesized using calB, obtained in $CDCl_3$ at 23 °C.



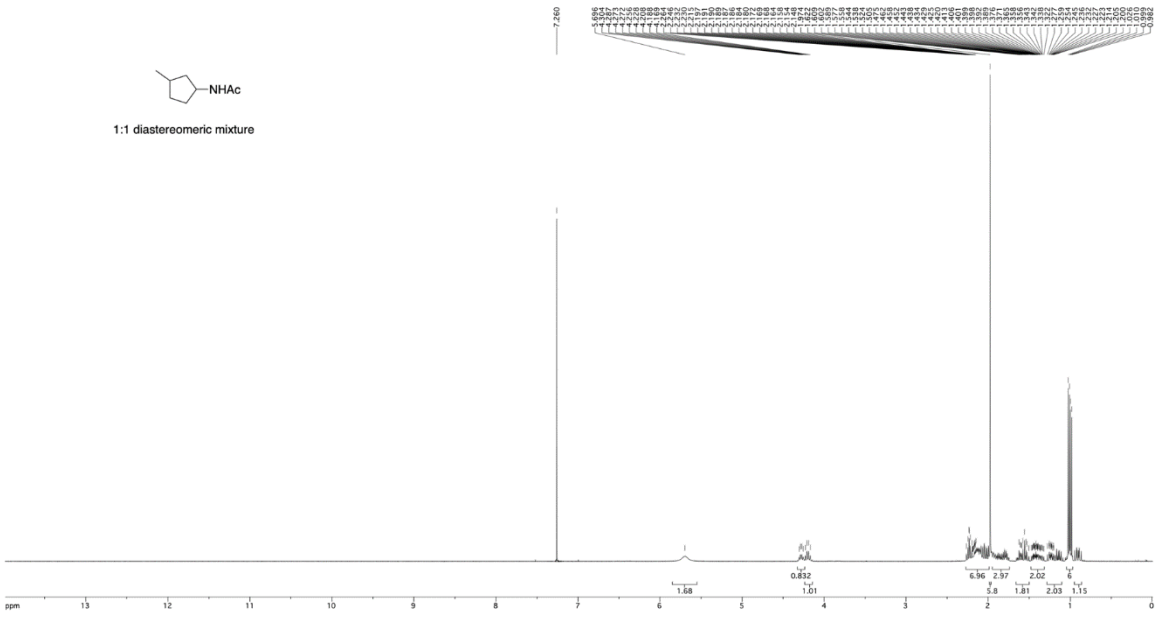
^{1}H-NMR spectrum of the diastereomeric mixture of 1-*N*-acetylamino-3-ethylcyclohexane **10**[1,3] obtained in $CDCl_3$ at 23 °C.



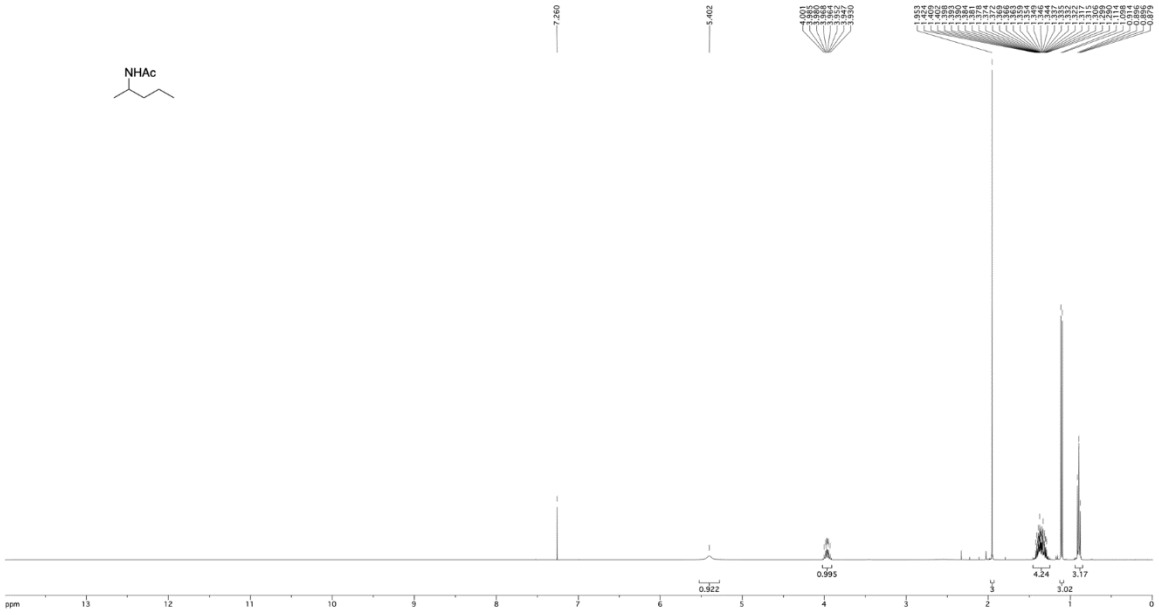
^{1}H-NMR spectrum of 1-*N*-acetylamino-3-ethylcyclohexane **9-*N-Ac*** synthesized using calB, obtained in $CDCl_3$ at 23 °C.



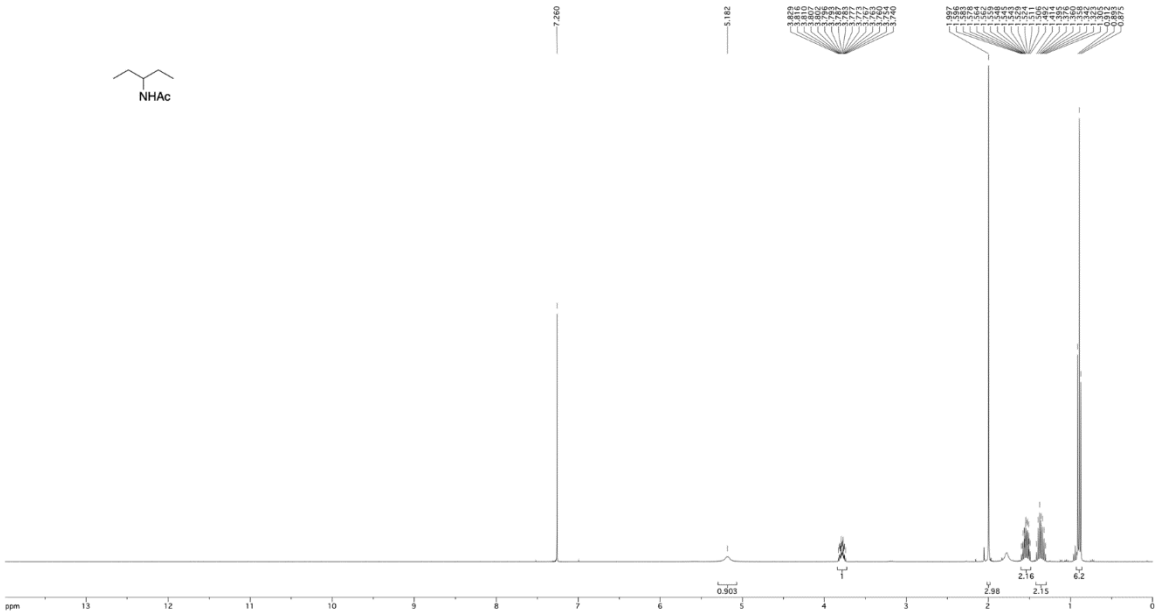
^{1}H-NMR spectrum of the diastereomeric mixture of 1-*N*-acetylamino-2-methylcyclopentane **12**^{1,2} obtained in $CDCl_3$ at 23 °C.



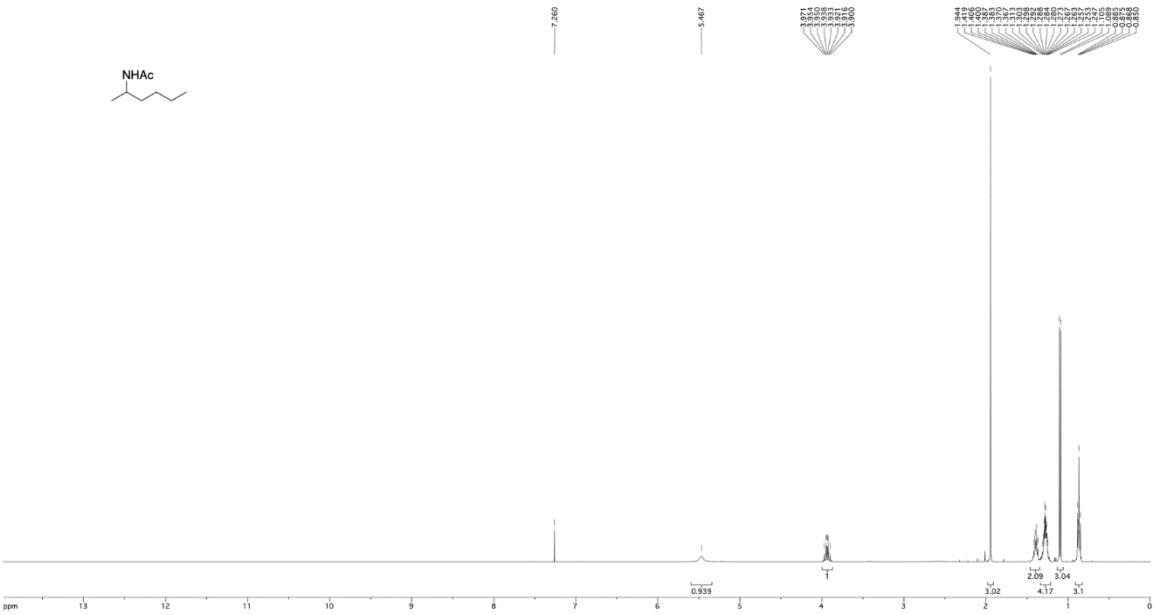
^{1}H-NMR spectrum of the diastereomeric mixture of 1-*N*-acetylamino-3-methylcyclopentane **12**^{1,3} obtained in $CDCl_3$ at 23 °C.



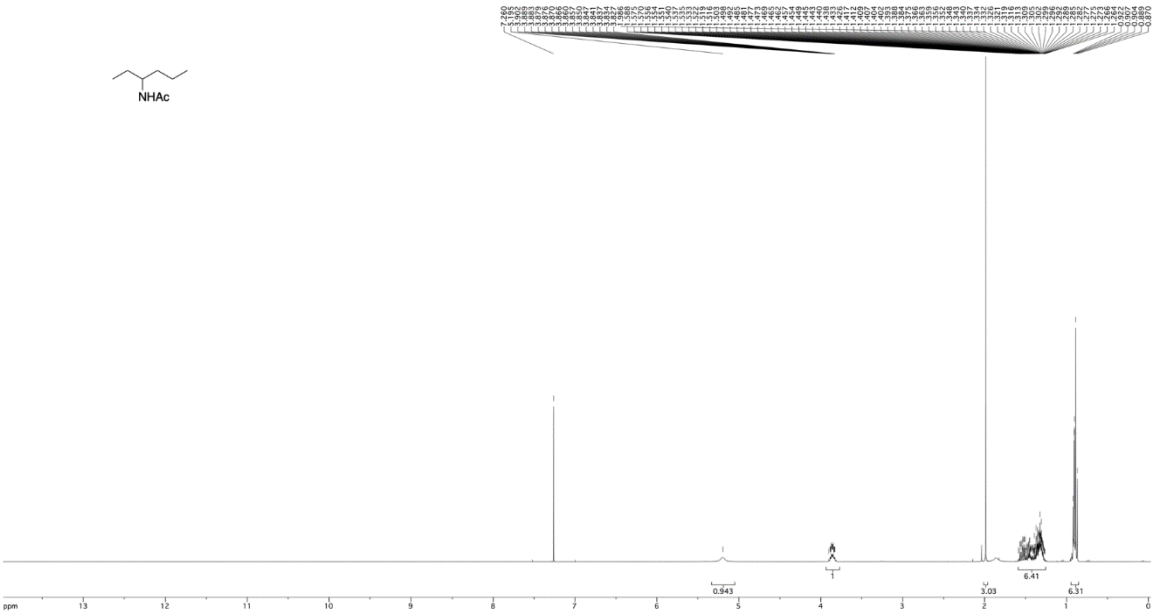
^{1}H-NMR spectrum of 2-*N*-acetylaminopentane **19^{b}** obtained in $CDCl_3$ at 23 °C.



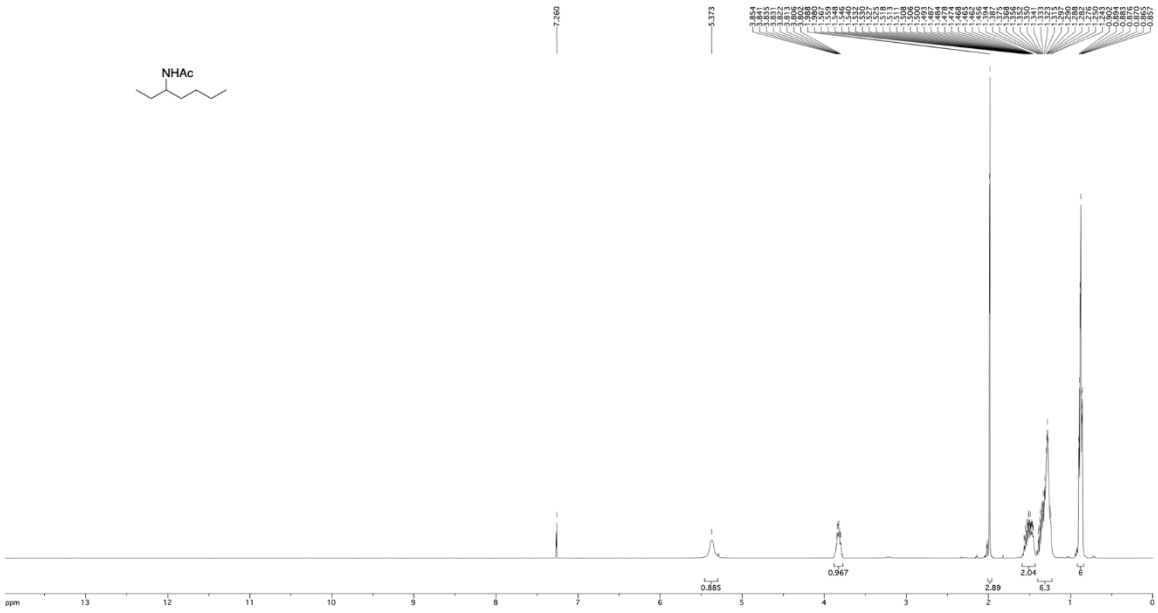
^{1}H-NMR spectrum of 3-*N*-acetylaminopentane **19^{a}** obtained in $CDCl_3$ at 23 °C.



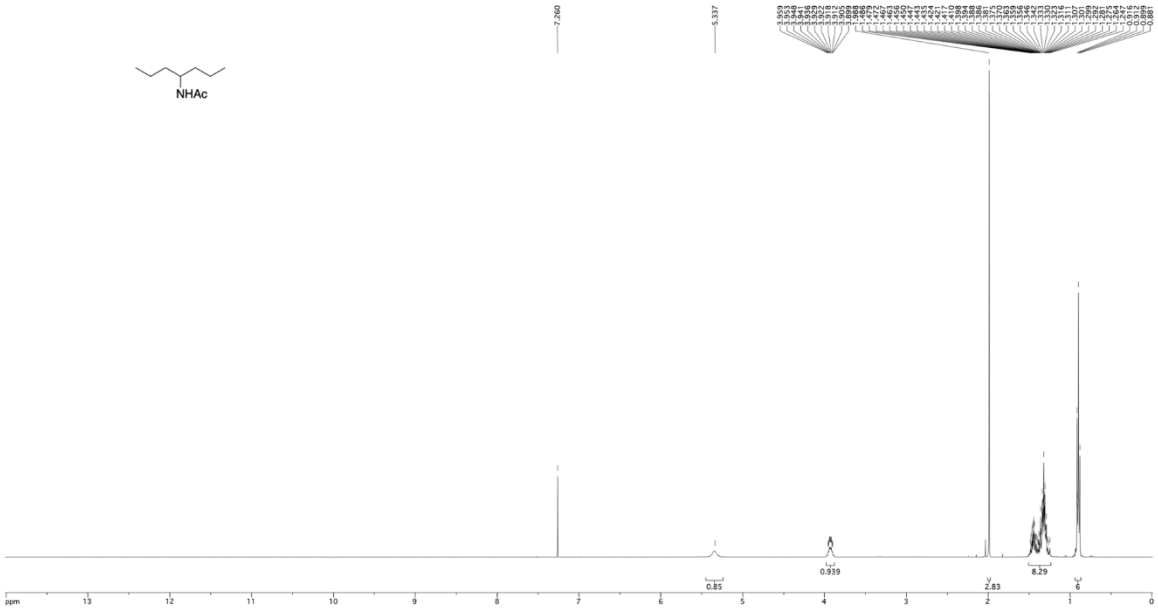
^{1}H-NMR spectrum of 2-*N*-acetylaminohexane **20^{b}** obtained in $CDCl_3$ at 23 °C.



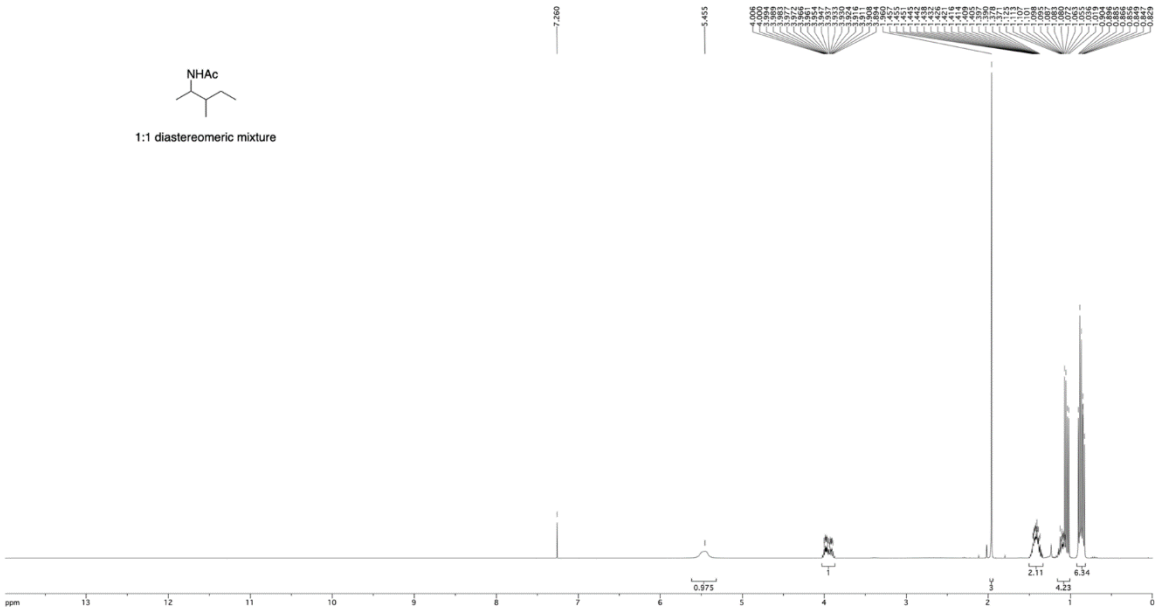
^{1}H-NMR spectrum of 3-*N*-acetylaminohexane **20^{a}** obtained in $CDCl_3$ at 23 °C.



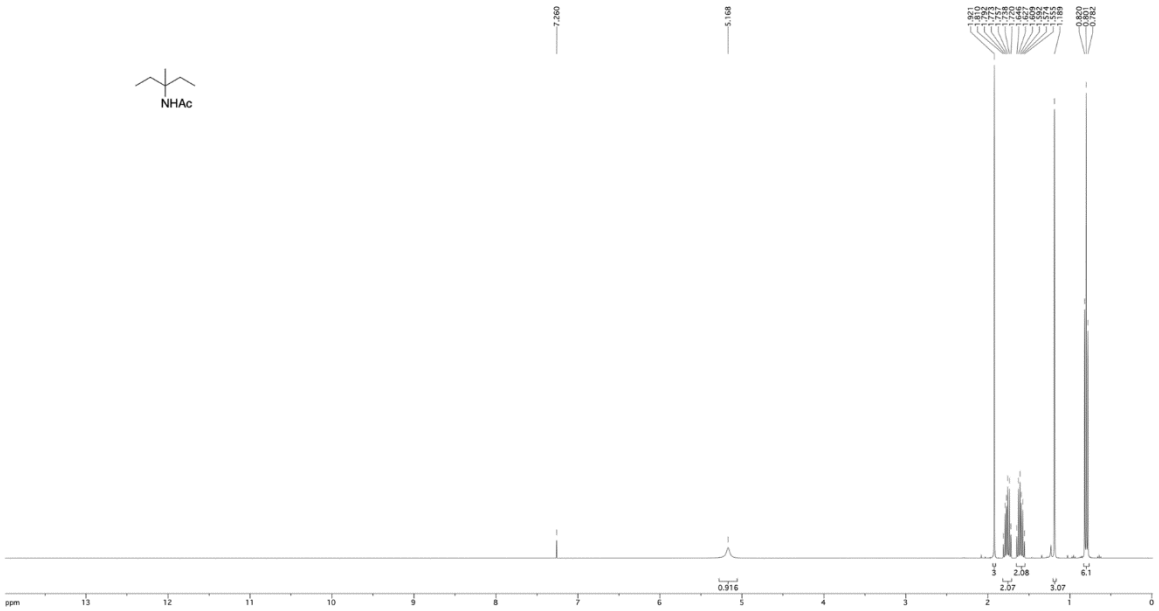
^{1}H-NMR spectrum of 3-*N*-acetylaminoheptane **21^b** obtained in $CDCl_3$ at 23 °C.



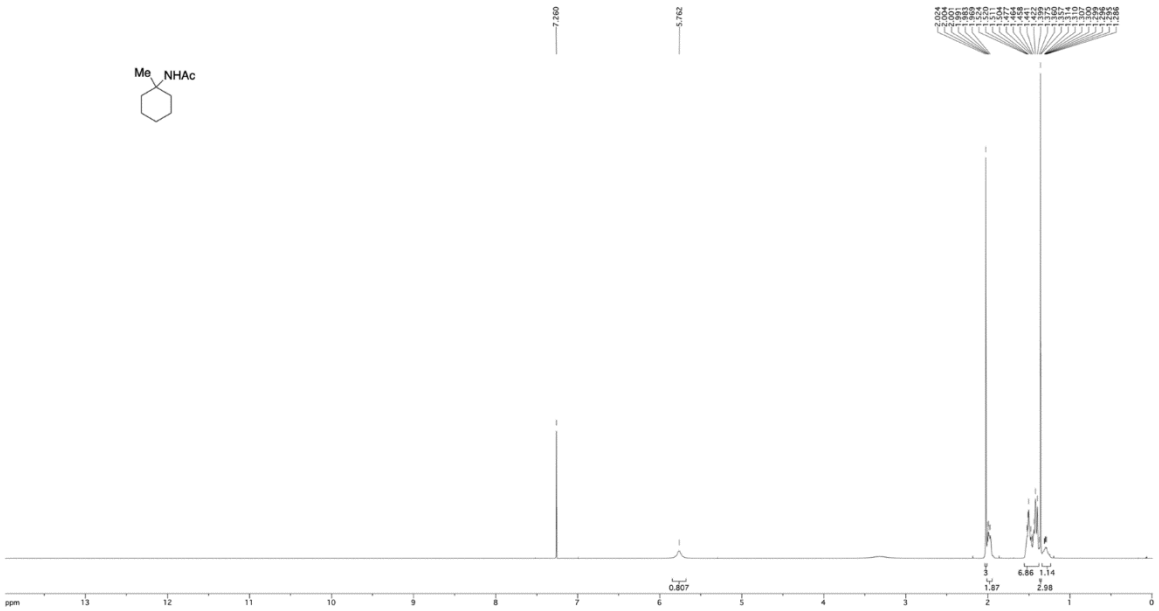
^{1}H-NMR spectrum of 4-*N*-acetylaminoheptane **21^a** obtained in $CDCl_3$ at 23 °C.



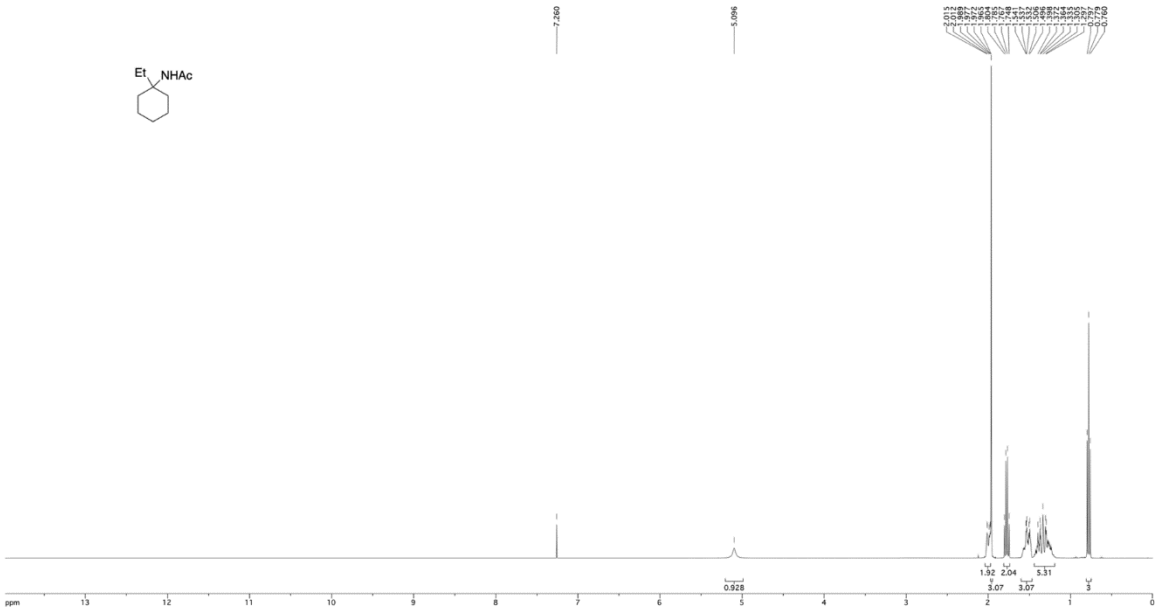
^{1}H-NMR spectrum of 2-*N*-acetylamino-3-methylpentane **22^{b}** obtained in $CDCl_3$ at 23 °C.



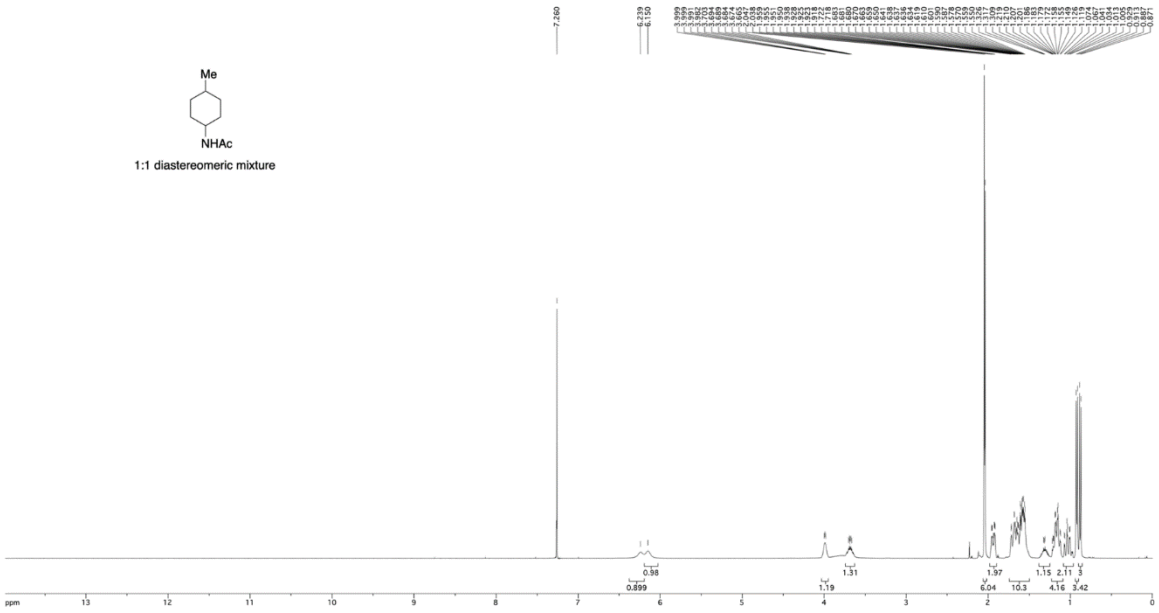
^{1}H-NMR spectrum of 3-*N*-acetylamino-3-methylpentane **22^{a}** obtained in $CDCl_3$ at 23 °C.



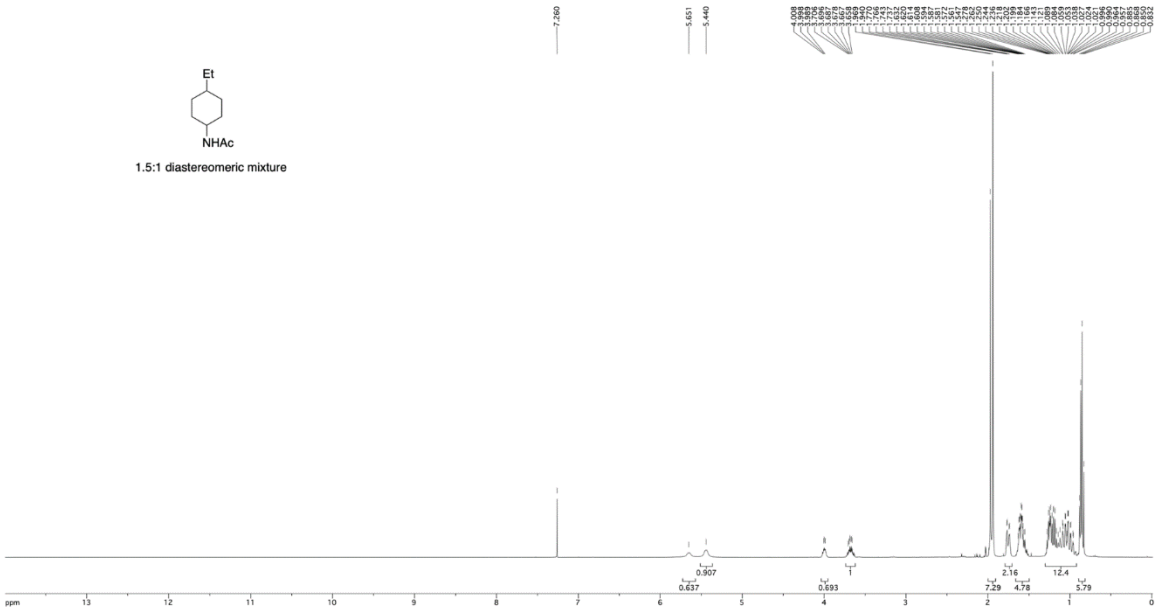
^{1}H-NMR spectrum of 1-*N*-acetylamino-1-methylcyclohexane **7**[1,1] obtained in $CDCl_3$ at 23 °C.



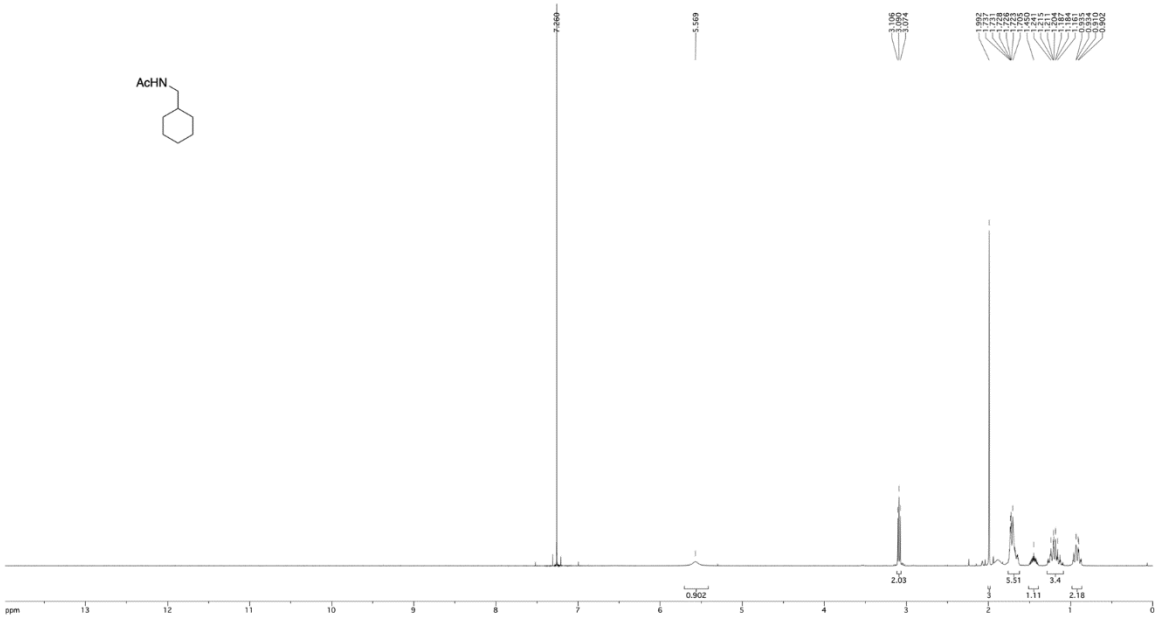
^{1}H-NMR spectrum of 1-*N*-acetylamino-1-ethylcyclohexane **10**[1,1] obtained in $CDCl_3$ at 23 °C.



^{1}H-NMR spectrum of the diastereomeric mixture of 1-*N*-acetylamino-4-methylcyclohexane **7**[1,4] obtained in $CDCl_3$ at 23 °C.



^{1}H-NMR spectrum of the diastereomeric mixture of 1-*N*-acetylamino-4-ethylcyclohexane **10**[1,4] obtained in $CDCl_3$ at 23 °C.



^{1}H-NMR spectrum of *N*-(cyclohexylmethyl)acetamide **7^{Me}** obtained in $CDCl_3$ at 23 °C.

B.11. X-ray Crystallography

Low-temperature diffraction data (ϕ-and ω-scans) were collected on a Bruker AXS D8 VENTURE KAPPA diffractometer coupled to a PHOTON II CPAD detector with Cu K_α radiation (λ = 1.54178 Å) from an IμS micro-source for the structure of compound V22057. The structure was solved by direct methods using SHELXS[17] and refined against F^2 on all data by full-matrix least squares with SHELXL-2017[18] using established refinement techniques.[19] All non-hydrogen atoms were refined anisotropically. 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).

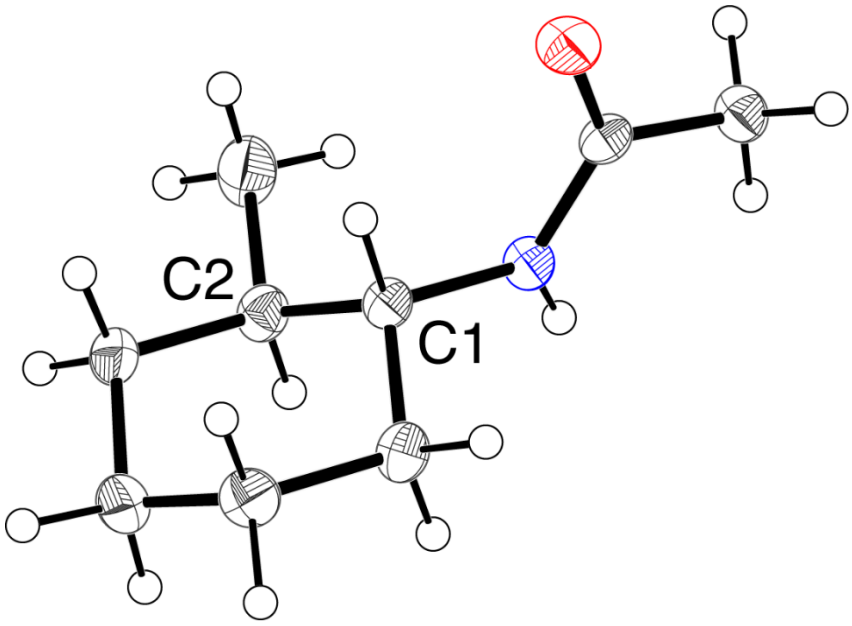


Figure B-3. Displacement ellipsoid plot for **ent-8** plotted at 50% probability. Compound **ent-8** crystallizes in the monoclinic space group $P2_1$ with two molecules in the asymmetric unit. The coordinates for the hydrogen atoms bound to N1 and N11 were located in the difference Fourier synthesis and refined semi-freely with the help of a restraint on the N-H distance (0.88(4) Å).

Table B-1. X-ray experimental details of **ent-8** (CCDC **2194386**).

Crystal data	
Chemical formula	$C_{18}H_{34}N_2O_2$
M_r	310.47
Crystal system, space group	Monoclinic, $P2_1$
Temperature (K)	100
a, *b*, *c* (Å)	4.7814 (5), 27.534 (2), 7.4464 (6)
b (°)	108.100 (5)
V (Å^3)	931.81 (15)
Z	2
Radiation type	Cu *K*a
m (mm^{-1})	0.56
Crystal size (mm)	0.40 × 0.10 × 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}	0.551, 0.754
No. of measured, independent and observed [$I > 2s(I)$] reflections	23144, 3759, 3498
R_{int}	0.069
$(\sin q/l)_{max}$ (Å^{-1})	0.625
Refinement	
$R[F^2 > 2s(F^2)]$, $wR(F^2)$, S	0.045, 0.119, 1.04
No. of reflections	3759
No. of parameters	209
No. of restraints	3
H-atom treatment	H atoms treated by a mixture of independent and constrained refinement
Γ_{max}, Γ_{min} (e Å^{-3})	0.21, -0.19
Absolute structure	Flack x determined using 1575 quotients [(I+)-(I-)]/[(I+)+(I-)] (Parsons, Flack and Wagner, Acta Cryst. B69 (2013) 249-259).
Absolute structure parameter	-0.3 (2)

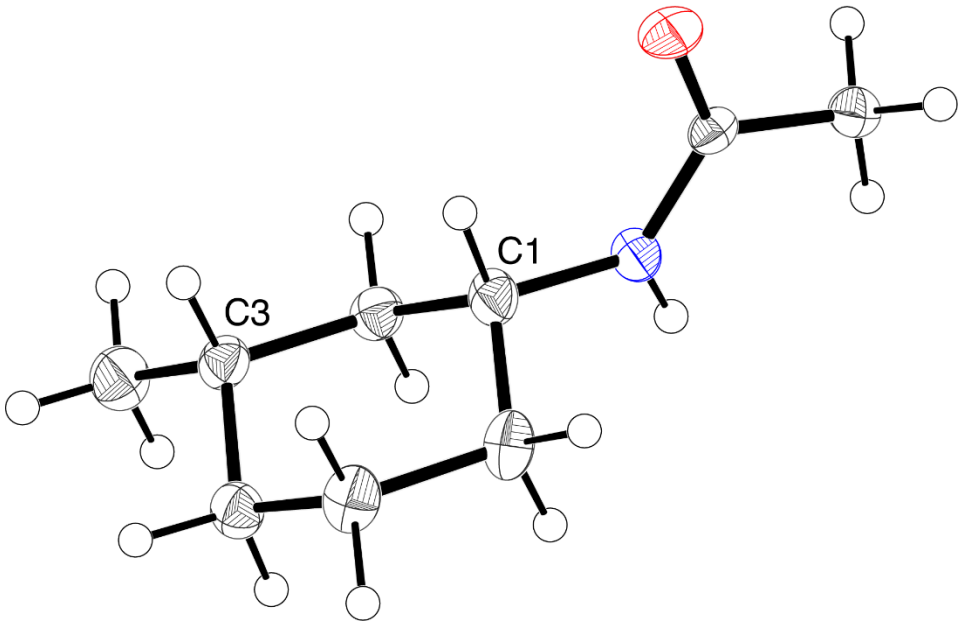


Figure B-4. Displacement ellipsoid plot for **6-N-Ac** plotted at 50% probability. Compound **6-N-Ac** crystallizes in the orthorhombic space group $P2_12_12_1$ with one molecule in the asymmetric unit. The coordinates for the hydrogen atom bound to N1 were located in the difference Fourier synthesis and refined semi-freely with the help of a restraint on the N-H distance (0.88(4) Å).

Table B-2. X-ray experimental details of **6-N-Ac** (CCDC **2194387**).

Crystal data	
Chemical formula	$C_9H_{17}NO$
M_r	155.23
Crystal system, space group	Orthorhombic, $P2_12_12_1$
Temperature (K)	100
a, b, c (Å)	4.8599 (6), 9.9649 (12), 19.3200 (18)
V (Å^3)	935.64 (18)
Z	4
Radiation type	Cu Ka
m (mm^{-1})	0.56
Crystal size (mm)	0.30 × 0.15 × 0.15
Data collection	
Diffractometer	Bruker D8 VENTURE Kappa Duo PHOTON II CPAD
Absorption correction	Multi-scan *SADABS2016*/2 (Sheldrick, 2014)
T_{min}, T_{max}	0.635, 0.754
No. of measured, independent and observed [I > 2s(I)] reflections	9348, 1896, 1877
R_{int}	0.033
(sin q/l)$_{max}$ (Å^{-1})	0.625
Refinement	
$R[F^2 > 2s(F^2)]$, $wR(F^2)$, S	0.035, 0.094, 1.10
No. of reflections	1896
No. of parameters	105
No. of restraints	1
H-atom treatment	H atoms treated by a mixture of independent and constrained refinement
Γ_{max}, Γ_{min} (e Å^{-3})	0.26, -0.18
Absolute structure	Flack x determined using 736 quotients [(I+)-(I-)]/[(I+)+(I-)] (Parsons, Flack and Wagner, Acta Cryst. B69 (2013) 249-259).
Absolute structure parameter	0.01 (9)

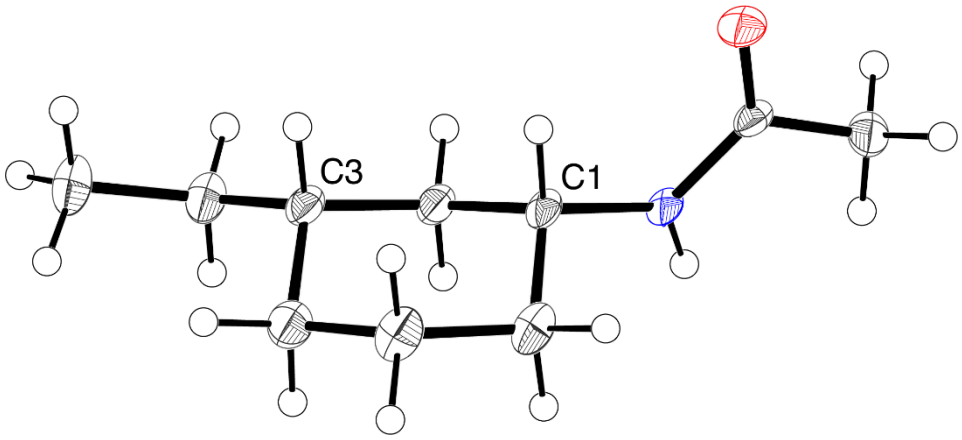


Figure B-5. Displacement ellipsoid plot for **9-N-Ac** plotted at 50% probability. Compound **9-N-Ac** crystallizes in the monoclinic space group $P2_1$ with one molecule in the asymmetric unit. The coordinates for the hydrogen atom bound to N1 were located in the difference Fourier synthesis and refined semi-freely with the help of a restraint on the N-H distance (0.88(4) Å).

Table B-3. X-ray experimental details of **9-N-Ac** (CCDC **2194384**).

Crystal data	
Chemical formula	$C_{10}H_{19}NO$
M_r	169.26
Crystal system, space group	Monoclinic, $P2_1$
Temperature (K)	100
a, *b*, *c* (Å)	10.2124 (8), 4.7908 (3), 10.8829 (10)
b (°)	107.325 (8)
V (Å^3)	508.29 (7)
Z	2
Radiation type	Cu *K*a
m (mm^{-1})	0.55
Crystal size (mm)	0.40 × 0.20 × 0.15
Data collection	
Diffractometer	Bruker D8 VENTURE Kappa Duo PHOTON II CPAD
Absorption correction	Multi-scan *SADABS2016*/2 (Sheldrick, 2014)
T_{min}, T_{max}	0.371, 0.754
No. of measured, independent and observed [$I > 2s(I)$] reflections	8913, 2022, 1966
R_{int}	0.055
$(\sin q/l)_{max}$ (Å^{-1})	0.625
Refinement	
$R[F^2 > 2s(F^2)]$, $wR(F^2)$, S	0.066, 0.185, 1.09
No. of reflections	2022
No. of parameters	114
No. of restraints	2
H-atom treatment	H atoms treated by a mixture of independent and constrained refinement
Γ_{max}, Γ_{min} (e Å^{-3})	0.30, -0.37
Absolute structure	Classical Flack method preferred over Parsons because s.u. lower.
Absolute structure parameter	-0.3 (5)

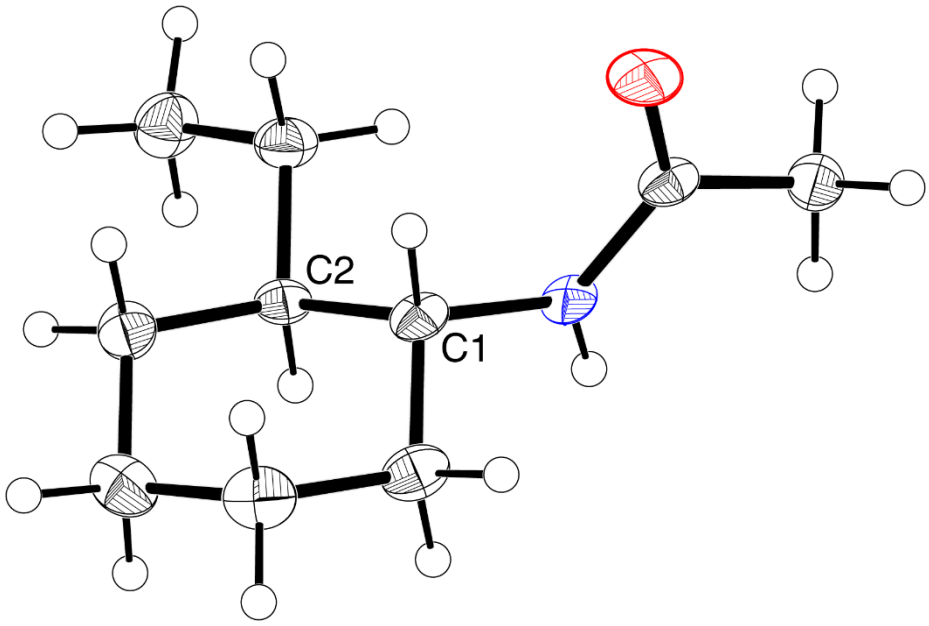


Figure B-6. Displacement ellipsoid plot for **ent-10** plotted at 50% probability. Compound **ent-10** crystallizes in the orthorhombic space group $P2_12_12_1$ with one molecule in the asymmetric unit. The coordinates for the hydrogen atom bound to N1 were located in the difference Fourier synthesis and refined semi-freely with the help of a restraint on the N-H distance (0.88(4) Å).

Table B-4. X-ray experimental details of **9-N-Ac** (CCDC **2194385**).

Crystal data	
Chemical formula	$C_{10}H_{19}NO$
M_r	169.26
Crystal system, space group	Orthorhombic, $P2_12_12_1$
Temperature (K)	100
a, *b*, *c* (Å)	4.8591 (5), 13.7522 (16), 15.4412 (16)
V (Å^3)	1031.83 (19)
Z	4
Radiation type	Cu *K*a
m (mm^{-1})	0.54
Crystal size (mm)	0.40 × 0.05 × 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}	0.460, 0.754
No. of measured, independent and observed [$I > 2s(I)$] reflections	14915, 2100, 1967
R_{int}	0.081
$(\sin q/l)_{max}$ (Å^{-1})	0.625
Refinement	
$R[F^2 > 2s(F^2)]$, $wR(F^2)$, S	0.045, 0.121, 1.06
No. of reflections	2100
No. of parameters	114
No. of restraints	1
H-atom treatment	H atoms treated by a mixture of independent and constrained refinement
Γ_{max}, Γ_{min} (e Å^{-3})	0.19, -0.23
Absolute structure	Flack x determined using 773 quotients [(I+)-(I-)]/[(I+)+(I-)] (Parsons, Flack and Wagner, Acta Cryst. B69 (2013) 249-259).
Absolute structure parameter	-0.6 (3)

B.12. Computational Studies

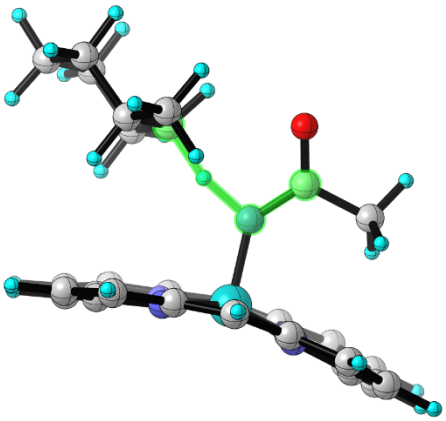
Density functional theory (DFT) was used to study the mechanism of iron porphyrin catalyzed nitrene insertion. UB3LYP[20] functional as implemented in Gaussian16 with 6-31G(d)[21] basis set for C, H, N, O and LANL2DZ[22] for Fe was used for geometry optimizations, transition state (TS) optimizations, quick reaction coordinates and computing vibrational frequencies. This level of theory was employed in other studies[1,23,24] and was shown to be appropriate via benchmarks against KS-UCCSD(T)/aug-cc-pVTZ. The key word guess=mix was used for all the open shell singlet computations. Chlorobenzene was used as the solvent to mimic the enzymatic environment and was approximated via polarized continuum model. The stability of the wave functions were verified using the key word stable=opt. Thermal corrections were computed using Grimme's quasi-rigid rotor harmonic oscillator approximation.[25] Single point energies were computed at PCM(Chlorobenzene)-UB3LYP-D3BJ/6-311+G(d,p)/LANL2TZ(f) (Fe)[20-22]and PCM(Chlorobenzene)-UWB97XD/6-311+G(d,p)/LANL2TZ(f) (Fe)[21,22,26] on the amination and amidation systems respectively and were shown to produce accurate results for these systems by Head-Gordon[27] and Liu and Arnold.[24] ISOEFF code developed by Paneth and coworkers[28-30] was used to compute the KIEs. Molecular dynamics simulations were performed using GROMACS. PDB file 5UCW was extracted from the database and molecular dynamics simulations (MD) were carried out on this after incorporating the appropriate mutations using CHARMM-GUI. An 8.5nm cubic water box was used for solvation. TIP3P water was used for simulating the waters. 12 sodium atoms were used for neutralizing the charges. CHARMM36m force field parameters were used

for the protein and HEME, while SWISSParm was used to generate parameters for the substrate. Nitrene parameters were generated using Gaussian 16. Two stage energy minimization was implemented. In the first step, position restraints with a harmonic potential of 500 kcal/mol/Å2 was applied on the protein backbone, HEME, Nitrene, and the substrate heavy atoms. Solvent and sodium ions were allowed to equilibrate without any constraints. In the second step, the whole system is minimized without any restraints. In this step, the pull code for the distances between significant atoms involved in the reaction is implemented and carried forward to equilibration and production runs. Steepest descent algorithm was used for all energy minimizations and Verlet cut-off scheme with particle mesh ewald summation and 8Å cut-off for electrostatic and van der waal was employed. A force switch VDW modifier was used to smoothly truncate the VDW at eh cut-off after 6Å. Hydrogen bonds were constrained by the LINCS algorithm, and the SHAKE algorithm was used to constraint the water molecule geometry. Similar constraints, cut-offs, and cut-off schemes were used for equilibration and production runs. Annealing was implemented with a 2 ns equilibration run with 1fs time step at 300K using the V-rescale (modified Berendsen) thermostat under constant temperature (NVT) conditions to take the system from 0 to 300K. The production run of 100 ns with 2 fs time step at 300K and 1 bar was carried out using the Nosé-Hoover thermostat and Parinello-Rahman barostat under constant pressure (NPT) conditions. In the production run, a harmonic potential constraint of 30 kcal/mol/Å2 was implemented on substrate-nitrene (2.7 Å), nitrene-iron (1.95 Å). The last 20 ns were then used to obtain an average geometry for the protein, HEME, and substrate. A 10 Å spherical pocket centered around the substrate

was then extracted from these geometries to obtain the model theozyme. This theozyme was protonated appropriately and was used to perform constrained ONIOM optimizations at B3LYP-D3BJ/6-31G*/LANL2DZ:UFF level of theory. The distance between C-H and H-N were constrained based on the model QM computations. Frequency computations were performed to verify that only one imaginary mode corresponding to HAT was observed in these constrained optimizations. These geometries were then used for generating the qualitative steric map and dispersion interaction data. Steric maps were generated using the SEQCROW package of CHIMERAX.[31] Molecular graphics were generated using CYLView. Fragmentation analysis was performed at ωB97XD/6-311+G(d)/LANL2TZ(f) level of theory by substituting the C-H…O interaction of note with CH…H interaction.

B.12.1. Conformational Sampling:

Conformational sampling was performed by rotating about the highlighted bonds in scheme B-1 and through methylcyclohexane ring flips combinatorially. For the amidation reaction, a total of 8 unique combinations were identified while for the amination reaction 4 unique combinations were obtained. These combinations were then subjected to further permutations based on the regioisomers, the relative positions of the methyl group and the incoming nitrenoid groups, and the various electron surfaces leading to over 500 TSs.



Scheme B-1. Algorithm for the conformational search.

B.12.2. Catalytic Cycle for Amidation:

The general catalytic cycle for the iron porphyrin catalyzed amidation is presumed to be similar to uAMD8 catalyzed nitrene insertion owing to the similarities in the structure. A general catalytic cycle for this process is shown in figure B-7.

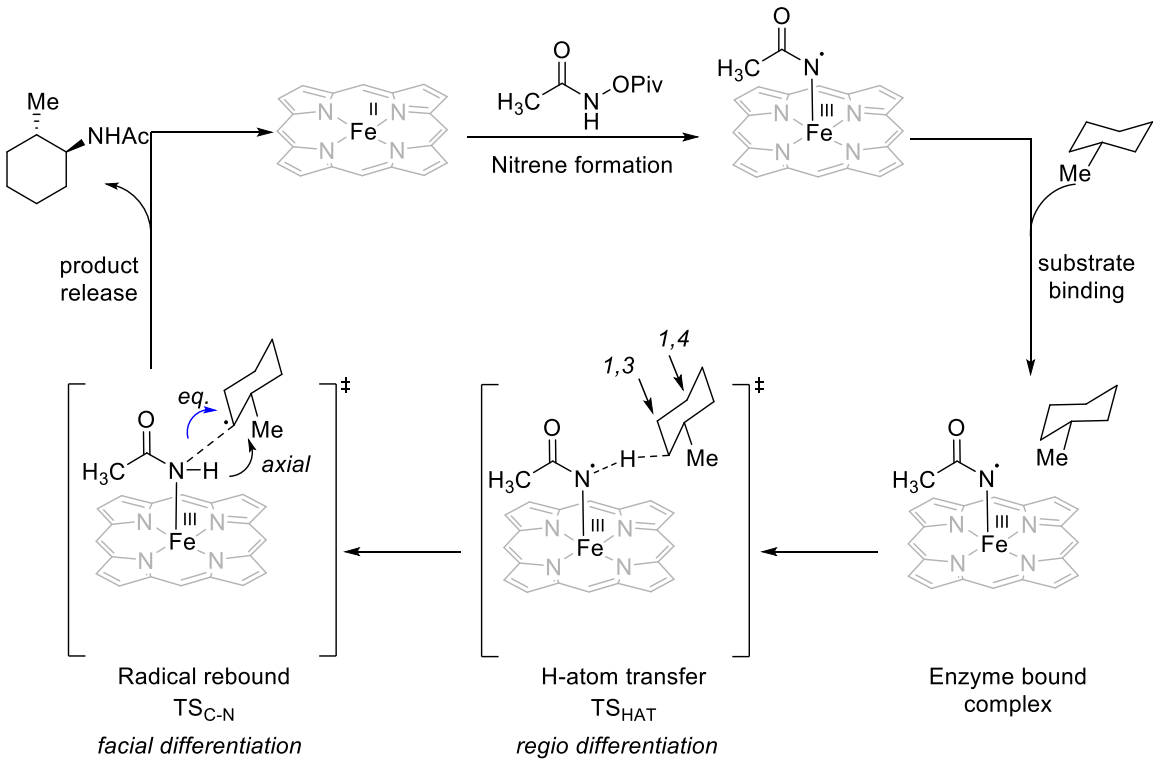


Figure B-7. Proposed catalytic cycle for iron porphyrin catalyzed amidation.

Singlet, triplet, and quintet electronic surfaces were explored for both the HAT step and the radical rebound step. The transition states for the singlet surface were found to stable on the open shell singlet surface (Figure B-7). The barrier for HAT TS on the singlet surface is comparable to that on the triplet surface. However, radical rebound TS on the singlet surface is disfavored by 3.2 kcal/mol over the lowest-lying radical rebound transition state on the triplet surface. The HAT barriers for different regioisomers were

computed at PCM(Chlorobenzene)-UWB97XD/6-311+G(d,p)/LANL2TZ(f)(Fe)//PCM(chlorobenzene)-B3LYP/6-31(d)/LANL2DZ(Fe). HAT TSs were located on the triplet surface for all the regioisomers. The barriers are all within 2.0 kcal/mol of each other for 1,2- , 1,3-, and 1,4- substituted products as shown in Figure B-7. However, the lowest energy HAT TS leading to the1,1-product is disfavored by 6.2 kcal/mol compared to the lowest energy HAT TS regioisomer which leads to the1,2-product. Furthermore, the axial and equatorial hydrogen atom abstractions are within 1 kcal/mol of each other for all the regioisomeric TSs.

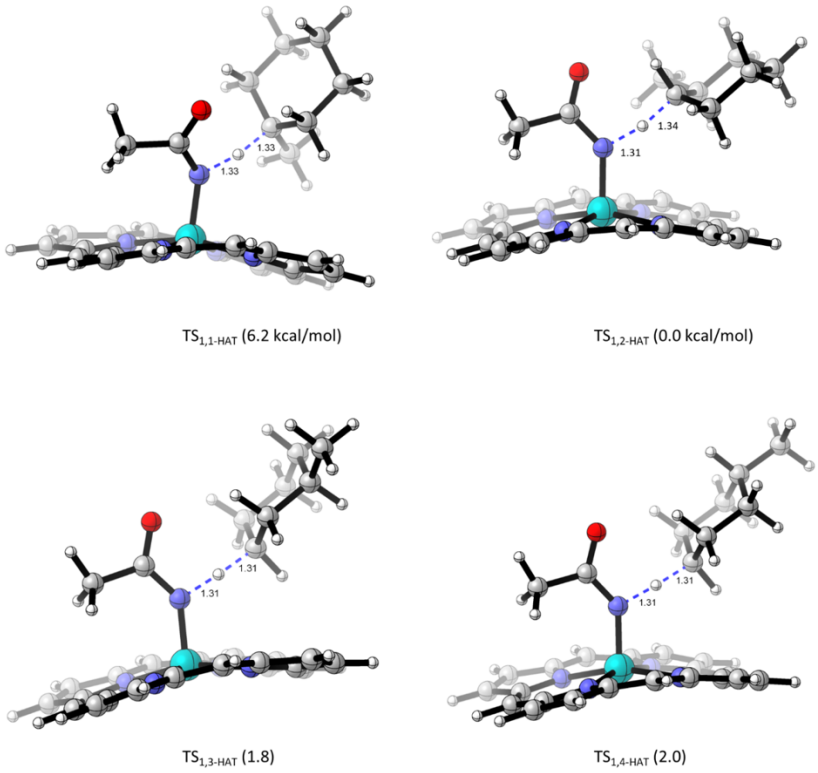


Figure B-8. Free energy barriers for the lowest energy HAT TSs leading to each of the possible regioisomeric products in the iron porphyrin catalyzed nitrene insertion located on the triplet surface (energies in kcal/mol).

The second stereocenter is determined in the radical rebound step, thereby making it the likely diastereodeterming TS based upon the relative heights of the barriers. The diastereoselectivity originates from either the axial or the equatorial attack of the nitrenoid intermediate or a possible ring flip. However, the radical rebound regioisomeric TS was found to favor the *equatorial* attack of the nitrenoid intermediate by over 3 kcal/mol (Figure B-8). Furthermore, these regioisomeric TSs were all within 4 kcal/mol of each other. These TSs were located on the triplet and quintet surfaces and the energy difference between the quintet and triplet surfaces is approximately 0.5 kcal/mol for most of the regioisomeric TSs – similar to our observations in the nitrene insertion of benzylic substrates.[32]

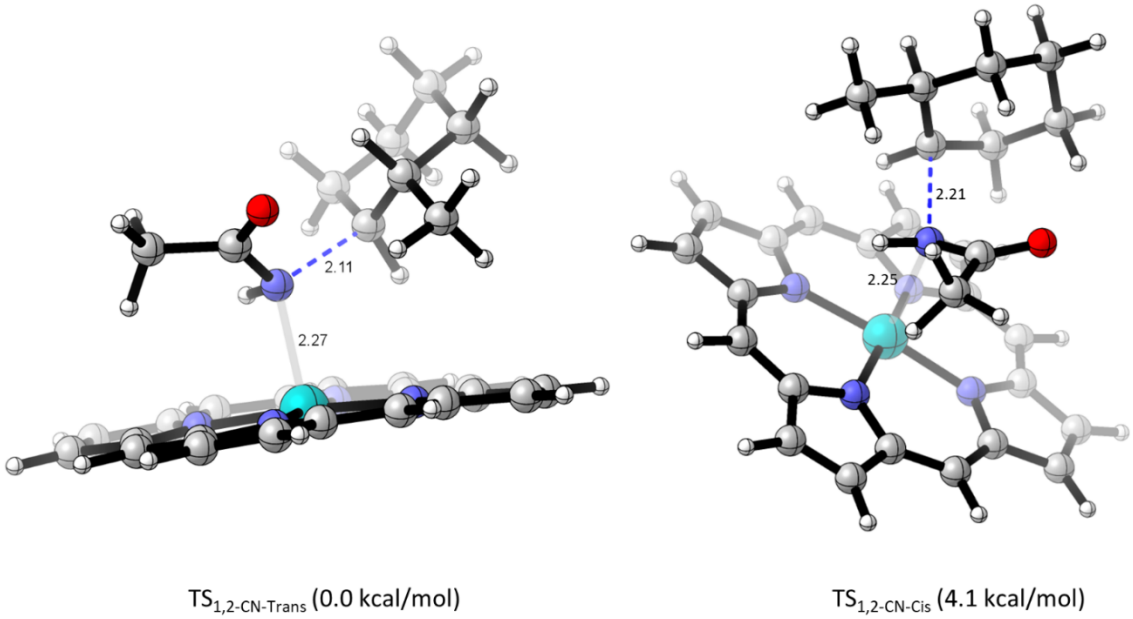


Figure B-9. Key radical rebound TSs for the iron porphyrin catalyzed nitrene insertion on the triplet surface for the 1,2-trans diastereomer resulting from equatorial attack and the 1,2-cis diastereomer resulting from axial attack. Computed at PCM(Chlorobenzene)-UWB97XD/6-311+G(d,p)/LANL2TZ(f)(Fe)//PCM(chlorobenzene)-B3LYP/6-31(d)/LANL2DZ(Fe) (energies in kcal/mol).

Furthermore, we carried out transition state optimization of the theozyme obtained from a 10 Å pocket of the geometries from molecular dynamics simulations of the docked-TS structures by freezing the bond forming and bond breaking distances. These geometries were then used to obtain qualitative insights into the active site of uAMD8 enzyme. Moreover, for the major regioisomeric TS ($TS_{1,2\text{-HAT}}$), a full enzyme optimization was performed with the key bond forming and bond breaking distances frozen. The steric maps obtained from both the ONIOM of the full enzyme and the theozyme were found to be qualitatively identical for the major regioisomeric TS (Figure B-9). Therefore, analysis was performed on the theozymes in the manuscript for the minor products.

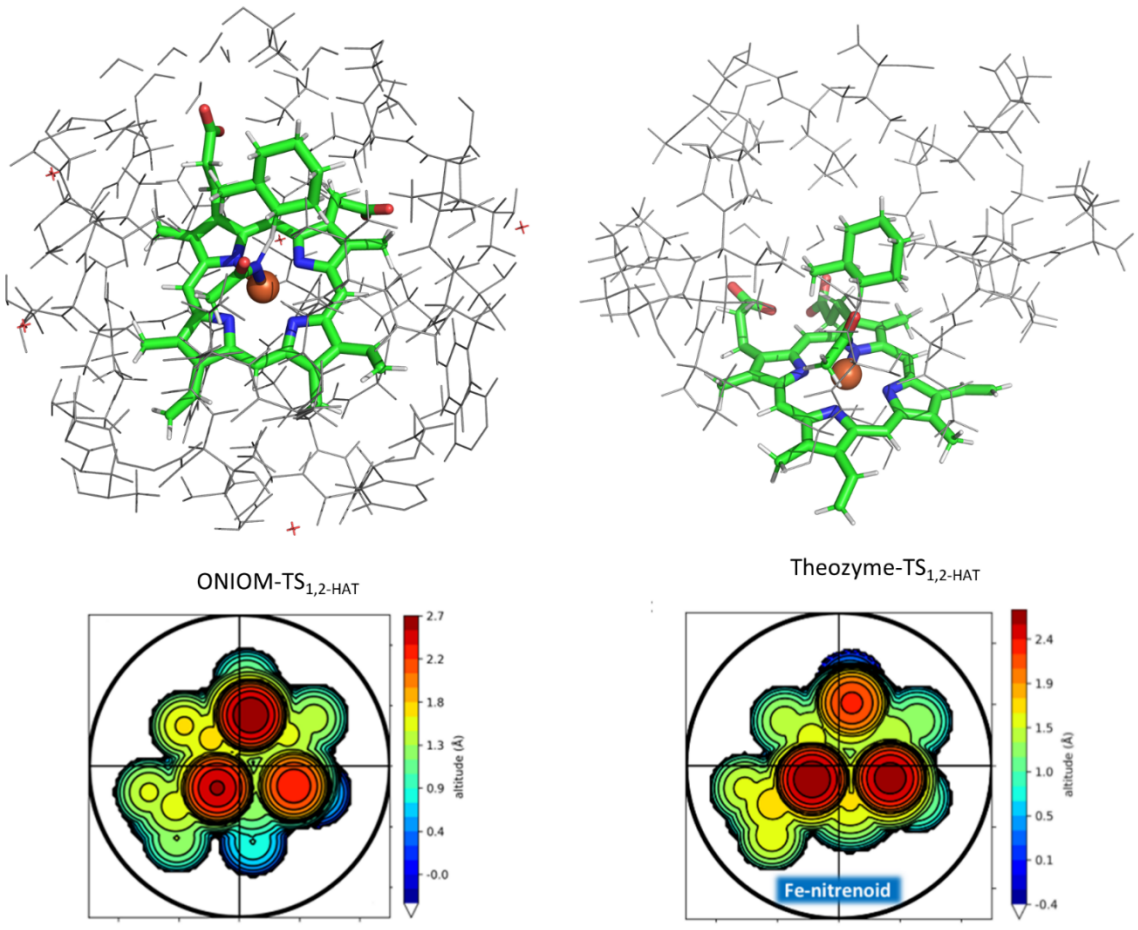


Figure B-10. Steric analysis and the constrained TS for the full enzyme (only the 12A pocket is shown in the figure for clarity) and the 10Å theozyme.

CHIMERAX was used to obtain the volume of the pocket of this theozyme and the 10 Å pocket extracted from the QMMM computation of the full enzyme. The volume data obtained from both computations were within error. The volume of methylcyclohexane and nitrene were obtained from the cavity volumes used in the PCM cavity. This was then used to estimate the free volume of the pocket (Table B-5).

Entry	Volume ($Å^3$)
Cavity	418.9
Nitrene	187.5
Substrate	110

Table B-5. Table with the volume data for the Theozyme-$TS_{1,2\text{-HAT}}$.

B.12.3. Dispersion Interactions:

Starting from the ONIOM constrained optimization geometries leading to the 1,2-, 1,3-, and 1,4-amidation products, we utilized Grimme and coworkers'[33] DFT-D3 code with Becke-Johnson[34] correction to compute the favorable dispersion interactions[35] experienced by the substrate (methylcyclohexane) by the nearby residues in the absence of water. All residues within 6A of the substrate, including the iron nitrenoid species, were included in the dispersion interaction computation. The substrate was then removed from this pocket and dispersion computations were recomputed. The difference between these two values was used to evaluate the favorable dispersion interactions experienced by the substrate (Table B-6).

Entry	D4-Value
Theozyme-$TS_{1,2\text{-HAT}}$	10.1
Theozyme-$TS_{1,3\text{-HAT}}$	9.0
Theozyme-$TS_{1,4\text{-HAT}}$	6.8

Table B-6. Table with the dispersion data for the Theozyme-$TS_{1,2\text{-HAT}}$, Theozyme-$TS_{1,3\text{-HAT}}$, and Theozyme-$TS_{1,4\text{-HAT}}$. Energies in kcal/mol.

Finally, the stabilizing energy of the C-H…O interactions in each of the theozyme regioisomeric TSs was estimated using fragmentation analysis (by replacing the oxygen of the nitrenoid with a hydrogen and evaluating the energy). This approach resulted yielded a stabilization energy of 0.6 kcal/mol for Theozyme-$TS_{1,2\text{-HAT}}$, 0.2 kcal/mol for Theozyme-$TS_{1,3\text{-HAT}}$, and 0.3 kcal/mol for Theozyme-$TS_{1,4\text{-HAT}}$.

B.12.4. Catalytic Cycle for Amination:

The general catalytic cycle for the iron porphyrin catalyzed amidation is presumed to be similar to uAMD8 catalyzed nitrene insertion owing to the similarities in the structure. A general catalytic cycle for this process is shown in Figure B-11.

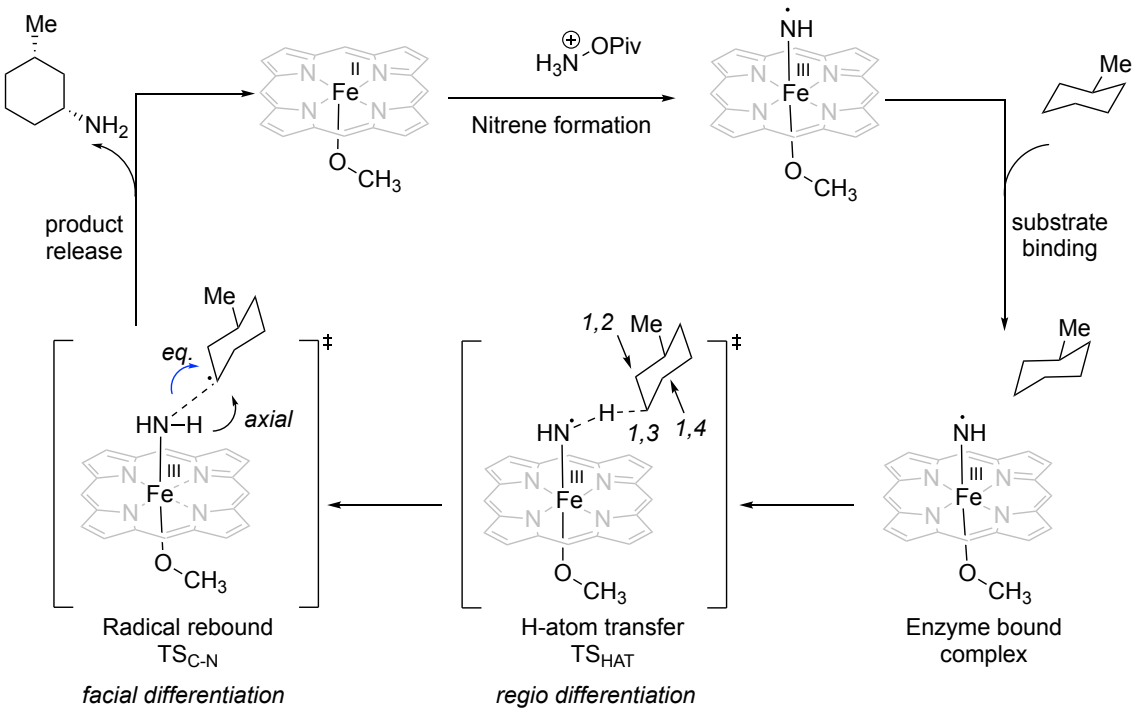


Figure B-11. Proposed catalytic cycle for iron porphyrin catalyzed amination.

Singlet, triplet, and quintet electronic surfaces were explored for both the HAT step and the radical rebound step. The transition states for the singlet surface were found to be stable on the open shell singlet surface. These TS were computed at PCM(Chlorobenzene)-UB3LYP-D3BJ/6-311+G(d,p)/LANL2TZ(f)(Fe)//PCM(chlorobenzene)-UB3LYP/6-31(d)/LANL2DZ(Fe). Similar to the amidation reaction, the barriers for the amination reaction with the model porphyrin ligand are all within 1.0 kcal/mol of each other for 1,1-

, 1,2- , 1,3-, and 1,4- substituted products as shown in Figure B-12. Furthermore, minimal energetic preference was observed for the axial or the equatorial hydrogen atom abstraction by the nitrenoid. Unsurprisingly, for the amination reaction with the model porphyrin ligand, the TS leading to 1,1-product was favored over the TS leading to other regioisomers owing to the more stable tertiary radical and the lack of sterically demanding acetyl group of the amidation reaction. Moreover, the TSs for 1,1-, 1,2-, 1,3-, and 1,4-products all had similar C…H and N…H bond-breaking and bond-forming distances (see Figure B-12).

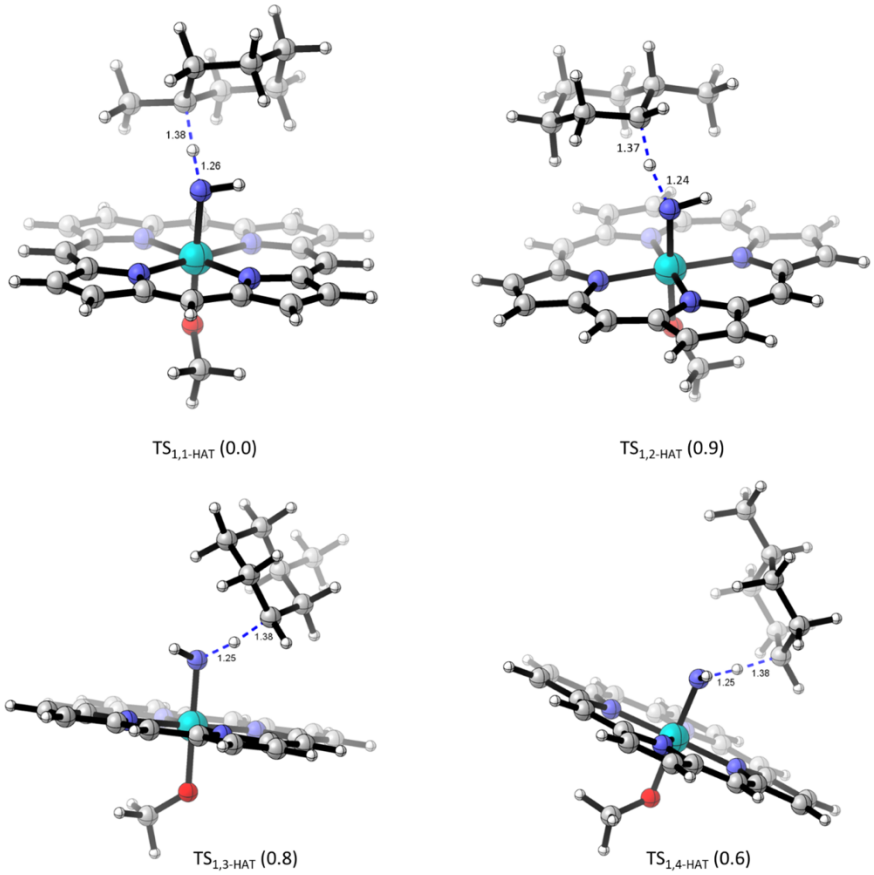


Figure B-12. TSs for the HAT step for the amination reaction of methylcyclohexane leading to different regioisomers of product, computed at B3LYP-D3BJ/6-311+G(d,p)/LANL2TZ(f) (Fe)//B3LYP/6-31G(d)/LANL2DZ (Fe) level of theory. Energies are in kcal/mol.

Finally, the diastereo-determining radical rebound TSs were found to be on the triplet surface similar to the reported literature.[36] Similar to the amidation of methylyclohexane on the pophyrin model system, DFT computations exhibit a strong preference for the equatorial approach of the nitrenoid (~2.6 kcal/mol) over the axial approach, as shown in Figure B-13.

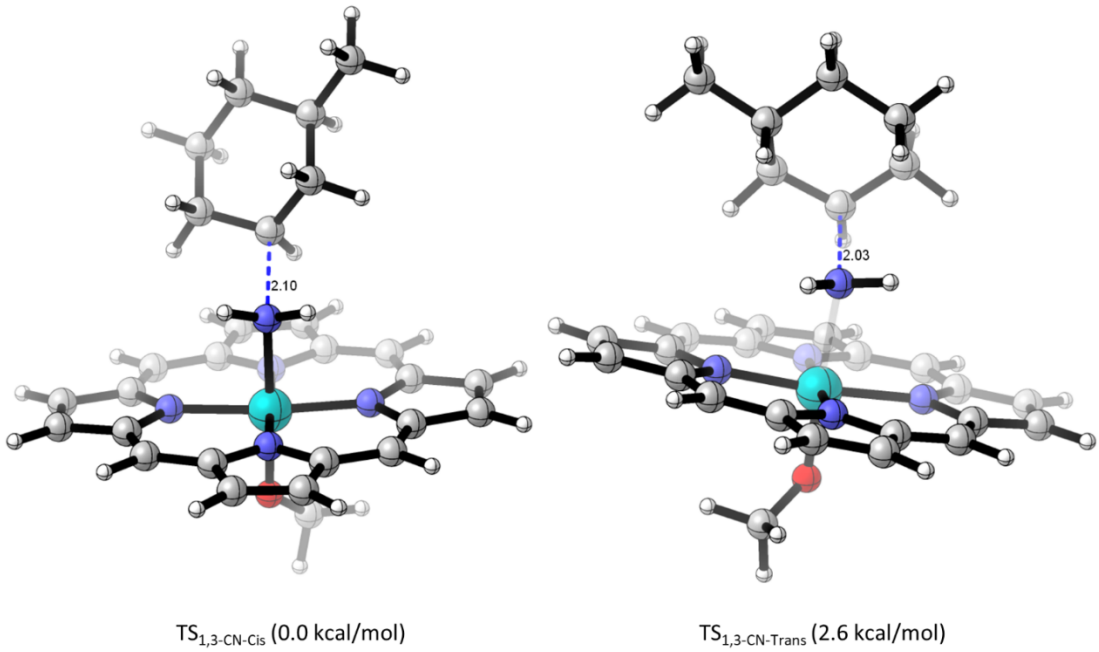
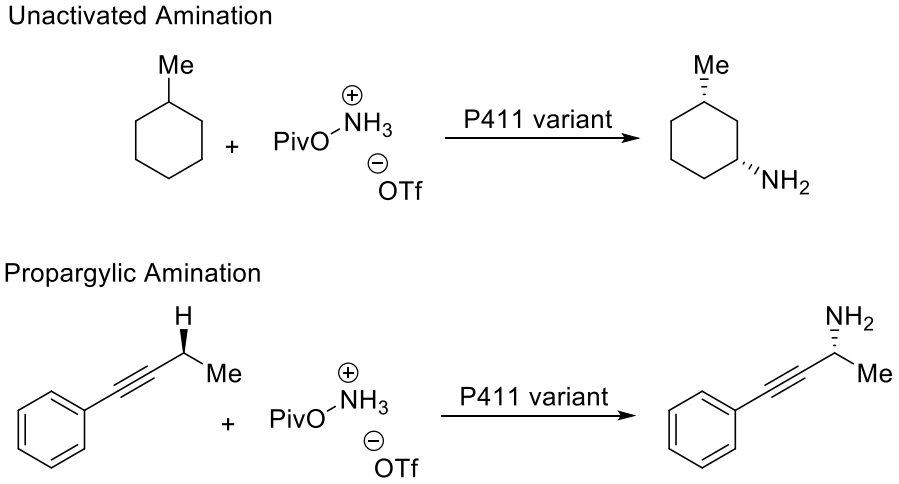


Figure B-13. TSs for diastereo-determining radical rebound step for the amination reaction leading to the 1,3-*cis* product resulting from equatorial attack and the 1,3-*trans* product resulting from axial attack, computed at B3LYP-D3BJ/6-311+G(d,p)/LANL2TZ(f) (Fe)//B3LYP/6-31G(d)/LANL2DZ (Fe) level of theory. Energies are in kcal/mol.

Comparing the unactivated uPA catalyzed amination of methylcyclohexane and the propargylic amination free energy barriers reported by Garcia-Borràs, Houk, and Arnold[36] computed at B3LYP-D3BJ/6-311+G(d,p)/LANL2TZ(f)//B3LYP /6-31G(d)/LANL2DZ (Fe) level of theory on model porphyrin systems further demonstrates the importance of

the enzyme active site in lowering the barriers for HAT (Figure B-14). Moreover, unlike the potential energy surface for the amidation reaction, for amination, we did not observe any surface crossing. The reaction for both propargylic amination and unactivated amination both proceeded on the triplet surface.



Scheme B-2. Scheme showing the unactivated amination investigated in this study and the activated (propargylic) amination reported by Garica-Borràs, Houk, and Arnold.[36]

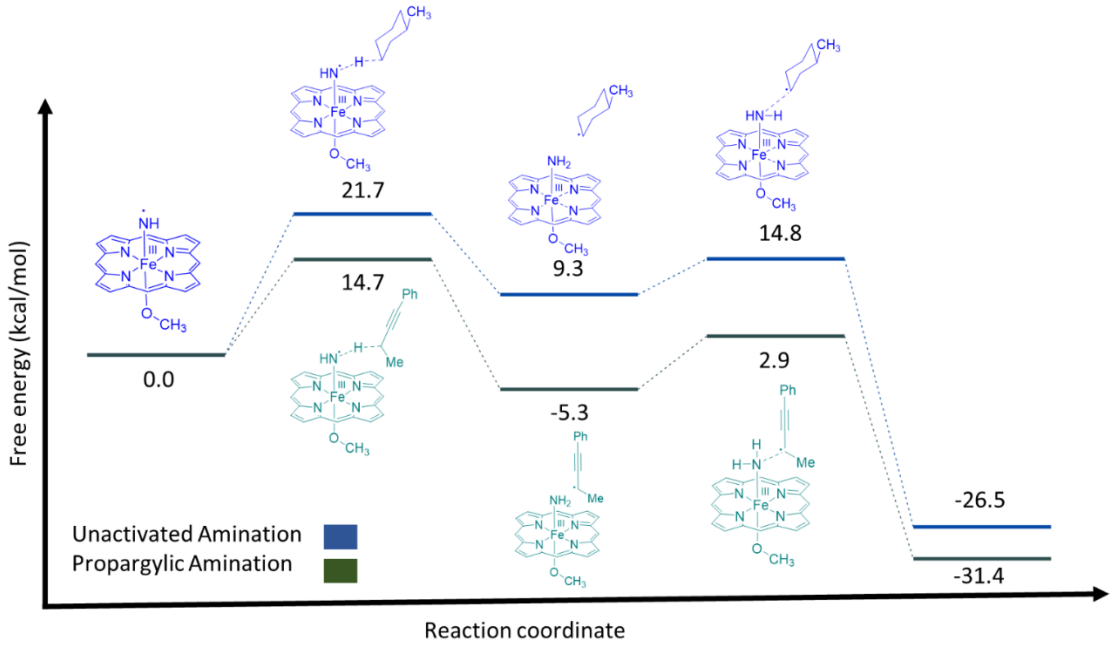


Figure B-14. Comparison of the potential energy surface of the model porphyrin catalyzed unactivated and propargylic amination reactions. Energies are in kcal/mol.

B.13. References

1. Athavale, S. V.; Gao, S.; Liu, Z.; Mallojjala, S. C.; Hirschi, J. S.; Arnold, F. H. Biocatalytic, Intermolecular C−H Bond Functionalization for the Synthesis of Enantioenriched Amides. *Angew. Chem. Int. Ed.* **2021**, *60* (47), 24864–24869.
2. (a) Makai, S.; Falk, Eric.; Morandi, B. Preparation of O-Pivaloyl hydroxylamine triflic acid. *Org. Synth.* **2020**, 97, 207-216. (b) Jia, Z.-J.; Gao, S.; Arnold, F. H. Enzymatic Primary Amination of Benzylic and Allylic C(Sp^3)–H Bonds. *J. Am. Chem. Soc.* **2020**, *142* (23), 10279–10283.
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* (2), 83–92.
4. 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* (5), 343–345.
5. Boville, C. E.; Scheele, R. A.; Koch, P.; Brinkmann-Chen, S.; Buller, A. R.; Arnold, F. H. Engineered Biosynthesis of β-Alkyl Tryptophan Analogues. *Angew. Chem. Int. Ed.* **2018**, *57* (45), 14764–14768.
6. Zhao, H.; Giver, L.; Shao, Z.; Affholter, J. A.; Arnold, F. H. Molecular Evolution by Staggered Extension Process (StEP) in Vitro Recombination. *Nat Biotechnol* **1998**, *16* (3), 258–261.
7. Barr, I.; Guo, F. Pyridine Hemochromagen Assay for Determining the Concentration of Heme in Purified Protein Solutions. *BIO-PROTOCOL* **2015**, *5* (18). https://doi.org/10.21769/BioProtoc.1594.
8. Xu, F.; Wu, Q.; Chen, X.; Lin, X.; Wu, Q. A Single Lipase-Catalysed One-Pot Protocol Combining Aminolysis Resolution and Aza-Michael Addition: An Easy and Efficient Way to Synthesise β-Amino Acid Esters: Synthesis of β-Amino Acid Esters. *Eur. J. Org. Chem.* **2015**, *2015* (24), 5393–5401.
9. Zhou, J.; List, B. Organocatalytic Asymmetric Reaction Cascade to Substituted Cyclohexylamines. *J. Am. Chem. Soc.* **2007**, *129* (24), 7498–7499.
10. Alalla, A.; Merabet-Khelassi, M.; Aribi-Zouioueche, L.; Riant, O. Green Synthesis of Benzamides in Solvent- and Activation-Free Conditions. *Synthetic Communications* **2014**, *44* (16), 2364–2376.
11. Roice, M.; Christensen, S. F.; Meldal, M. ULTRAMINE: A High-Capacity Polyethylene-Imine-Based Polymer and Its Application as a Scavenger Resin. *Chem. Eur. J.* **2004**, *10* (18), 4407–4415.
12. Zhang, Z.; Zheng, M.; Xue, X.; Marek, I.; Zhang, F.; Ma, J. Catalytic Enantioselective Cyclopropenation of Internal Alkynes: Access to Difluoromethylated Three-Membered Carbocycles. *Angew. Chem. Int. Ed.* **2019**, *58* (50), 18191–18196.
13. Kiyokawa, K.; Watanabe, T.; Fra, L.; Kojima, T.; Minakata, S. Hypervalent Iodine(III)-Mediated Decarboxylative Ritter-Type Amination Leading to the Production of α-Tertiary Amine Derivatives. *J. Org. Chem.* **2017**, *82* (22), 11711–11720.

14. Strick, B. F.; Mundal, D. A.; Thomson, R. J. An Oxidative [2,3]-Sigmatropic Rearrangement of Allylic Hydrazides. *J. Am. Chem. Soc.* **2011**, *133* (36), 14252–14255.
15. Mondal, A.; Subaramanian, M.; Nandakumar, A.; Balaraman, E. Manganese-Catalyzed Direct Conversion of Ester to Amide with Liberation of H_2. *Org. Lett.* **2018**, *20* (11), 3381–3384.
16. Hurtado-Rodrigo, C.; Hoehne, S.; Muñoz, M. P. A New Gold-Catalysed Azidation of Allenes. *Chem. Commun.* **2014**, *50* (12), 1494–1496.
17. Sheldrick, G. M. *Acta Cryst.* **1990**, A46, 467-473.
18. Sheldrick, G. M. *Acta Cryst.* **2015**, C71, 3-8.
19. Müller, P. *Crystallography Reviews* **2009**, *15*, 57-83.
20. Becke, A. D., Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993,** *98* (7), 5648-5652.
21. Hehre, W. J.; Stewart, R. F.; Pople, J. A., Self-Consistent Molecular-Orbital Methods. I. Use of Gaussian Expansions of Slater-Type Atomic Orbitals. *The Journal of Chemical Physics* **1969,** *51* (6), 2657-2664.
22. Hay, P. J.; Wadt, W. R., Ab initio effective core potentials for molecular calculations. Potentials for the transition metal atoms Sc to Hg. *J. Chem. Phys.* **1985,** *82* (1), 270-283.
23. Wang, J.; Gao, H.; Yang, L.; Gao, Y. Q., Role of Engineered Iron-haem Enzyme in Reactivity and Stereoselectivity of Intermolecular Benzylic C–H Bond Amination. *ACS Catal* **2020,** *10* (9), 5318-5327.
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(sp(3))-H bonds. *Nat Chem* **2019,** *11* (11), 987-993.
25. Grimme, S., Supramolecular binding thermodynamics by dispersion-corrected density functional theory. *Chem. Eur. J.* **2012,** *18* (32), 9955-64.
26. Chai, J. D.; Head-Gordon, M., Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008,** *10* (44), 6615-20.
27. Mardirossian, N.; Head-Gordon, M., Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals. *Molecular Physics* **2017,** *115* (19), 2315-2372.
28. Wigner, E., Crossing of potential thresholds in chemical reactions. *Z. Phys. Chem. B* **1932,** *19*, 203.
29. Bigeleisen, J.; Mayer, M. G., Calculation of Equilibrium Constants for Isotopic Exchange Reactions. *J. Chem. Phys.* **1947,** *15* (5), 261-267.
30. Anisimov, V.; Paneth, P., A Program for Studies of Isotope Effects Using Hessian Modifications. *J. Mathem. Chem.* **1999,** *26* (1/3), 75-86.
31. Ingman, V. M.; Schaefer, A. J.; Andreola, L. R.; Wheeler, S. E., QChASM Quantum chemistry automation and structure manipulation. *WIREs Computational Molecular Science* **2020,** *11* (4).

32. Athavale, S. V.; Gao, S.; Liu, Z.; Mallojjala, S. C.; Hirschi, J. S.; Arnold, F. H., Biocatalytic, Intermolecular C-H Bond Functionalization for the Synthesis of Enantioenriched Amides. *Angew Chem Int Ed Engl* **2021,** *60* (47), 24864-24869.
33. 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* (15), 154104.
34. Grimme, S.; Ehrlich, S.; Goerigk, L., Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011,** *32* (7), 1456-65.
35. Wagner, J. P.; Schreiner, P. R., London dispersion in molecular chemistry--reconsidering steric effects. *Angew Chem Int Ed Engl* **2015,** *54* (42), 12274-96.
36. Liu, Z.; Qin, Z. Y.; Zhu, L.; Athavale, S. V.; Sengupta, A.; Jia, Z. J.; Garcia-Borras, M.; Houk, K. N.; Arnold, F. H., An Enzymatic Platform for Primary Amination of 1-Aryl-2-alkyl Alkynes. *J Am Chem Soc* **2022,** *144* (1), 80-85.

Chapter IV

ENZYMATIC NITROGEN INCORPORATION USING HYDROXYLAMINE

Material from this chapter appears in: "**Gao, S.**; Das, A.; Alfonzo, E.; Sicinski, K. M.; Rieger, D.; Arnold, F. H. * Enzymatic Nitrogen Incorporation Using Hydroxylamine. Submitted."

S. G. discovered the initial activity. S. G. participated in the design and execution of the research. S. G. prepared the manuscript.

ABSTRACT

Hydroxylamine-derived reagents have enabled versatile nitrene transfer reactions for introducing nitrogen-containing functionalities in small-molecule catalysis as well as biocatalysis. However, the state-of-art *oxygen*-substituted hydroxylamines result in poor atom economy and stoichiometric organic waste. Activating hydroxylamine (NH_2OH) for nitrene transfer offers a low-cost and sustainable route to amine synthesis, since water is the sole byproduct. Despite its presence in nature, however, hydroxylamine is not known to be used for enzymatic nitrogen incorporation in biosynthesis. Here, we report an engineered heme enzyme that can utilize hydroxylammonium chloride, an inexpensive commodity chemical, for nitrene transfer. Directed evolution of *Pyrobaculum arsenaticum* protoglobin generated efficient enzymes for benzylic C–H primary amination and styrene aminohydroxylation. Mechanistic studies supported a stepwise radical pathway involving rate-limiting hydrogen atom transfer. This unprecedented activity is a useful addition to the 'nitrene transferase' repertoire and hints at possible future discovery of natural enzymes that use hydroxylamine for amination chemistry.

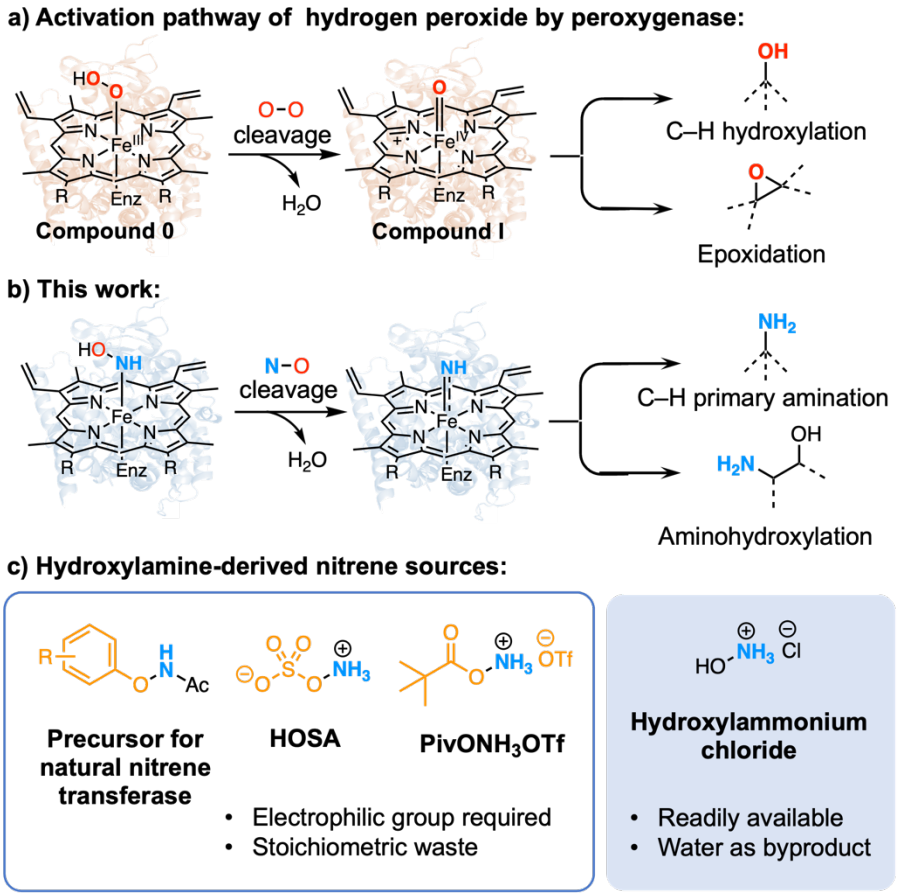
4.1. Introduction

Nature's ability to use simple and abundant precursors for enzymatic functionalization of biomolecules has long inspired reaction discovery. Heme-dependent peroxygenases activate hydrogen peroxide (H_2O_2) to access Compound I, for example, and effect a rich repertoire of oxidation reactions (Scheme 1a).[1] Whereas the use of H_2O_2 for C–H hydroxylation and epoxidation,[2] among other reactions,[3,4,5] has been studied extensively with chemo- and biocatalysts, the development of sustainable and efficient strategies for nitrogen incorporation from the analogous amine feedstock hydroxylamine (NH_2OH) is underexplored.[6] NH_2OH is a structural and isoelectronic analog of H_2O_2, and forming an iron-bound nitrene intermediate from NH_2OH would be analogous to accessing Compound I from the hydroperoxo intermediate during the peroxygenase catalytic cycle.[7] These shared features intimate that heme enzymes could possibly process NH_2OH for nitrogen incorporation (Scheme 1b), enabling amination reactions while generating water as the sole byproduct.

In soil microorganisms, NH_2OH serves as an intermediate in the nitrogen cycle which is oxidized to nitrates, nitrites and nitrous oxide.[8] Despite its natural incidence and frequent exposure to metalloenzymes,[9,10] however, NH_2OH is not known to serve as an amine source for nitrogen incorporation. The only natural 'nitrene transferase' known to date, cytochrome P450BezE, operates on an *O*-acetylated hydroxylamine precursor (Scheme 1c).[11] Similarly, the analogous amination reactions in synthetic chemistry and biocatalysis that have emerged recently use oxygen-substituted hydroxylamine-derived reagents.[12,13]

We previously reported that engineered heme proteins can activate hydroxylamine-derived reagents to form and transfer nitrenes to hydrocarbons.[14] These reagents rely on electron-withdrawing groups on the oxygen to weaken the N–O bond and pull away the bonding pair of electrons in the bond-breaking event *en route* to the nitrene intermediate. While effective in generating myriad C–N bonds, these reagents also generate copious amount of hazardous waste, thus limiting application at scale.

Scheme 4-1. Background and Summary



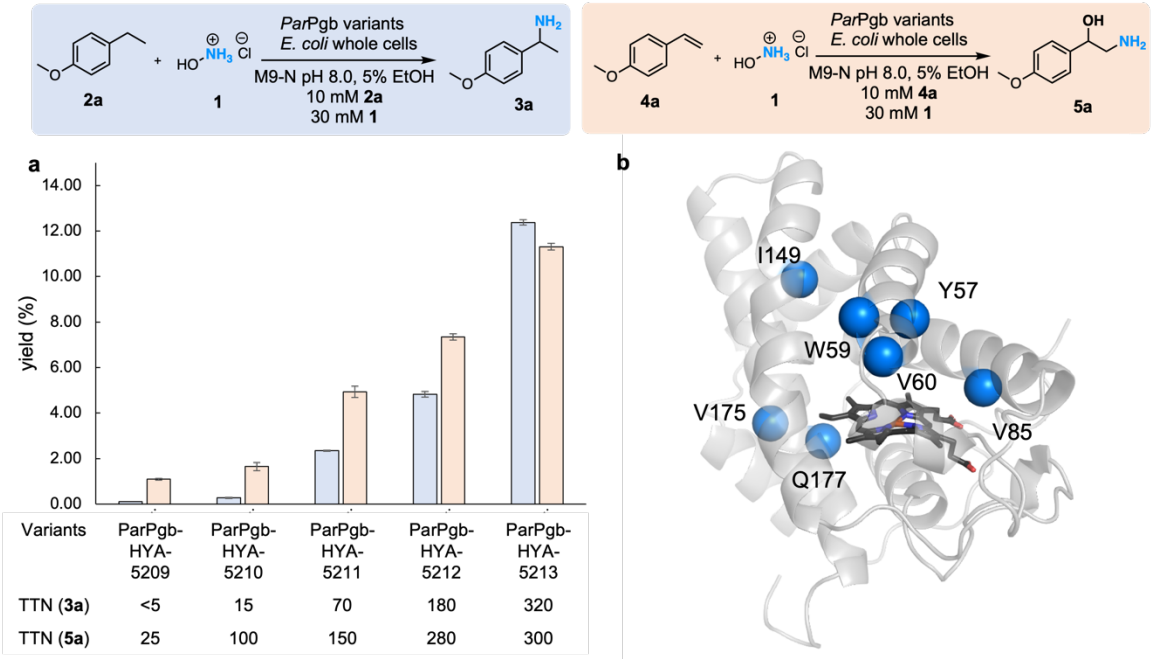
Inspired by peroxygenases, we sought to investigate the activity of heme proteins toward NH_2OH for nitrogen incorporation. To do this, a heme protein must overcome two main challenges. First, compared to electrophilic hydroxylamine-derived reagents, the absence of an electron-withdrawing group in NH_2OH results in an elevated N–O bond dissociation energy. [15,16] Additionally, the heme must selectively bind the nitrogen in NH_2OH instead of oxygen to form a nitrene intermediate.[6] We hypothesized that an enzyme active site could position NH_2OH for efficient bond cleavage and facilitate the release of water. The active sites of peroxygenases contain a highly conserved acid-base catalytic pair that facilitates efficient O–O bond cleavage of H_2O_2 through a proton relay mechanism.[1,7] Indeed, introducing this catalytic pair can even switch the activity of oxygen-dependent heme enzymes to that of peroxygenases.[17,18] We posited that a similar activation mechanism with NH_2OH could be reprised in a heme protein to create a pathway for amine synthesis with NH_2OH that parallels the reactivity of peroxygenases.

4.2. Results and Discussion

We searched for hints of this catalytic activity using the bench-stable hydrochloride salt of NH_2OH **(1)** for two different amination reactions enabled by nitrene chemistry, the benzylic C–H primary amination of 4-ethylanisole **2a** and the aminohydroxylation of 4-vinylanisole **4a**. These two reactions, which share a common reaction pathway via N–O bond cleavage of hydroxylamine precursors, were previously realized using *O*-pivaloylhydroxylammonium triflate ($PivONH_3OTf$) as the amine source,[14a,b] and could reveal a general starting point for exploring enzymatic activity with NH_2OH. A collection

of hemoproteins including cytochromes P450, cytochromes *c*, and globins was examined for both transformations. To our surprise, several engineered enzyme variants exhibited trace levels of product formation (see SI, Section 2.). Among the active protein variants, only *Pyrobaculum arsenaticum* protoglobin W59L V60Q, renamed *Par*Pgb-HYA-5209 (*Par*Pgb Hydroxylamine-dependent Aminase) exhibited measurable reactivity for both reactions, furnishing the benzylic C–H primary amination product **3a** with trace activity and aminohydroxylation product **5a** with ~1% yield. Control experiments with free hemin and bovine serum albumin showed no product formation. Wild type *Par*Pgb, a small, dimeric, gas-binding hemoprotein whose native function is unknown, [19] also does not catalyze these reactions.

*Par*Pgb-HYA-5209 has two mutations that are presumed to create a more open active site, allowing non-native substrates to enter and react. [20] Previously, we reported protoglobin variants that catalyze carbene transfer for cyclopropanation and X–H bond insertions,[21,22,23] but nitrene transfer activity, especially for C–H functionalization, had not been demonstrated. *Par*Pgb-HYA-5209 is an appealing starting point for enzyme engineering due to its high thermostability,[24] which allows it to tolerate activating, but often destabilizing, mutations.[25] Additionally, protoglobin-based biocatalysts can be produced readily in *Escherichia coli* at high levels.



Variants	ParPgb-HYA-5209	ParPgb-HYA-5210	ParPgb-HYA-5211	ParPgb-HYA-5212	ParPgb-HYA-5213
TTN (**3a**)	<5	15	70	180	320
TTN (**5a**)	25	100	150	280	300

Figure 4-1. ***a)*** Evolution trajectory of benzylic C–H primary amination with 4-ethylanisole **2a** (blue) and reactivity of the lineage for aminohydroxylation with 4-vinylanisole **4a** (orange). Experimental details can be found in SI Section C.3.3. **b)** The amino acid positions mutated in *Par*Pgb-HYA-5213 are shown (blue spheres) on the homology model of wildtype *Par*Pgb (grey) based on the crystal structure of *Methanosarcina acetivorans* protoglobin (PDB:2VEE).

Once a promiscuous activity for a new-to-nature reaction is identified in a protein, directed evolution can often enhance that activity.[26,27] We evolved *Par*Pgb-HYA-5209 for benzylic primary amination with **2a** and **1** using a combination of site-saturation mutagenesis (SSM) and random mutagenesis. We identified 20 new, improved variants containing 41 unique mutations (see SI, Section 3.1.). The best of those variants, *Par*Pgb-HYA-5210, was over 4-fold more active, giving 0.27% yield and 15 total turnover numbers (TTN) in whole-cell reactions. We shuffled the leading sequences using staggered extension process (StEP) recombination[28] and screening, for three iterative rounds (Figure 1a). The top variant from recombination, *Par*Pgb-HYA-5213, has seven amino acid substitutions (Y57D, W59L,

V60Q, V85I, I149F, V175A and Q177R) relative to the wild-type protein (Figure 1b). Further SSM targeting these seven amino acids gave no improvement. The final variant exhibited a 45-fold improvement in yield compared to *Par*Pgb-HYA-5210, and over 160-fold in yield compared to *Par*Pgb-HYA-5209 (12.4% yield and 320 TTN). Having established a lineage for benzylic C–H primary amination, we also investigated the activity of the engineered enzymes for aminohydroxylation: the activity with styrene **4a** was increased 10-fold, giving 11.3% yield and 300 TTN in whole-cell reactions (Figure 1a.)

We next investigated the reaction conditions with purified enzyme. During protein purification, we discovered that the heme occupancy of the protoglobins was only ~25% by comparing the protein and heme concentrations. We reasoned that heme production in *E. coli* does not match the high protoglobin protein expression level.[29] Fortunately, we were able to reconstitute the fully charged enzymes by supplementing exogenous heme *b* cofactor to the apo protoglobins.[22,30] The amination reactions can proceed aerobically but are more efficient under anaerobic conditions, increasing the yield of **3a** by ~10% (see SI, Section 9.1.). Further reaction optimization using heme-loaded purified enzyme of variant *Par*Pgb-HYA-5213 with sodium dithionite as reducing reagent under anaerobic conditions greatly improved the yield for benzylic C–H primary amination with **2a**: the reaction with 0.5 mol% enzyme afforded 95.3% analytical yield and 69% isolated yield of primary amine **3a**. The enzymatic product was derivatized to measure enantiomeric excess (ee) (see SI, Section 5.). The reaction gave excellent enantioselectivity of >99% ee, favoring the (*R*)-enantiomer.

To gain insight into how directed evolution improved the reactivity of *Par*Pgb using NH_2OH, we compared the kinetic parameters of *Par*Pgb-HYA-5213 and its parent *Par*Pgb-HYA-5209 for benzylic C–H primary amination using **1** and ethylbenzne **2c** (see SI, Section 7.). The observed turnover number *(*k_{cat}) of *Par*Pgb-HYA-5213 (1.5 s^{-1}) increased 180-fold compared to *Par*Pgb-HYA-5209 (0.0087 s^{-1}), which is likely the major factor contributing to the increase in overall yield observed along the evolutionary lineage. In addition, the observed K_M value for NH_2OH decreased from 5.4 mM to 0.30 mM, suggesting the amino acids introduced during the evolution campaign may improve NH_2OH binding. Finally, we observed that the rate decreases at high concentrations of NH_2OH, reflecting enzyme inactivation and/or substrate inhibition.

We next evaluated the enzymatic activity for benzylic C–H primary amination on various substrates (Figure 2a). Overall, substrates bearing *para* electron-rich and electron-deficient substituents (**3a-e**) are well tolerated, furnishing primary amine products in overall good yield and excellent selectivity. We also examined the substrate scope of aminohydroxylation with styrenes (Figure 2b). Styrenes bearing both electron-rich and electron-deficient groups on the aryl moiety furnished amino alcohol products in good yields (**5a-e**). The enantioselectivities of the amino alcohol products vary from highly enantioselective to racemic. We reasoned that the electron-rich aziridine products undergo hydration via a ring-opened carbocation intermediate whereas electron-deficient substrates likely undergo hydration through an S_N2 pathway.[31]

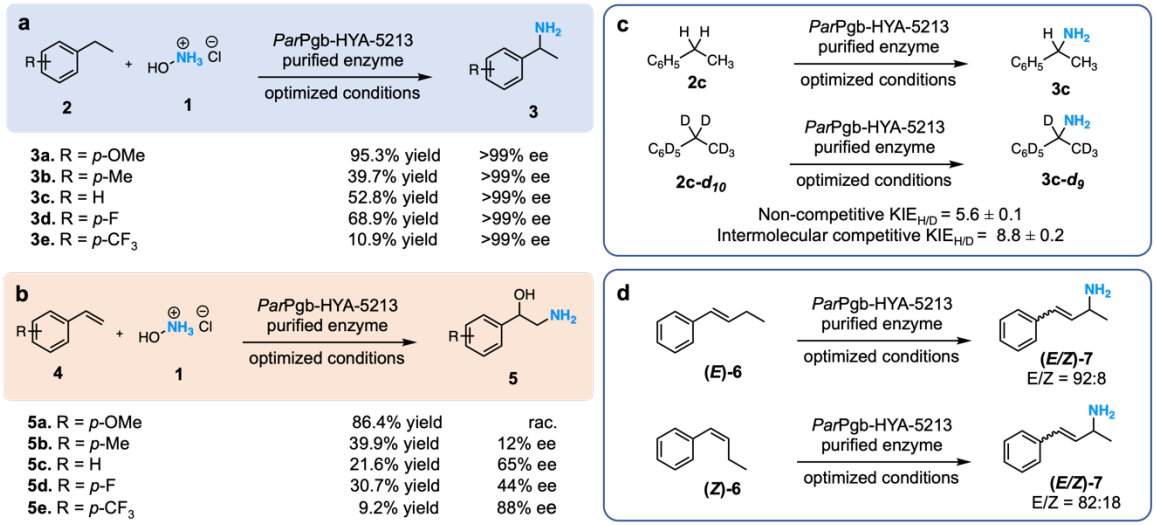


Figure 4-2. Substrate scope and mechanistic study. Experimental details can be found in SI Section C.4, 5, 8 and 9.2. **a)** Substrate scope of the benzylic C–H substrate **b)** Substrate scope of the styrenyl substrate. **c)** KIE experiment using **2c** and **2c-*d*$_{10}$.** **d)** Radical probe experiment using **(*E*)-** and **(*Z*)-6**.

The mechanism of benzylic C–H primary amination with NH_2OH was investigated using kinetic isotope effects (KIE) experiments and a radical probe. A non-competitive KIE value of 5.6 and intermolecular competitive KIE of 8.8 was obtained, indicating C–H bond cleavage as the rate-determining step in the final variant *Par*Pgb-HYA-5213 (Figure 2c). Interestingly, in previous studies of enzymatic benzylic C–H amination with hydroxylamine-derived and azide precursors, the non-competitive KIE values were near unity, indicating nitrene formation was rate-determining.[14c,32] The primary non-competitive KIE value here suggests that nitrene formation with NH_2OH by the final variant is relatively facile. In addition, the relatively large KIE values also suggest a radical-related reaction pathway involving an open-shell nitrene intermediate.[33,34] This agrees with the isomerization of radical probe substrate **(*E/Z*)-6** (Figure 2d), which is consistent with the formation of a carbon-centered radical at the allylic position. Collectively, these results

provide evidence in favor of a stepwise radical mechanism for the C–H amination process.[12,35,36]

Given that wild-type *Par*Pgb cannot use hydroxylamine for amination, we examined the mutations responsible for this novel reactivity. After three rounds of recombination, mutation Y57D, together with mutations W59L and V60Q, were conserved among all the final, most active variants. Furthermore, these residues are directly above the heme cofactor and point towards the heme iron in the generated homology model (see SI, Figure S4). We hypothesized that, much like the key active-site residues in peroxygenases, mutations Y57D and V60Q could serve as an acid-base catalytic pair to activate hydroxylamine through a putative proton relay mechanism, while mutations Y57D and W59L could create a more open active site. We reasoned that introducing these three mutations to other wild type protoglobin homologs could enable them to use NH_2OH. Eleven homologous variants from archaea and bacteria with wild-type protein sequence identities ranging from 51.83% to 82.72% to *Par*Pgb were selected for this mutation transfer experiment (Table 1). These three mutations were introduced at their homologous positions by site-directed mutagenesis to yield eleven *P*gb DLQ variants. All eleven wild-type protoglobins expressed and showed no reactivity towards benzylic C–H primary amination with **2a** and **1**. In contrast, eight out of eleven *P*gb DLQ variants showed activity, although none were more active than *Par*Pgb DLQ. The *Pyrobaculum ferrireducens* protoglobin DLQ variant with 82.72% sequence identity to ParPgb gave the highest activity in terms of TTN, comparable to that of ParPgb DLQ. These results suggest the key role of these mutations in activating hydroxylamine. This diverse library of hydroxylamine-dependent

protoglobins could serve as a starting point for exploring new-to-nature nitrene transfer reactions.

Table 4-1. Activity of mutations transferred variants with 4-ethylanisole **2a** and hydroxylammonium chloride **1**[a]

Variant	Sequence identity (%)	Protein expression (mg per L of cells)	TTN for 3a
*Tpe*Pgb-HYA-5225	51.83	165	< 1
*Mac*Pgb-HYA-5217	53.68	35	0
*Tda*Pgb-HYA-5223	54.21	185	1.5
*Pme*Pgb-HYA-5220	54.45	10	0
*Cth*Pgb-HYA-5216	54.74	180	< 1
*Thu*Pgb-HYA-5224	57.07	180	< 1
*Aau*Pgb-HYA-5214	58.12	195	2.8
*Tam*Pgb-HYA-5221	58.12	170	2.9
*Tar*Pgb-HYA-5222	59.59	190	4.4
*Ape*Pgb-HYA-5215	62.83	20	0
*Pfe*Pgb-HYA-5219	82.72	130	25
*Par*Pgb-HYA-5218	100	220	30

[a]Experimental details can be found in SI Section C.6.

In conclusion, we have demonstrated that a hemoprotein can evolve to use a feedstock chemical, hydroxylammonium chloride, for direct incorporation of nitrogen. This unprecedented reactivity enables benzylic C–H primary amination and alkene aminohydroxylation, with water as the sole byproduct. Mechanistic studies revealed a radical reaction pathway with C–H bond cleavage as the rate-determining step. We envision this discovery will enable new sustainable and economical biocatalytic routes to nitrogen-containing molecules. Finally, we anticipate this work will presage the discovery of natural enzymes that utilize NH_2OH as amine sources.

Chapter IV Bibliography

1. Sigmund, M. C.; Poelarends, G. J. Current state and future perspectives of engineered and artificial peroxygenases for the oxyfunctionalization of organic molecules. *Nat. Catal.* **2020**, *3*(9), 690–702.
2. a) Ullrich, R.; Hofrichter, M. The haloperoxidase of the agaric fungus Agrocybe aegerita hydroxylates toluene and naphthalene. *FEBS lett.* **2005**, *579*(27), 6247–6250.
   b) Peter, S.; Kinne, M.; Wang, X.; Ullrich, R.; Kayser, G.; Groves, J. T.; Hofrichter, M. Selective hydroxylation of alkanes by an extracellular fungal peroxygenase. *FEBS J.* **2011**, *278*(19), 3667–3675.
   c) Peter, S.; Kinne, M.; Ullrich, R.; Kayser, G.; Hofrichter, M. Epoxidation of linear, branched and cyclic alkenes catalyzed by unspecific peroxygenase. *Enzyme Microb. Technol.* **2013**, *52*(6-7), 370–376.
   d) Cheng, L.; Wang, H.; Cai, H.; Zhang, J.; Gong, X.; Han, W. Iron-catalyzed arene C−H hydroxylation. *Science* **2021**, *374*(6563), 77–81.
   e) Masferrer-Rius, E.; Borrell, M.; Lutz, M.; Costas, M.; Klein Gebbink, R. J. Aromatic C− H hydroxylation reactions with hydrogen peroxide catalyzed by bulky Manganese complexes. *Adv. Synth. Catal*. **2021**, *363*(15), 3783–3795.
   f) Mao, S.; Budweg, S.; Spannenberg, A.; Wen, X.; Yang, Y.; Li, Y. W.; Junge, K.; Beller, M. Iron-Catalyzed Epoxidation of Linear α-Olefins with Hydrogen Peroxide. *ChemCatChem.* **2022**, *14*(4), e202101668.
3. Bassanini, I.; Ferrandi, E. E.; Vanoni, M.; Ottolina, G.; Riva, S.; Crotti, M.; Brenna, E.; Monti, D. Peroxygenase-Catalyzed Enantioselective Sulfoxidations. *Eur. J. Org. Chem.* **2017**, *2017*(47), 7186–7189.
4. Bormann, S.; Baraibar, A. G.; Ni, Y.; Holtmann, D.; Hollmann, F. Specific oxyfunctionalisations catalysed by peroxygenases: opportunities, challenges and solutions. *Catal. Sci. Technol.* **2015**, *5*(4), 2038–2052.
5. Calvete, M. J.; Piñeiro, M.; Dias, L. D.; Pereira, M. M. Hydrogen peroxide and metalloporphyrins in oxidation catalysis: old dogs with some new tricks. *ChemCatChem*. **2018**, *10*(17), 3615–3635.
6. a) See, Y. Y.; Sanford, M. S. C–H Amination of Arenes with Hydroxylamine. *Org. Lett.* **2020**, *22*(8), 2931–2934.
   b) Kuznetsova, N. I.; Kuznetsova, L. I.; Detusheva, L. G.; Likholobov, V. A.; Pez, G. P.; Cheng, H. Amination of benzene and toluene with hydroxylamine in the presence of transition metal redox catalysts. *J. Mol. Catal*. **2000**, *161*(1-2), 1–9.
7. a) Huang, X.; Groves, J. T. Oxygen activation and radical transformations in heme proteins and metalloporphyrins. *Chem. Rev.* **2017**, *118*(5), 2491–2553.
   b) Ma, N.; Fang, W.; Liu, C.; Qin, X.; Wang, X.; Jin, L.; Wang, B. and Cong, Z. Switching an Artificial P450 Peroxygenase into Peroxidase via Mechanism-Guided Protein Engineering. *ACS Catal.* **2021**, *11*(14), 8449–8455.
8. Soler-Jofra, A.; Pérez, J.; van Loosdrecht, M. C. Hydroxylamine and the nitrogen cycle: A review. *Water Res.* **2021**, *190*, 116723.

9. Caranto, J. D.; Lancaster, K. M. Nitric oxide is an obligate bacterial nitrification intermediate produced by hydroxylamine oxidoreductase. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*(31), 8217–8222.
10. Caranto, J. D.; Vilbert, A. C.; Lancaster, K. M. Nitrosomonas europaea cytochrome P460 is a direct link between nitrification and nitrous oxide emission. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113*(51), 14704–14709.
11. Tsutsumi, H.; Katsuyama, Y.; Izumikawa, M.; Takagi, M.; Fujie, M.; Satoh, N.; Shin-Ya, K. and Ohnishi, Y. Unprecedented cyclization catalyzed by a cytochrome P450 in benzastatin biosynthesis. *J. Am. Chem. Soc.* **2018**, *140*(21), 6631–6639.
12. Gasser, V. C.; Makai, S.; Morandi, B. The advent of electrophilic hydroxylamine-derived reagents for the direct preparation of unprotected amines. *Chem. Commun.* **2022**, *58*(72), 9991–10003.
13. Gao, Y.; Li, H.; Zhao, Y.; Hu, X. Q. Nitrene transfer reaction with hydroxylamine derivatives. *Chem. Commun.* **2023**, 59, 1889–1906.
14. a) Cho, I.; Prier, C. K.; Jia, Z. J.; Zhang, R. K.; Görbe, T.; Arnold, F. H. Enantioselective aminohydroxylation of styrenyl olefins catalyzed by an engineered hemoprotein. *Angew. Chem. Int. Ed*. **2019**, *58*(10), 3138–3142.
b) Jia, Z. J.; Gao, S.; Arnold, F. H. Enzymatic primary amination of benzylic and allylic C (sp3)–H bonds. *J. Am. Chem. Soc.* **2020**, *142*(23), 10279–10283.
c) Athavale, S. V.; Gao, S.; Liu, Z.; Mallojjala, S. C.; Hirschi, J. S.; Arnold, F. H. Biocatalytic, intermolecular C− H bond functionalization for the synthesis of enantioenriched amides. *Angew. Chem. Int. Ed.* **2021**, *60*(47), 24864–24869.
d) Liu, Z.; Qin, Z.-Y.; Zhu, L.; Athavale, S. V.; Sengupta, A.; Jia, Z. J.; Garcia-Borràs, M.; Houk, K. N. and Arnold, F. H. An enzymatic platform for primary amination of 1-aryl-2-alkyl alkynes. *J. Am. Chem. Soc.* **2021**, *144*(1), 80-85.
e) Athavale, S.V.; Gao, S.; Das, A.; Mallojjala, S. C.; Alfonzo, E.; Long, Y.; Hirschi, J. S. and Arnold, F. H. Enzymatic nitrogen insertion into unactivated C–H bonds. *J. Am. Chem. Soc.* **2022**, *144*(41), 19097–19105.
15. Bach, R. D.; Schlegel, H. B. The bond dissociation energy of the N–O bond. *J. Phys. Chem.* **2021**,*125*(23), 5014–5021.
16. Bach, R. D.; Schlegel, H. B. Bond dissociation energy of peroxides revisited. *J. Phys. Chem.* **2020**,*124*(23), 4742–4751.
17. Shoji, O.; Fujishiro, T.; Nishio, K.; Kano, Y.; Kimoto, H.; Chien, S.C.; Onoda, H.; Muramatsu, A.; Tanaka, S.; Hori, A. and Sugimoto, H. A substrate-binding-state mimic of H_2O_2-dependent cytochrome P450 produced by one-point mutagenesis and peroxygenation of non-native substrates. *Catal. Sci. Technol.* **2016**, *6*(15), 5806–5811.
18. Ozaki, S. I.; Matsui, T.; Watanabe, Y. Conversion of myoglobin into a peroxygenase: a catalytic intermediate of sulfoxidation and epoxidation by the F43H/H64L mutant. *J. Am. Chem. Soc.* **1997**, *119*(28), 6666–6667.
19. Pesce, A.; Bolognesi, M.; Nardini, M. Protoglobin: structure and ligand-binding properties. *Adv. Microb. Physiol*. **2013**, *63*, 79–96.

20. Knight, A. M. *Expanding the Scope of Metalloprotein Families and Substrate Classes in New-to-Nature Reactions,* **2021**, Doctoral dissertation, California Institute of Technology.
21. Knight, A. M.; Kan, S. J.; Lewis, R. D.; Brandenberg, O. F.; Chen, K.; Arnold, F. H. Diverse engineered heme proteins enable stereodivergent cyclopropanation of unactivated alkenes. *ACS Cent. Sci.* **2018**, *4*(3), 372–377.
22. Porter, N. J.; Danelius, E.; Gonen, T.; Arnold, F. H. Biocatalytic carbene transfer using diazirines. *J. Am. Chem. Soc.* **2022**, *144*(20), 8892–8896.
23. Schaus, L.; Das, A.; Knight, A.M.; Jimenez-Osés, G.; Houk, K. N.; Garcia-Borràs, M.; Arnold, F. H. and Huang, X. Protoglobin-Catalyzed Formation of cis-Trifluoromethyl-Substituted Cyclopropanes by Carbene Transfer. *Angew. Chem. Int. Ed.* **2023**, *62*(4), e202208936.
24. Freitas, T. A. K.; Hou, S.; Dioum, E. M.; Saito, J. A.; Newhouse, J.; Gonzalez, G.; Gilles-Gonzalez, M.-A.; Alam, M. Ancestral hemoglobins in Archaea. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, 101 (17), 6675−6680.
25. Bloom, J. D.; Labthavikul, S. T.; Otey, C. R.; Arnold, F. H. Protein stability promotes evolvability. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, 103 (15), 5869−5874.
26. Arnold, F. H. Design by directed evolution. *Acc. Chem. Res.* **1998**, *31*(3), 125–131.
27. Arnold, F. H. Directed evolution: bringing new chemistry to life. *Angew. Chem. Int. Ed.* **2018**, *57*(16), 4143–4148.
28. Zhao, H.; Giver, L.; Shao, Z.; Affholter, J. A.; Arnold, F. H. Molecular Evolution by Staggered Extension Process (StEP) in Vitro Recombination. *Nat. Biotechnol.* **1998**, *16* (3), 258–261.
29. Hu, B.; Yu, H.; Zhou, J.; Li, J.; Chen, J.; Du, G.; Lee, S.Y.; Zhao, X. Whole-Cell P450 Biocatalysis Using Engineered Escherichia coli with Fine-Tuned Heme Biosynthesis. *Adv. Sci.* **2023**, *10*(6), 2205580.
30. Pesce, A.; Tilleman, L.; Donné, J.; Aste, E.; Ascenzi, P.; Ciaccio, C.; Coletta, M.; Moens, L.; Viappiani, C.; Dewilde, S.; Bolognesi, M.; Nardini, M. Structure and haem-distal site plasticity in Methanosarcina acetivorans protoglobin. *PLoS One.* **2013**, *8*(6), p.e66144.
31. Farwell, C. C.; Zhang, R. K.; McIntosh, J. A.; Hyster, T. K.; Arnold, F. H. Enantioselective enzyme-catalyzed aziridination enabled by active-site evolution of a cytochrome P450. *ACS Cent. Sci.* **2015**, *1*(2), 89–93.
32. 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*(7), 629–634.
33. Hong, S. Y.; Hwang, Y.; Lee, M.; Chang, S. Mechanism-Guided Development of Transition-Metal-Catalyzed C–N Bond-Forming Reactions Using Dioxazolones as the Versatile Amidating Source. *Acc. Chem. Res.* **2021**, *54*(11), 2683-2700.
34. Simmons, E. M.; Hartwig, J. F. On the Interpretation of Deuterium Kinetic Isotope Effects in C–H Bond Functionalizations by Transition-Metal Complexes. *Angew. Chem. Int. Ed.* **2012**, *51*(13), 3066-3072.

35. 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*(11), 987–993.
36. Chatterjee, S.; Harden, I.; Bistoni, G.; Castillo, R.G.; Chabbra, S.; van Gastel, M.; Schnegg, A.; Bill, E.; Birrell, J.A.; Morandi, B. Neese, F.; DeBeer, S. A Combined Spectroscopic and Computational Study on the Mechanism of Iron-Catalyzed Aminofunctionalization of Olefins Using Hydroxylamine Derived N–O Reagent as the "Amino" Source and "Oxidant". *J. Am. Chem. Soc.* **2022**. *144*(6), 2637–2656.

Appendix C
SUPPLEMENTARY INFORMATION FOR CHAPTER IV

C.1. General Procedures

Unless otherwise noted, all chemicals and reagents were obtained from commercial suppliers (Sigma-Aldrich, VWR, Alfa Aesar, Ambeed, Combi-Blocks, and Enamine) 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 Varian Inova 300 MHz or Bruker 400 MHz instrument in $CDCl_3$ and are referenced to residual proton solvent signals. Data for ^{1}H NMR are reported in the conventional form: chemical shift (δ ppm), multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, hept = heptet, m = multiplet, br = broad), coupling constant (Hz), and integration. Sonication was performed using a Qsonica Q500 sonicator. Chemical reactions were monitored using thin layer chromatography (Merck 60 silica gel plates) and a UV lamp for visualization, if possible. High-performance liquid-chromatography mass spectroscopy (HPLC-MS) for analysis was carried out using Agilent 1200 series instruments, with C18 (Agilent Poroshell 120 EC-C18, 4.6 × 50 mm, 5 μm) columns. Water and acetonitrile containing 0.1% acetic acid were used as eluents. High-performance liquid-chromatography (HPLC) for analysis was carried out using Agilent 1200 series instruments, with C18 (Agilent Poroshell 120 EC-C18, 4.6 × 50 mm, 100mm and 150mm, 5 μm) columns. Solvent A: 50mM Na_2HPO_4 buffer (pH 7.0) and solvent B: in 25 mM Na_2HPO_4 in 50% acetonitrile/water were used as eluents. Distilled water was utilized for growth media while double distilled water was employed

for all buffer preparations. Normal-phase chiral HPLC was performed using Daicel Chiralcel OJ-H columns (4.6 × 250 mm, 5 μm) with hexane and isopropanol as eluents.

The M9-N buffer described in this study was prepared as follows: A 5X stock solution was prepared by dissolving Na_2HPO_4 (34 g), KH_2PO_4 (15 g), and NaCl (2.5 g) in 1 L of water followed by sterilization by autoclaving. A 1X stock was then prepared by diluting 200 mL of the 5X solution to 1 L, followed by addition of 1 mL $CaCl_2$ (0.1 M), 2 mL $MgSO_4$ (1.0 M), and adjusting the pH to 8.0 by the addition of 40% aqueous NaOH.

The KPi buffer described in this study was prepared as follows: The buffer was prepared by dissolving K_2HPO_4 (8.1 g), KH_2PO_4 (0.44 g), and NaCl (8.8 g) in 1 L of water and adjusting the pH to 8.0 by the addition of 40% aqueous NaOH followed by sterilization by filtration.

C.1.1. Cloning and Mutagenesis

T7 Express competent *Escherichia coli* (BL21) (New England Biolabs) cells were utilized for all experiments. Luria-Bertani (LB) and Hyper-Broth™ (HB, AthenaES) media were used for growth. During growth, media were supplemented with ampicillin (100 μg/mL). HB was also supplemented with Glucose Mix (AthenaES) according to the manufacturer instructions. Protein induction was performed by supplementing with isopropyl β-D-1-thiogalactopyranoside (IPTG, 0.5 mM) and δ-aminolevulinic acid (ALA, 1.0 mM). Expression vector pET22b(+) (Novagen) was used for cloning and expression of all variants described in this paper.

Site-saturation mutagenesis (SSM) was performed using a modified QuikChange™ mutagenesis protocol using the 22-codon trick.[1] The PCR products were purified with New England Biolabs gel purification kit, and the gaps were repaired using Gibson Mix™.[2] Without further purification, 1 μL of the Gibson product was used to transform 50 μL of competent *E. coli* cells.

Random mutations were introduced using error-prone PCR with the addition of 200–500 μM $MnCl_2$ to a Taq PCR method as previously reported using the primers described below.[3]

Table C-1. Primer Used in Error-Prone PCR.

Primers	Sequence (5' → 3')
005	GAA ATA ATT TTG TTT AAC TTT AAG AAG GAG ATA TAC ATA TG
006	GCC GGA TCT CAG TGG TGG TGG TGG TGG TGC TCG AG
007	CAT ATG TAT ATC TCC TTC TTA AAG TTA AAC AAA ATT ATT TC
008	CTC GAG CAC CAC CAC CAC CAC CAC TGA GAT CCG GC

Primers 005 and 006 were used to amplify the region coding for the enzyme (full length) using Taq polymerase and the parent DNA template. Mutations within the full-length fragment were introduced by varying the concentration of $MnCl_2$ during PCR amplification. The backbone (i.e.; pET22b(+) vector) fragment was amplified following the SSM protocol using Primers 007 and 008. The excess DNA templates were digested with DpnI, and the amplified fragments were purified by gel electrophoresis (1% agarose gel). Both the amplified fragments were then assembled into a circular plasmid using the

Gibson assembly.[2] Libraries generated with 200, 300, 400, and 500 μM $MnCl_2$ were tested (one 96-well plate each) to determine which library gave the optimal balance of high diversity and low rate of inactivation (approximately 50% variants are inactive and 50% have similar or better activity compared to the parent). The chosen library was then tested further.

Staggered extension process (StEP) PCR was conducted by pooling plasmids of the variants to be recombined in equimolar fashion and using this mixture as the template DNA.[4] This method recombines multiple hits obtained from SSM and error-prone PCR libraries with high efficiency by performing special PCR cycles, where the elongation step is disrupted earlier by denaturation. The procedure has been designed to always require two PCRs: one to recombine the sequences as efficiently as possible and another to amplify the recombination product from the first PCR to create a useful amount of product. Primers 005 and 006 were used to amplify the region coding for the enzyme (full length) using Taq polymerase and the parent DNA template. After examined by gel electrophoresis (1% agarose gel), the successful recombination fragment with lowest the reaction temperature was chosen for library. The amplified fragments were then assembled into a circular plasmid using the Gibson assembly.[2]

C.1.2. Protein Expression for Whole-Cell Reactions

E. coli transformed with pET22b(+) constructs encoding various protoglobin variants were grown overnight in 5-mL LB medium supplemented with ampicillin (LB_{amp}). Subsequently, 1 mL of this preculture was used to inoculate 50 mL of HB medium in a

125-mL Erlenmeyer flask, supplemented with ampicillin (HB_{amp}) and Glucose Mix for expression. The expression culture was incubated at 37 °C and shaken at 220 rpm for 2.5 hours until the optical cell density at 600 nm (OD_{600}) was ~0.9 – 1.1. Then, the expression culture was cooled in an ice bath for 30–50 minutes and was treated with ALA and IPTG. Cells were grown at 22 °C and 130 rpm for 20–24 hours, and the shaking radius was 25 mm. Once expression was finished, the cultures were centrifuged (4,000*g*, 3 minutes, 4 °C) and the cell pellets were resuspended to an OD_{600} of 60 (or other appropriate OD as necessary) in M9-N buffer with pH adjusted to 8.0. This cell suspension was utilized to setup whole-cell reactions.

C.1.3. Determination of Heme and Protein Concentration

Heme concentration was determined by performing hemochrome assay with lysate or purified protein.[5] 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 or purified protein 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 heme concentration was calculated using the extinction coefficient: $\varepsilon^{[557\text{reduced}-540\text{oxidized}]} = 23.98\ mM^{-1}cm^{-1}$ and dilution factors. Please note that the hemochrome assay detects the levels of heme, not necessarily concentration of the enzyme. Purified protein concentration was determined

with a NanoDrop spectrophotometer ND-1000. The molar extinction coefficient ($mM^{-1}cm^{-1}$) and molar mass (kDa) of each enzyme variant was calculated by ProtParam. The protein concentration was calculated by the average of three measurements. We also attempted to use the BCA assay, but free heme gave background strong background interference.

The heme occupancy of purified enzyme was calculated by comparing the heme concentration by hemochrome assay and protein concentration by NanoDrop experiment.

C.1.4. Reaction Screening in 96-Well Plate Whole-Cell Format

After the library generation by SSM, error-prone PCR, or StEP recombination, single colonies from LB_{amp} agar plates (ampicillin) were randomly picked and cultured in 300 µL of LB medium with 0.1 mg/mL ampicillin (LB_{amp}) in a sterilized 96-well deep-well plate. Each plate typically contained six wells inoculated with single colonies expressing the parent enzyme (previous best variant) and two sterile wells. The cultures were grown at 37 °C, 250 rpm, and 80% relative humidity for 8–12 hours. A separate, sterilized 96-well deep-well plate was filled with 950 µL of HB medium containing 0.1 mg/mL ampicillin (HB_{amp}) in each well. Likewise, a glycerol stock replica plate of the preculture (50 µL of preculture added to 50 µL of 50% glycerol per well) was also prepared and stored at -80 °C for future reference. The plate with HB_{amp} was inoculated with the LB preculture (50 µL/well) and incubated at 37 °C, 250 rpm, and 80% relative humidity for 2.5 hours. The plate was cooled on an ice bath for 1 hour, induced with 0.5 mM IPTG and 1 mM ALA (final concentrations), and then expressed at 22 °C and 220 rpm for 20–22 hours. The cells

were pelleted (4,000*g*, 3 minutes, 4 °C), the supernatant was discarded, and 360 μL of M9-N (pH 8.0) were added to each well. After the cell pellets were fully resuspended by shaking at 500 rpm, we added 20 μL of each reactant stock solution (200 mM of benzylic substrate dissolved in ethanol and 600 mM of hydroxylammonium chloride dissolved in water) to each well of the 96-well plate. Thus, the final reactant concentrations in each well were 10 mM (hydrocarbon) and 30 mM (hydroxylammonium chloride). The plate was sealed with aluminum foil tape and shaken at 600 rpm at room temperature overnight. We performed the 96-well plate screening aerobically.

Once the seal was removed, ethanol (800 μL/well) was added. The resulting suspension in the wells was mixed by pipetting. The plate was then centrifuged (4,500*g*, 5 minutes, 4 °C) to precipitate proteins and cell debris. The supernatant (200 μL/well) was transferred to an analytical 96-well plate for reverse-phase HPLC-MS analysis. The MS product signals from each well were compared, and wells showing signals higher than the parent wells were identified. These 'hits' were recultured using the wells from the frozen replica glycerol stock plate and then sequenced. In this way, the better variants showing mutations were selected as new parents for the next round of directed evolution.

C.1.5. Protein Purification

E. coli transformed with pET22b(+) constructs encoding various protoglobin variants were grown overnight in 50-mL LB medium supplemented with ampicillin (LB_{amp}). Subsequently, 20 mL of this preculture were used to inoculate 1 L of HB medium in a 2-L Erlenmeyer flask, supplemented with ampicillin (HB_{amp}) and Glucose Mix, for expression.

The expression culture was incubated at 37 °C and shaken at 140 rpm for 2.5 hours, until the optical cell density at 600 nm (OD_{600}) was ~0.9–1.1. Then, the expression culture was cooled in an ice bath for 30–50 minutes and was treated with ALA and IPTG. Cells were grown at 22 °C and 90 rpm for 20–22 hours, and the shaking radius was 25 mm. Once expression was finished, the cultures were centrifuged (4000*g*, 5 minutes, 4 °C) and the cell pellet was frozen at -20 °C until further processing.

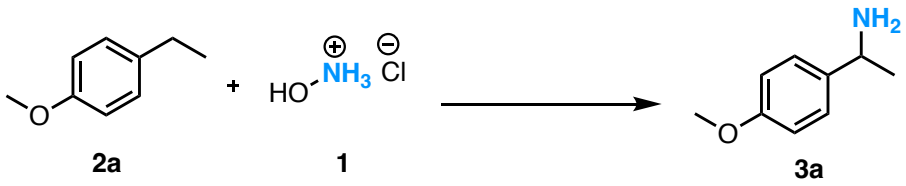
After freeze-thawing thrice, the cell pellet was resuspended in 60 mL of KPi buffer and 60 mg of lysozyme lysate powder from chicken egg white, 10 mg of DNase I lysate powder from bovine pancreas, and 60 μL of 2 M magnesium chloride solution (filtered) were added. The resuspended solution was incubated at 37 °C for one hour before it was centrifuged (14000*g*, 15 minutes, 4 °C). The supernatant was applied to a 5-mL Ni^{2+}-NTA column equilibrated with 50 mM KPi (pH 8.0) at 4 °C. Once all lysate had been applied, the column was washed with 50 to 100 mL of buffer A (KPi buffer with 20 mM imidazole, pH 8.0) until the flow-through was clear. The protein-of-interest was eluted with 5 to 10 mL buffer B (KPi buffer with 500 mM imidazole, pH 8.0). Fractions were pooled and concentrated to >2.5 mL with 10 kDa MWCO filter. The concentrated solution was dialyzed in 4 L of KPi buffer at 4 °C overnight. The protein concentration was measured by NanoDrop, and the heme concentration was measured by hemochrome assay. Next, exogenous heme *b* cofactor (20 mM stock solution in 0.1 M NaOH solution) was supplemented according to the difference between protein concentration and heme concentration. The protein solution was chilled on ice for 30 min and centrifuged (5000*g*, 10 minutes, 4 °C). The protein was further purified via size exclusion chromatography with

an ÄKTAxpress system equipped with a HiLoad 16/600 Superdex 75pg column to remove excess heme and other impurities. Finally, the protein concentration and the heme concentration were measured again. The concentrated, purified protein (0.4 – 1 mM stock) was flash-frozen and stored at -80 °C.

C.2. Discovery of Initial Activity

The Arnold lab collection of heme enzymes (including P450s, P411s, protoglobins, and cytochromes *c*) was screened for product formation in whole-cell reactions in 96-well plates (see section 1.4). The collection consisted of nearly 400 distinct variants accumulated from prior directed evolution campaigns. No product was detected in control reactions with free heme. Initial activity screening was performed anaerobically.

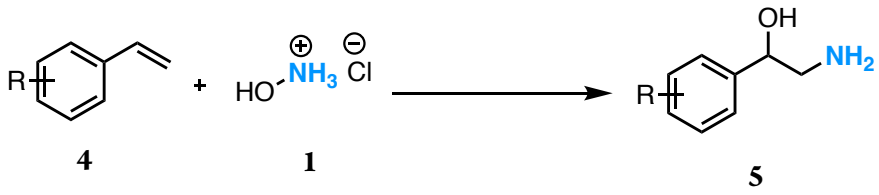
Benzylic primary amination variants showing product formation:



2a **1** **3a**

Variants	yield %
P411 CA-G8	1.9
ParPgb W59L V60Q	trace
5 mol% Heme + BSA	0
5 mol% Heme + BSA + 10 mM sodium dithionite	0

Aminohydroxylation variants showing product formation:



4 **1** **5**

Variants	yield %
P411-CIS variant[6]	5.4

P411 BPA3[7]	3.2
ApePgb AGAGW[8]	0.33
Rma cyt *c* TQL variant[9]	trace
Wildtype chloroperoxidase from *Caldariomyces fumago*	0.18
ParPgb W59L V60Q	1.1
5 mol% Heme + BSA	0
5 mol% Heme + BSA + 10 mM sodium dithionite	0

C.3. Directed Evolution for the *Par*Pgb-HYA Lineage

C.3.1. Summary of Directed Evolution

3.1.1. Summary of the Initial Round by SSM Libraries and Error-prone Library

The initial round of directed evolution was performed by SSM targeting the active site and random mutagenesis targeting the whole protein. The best variants from both libraries are shown below:

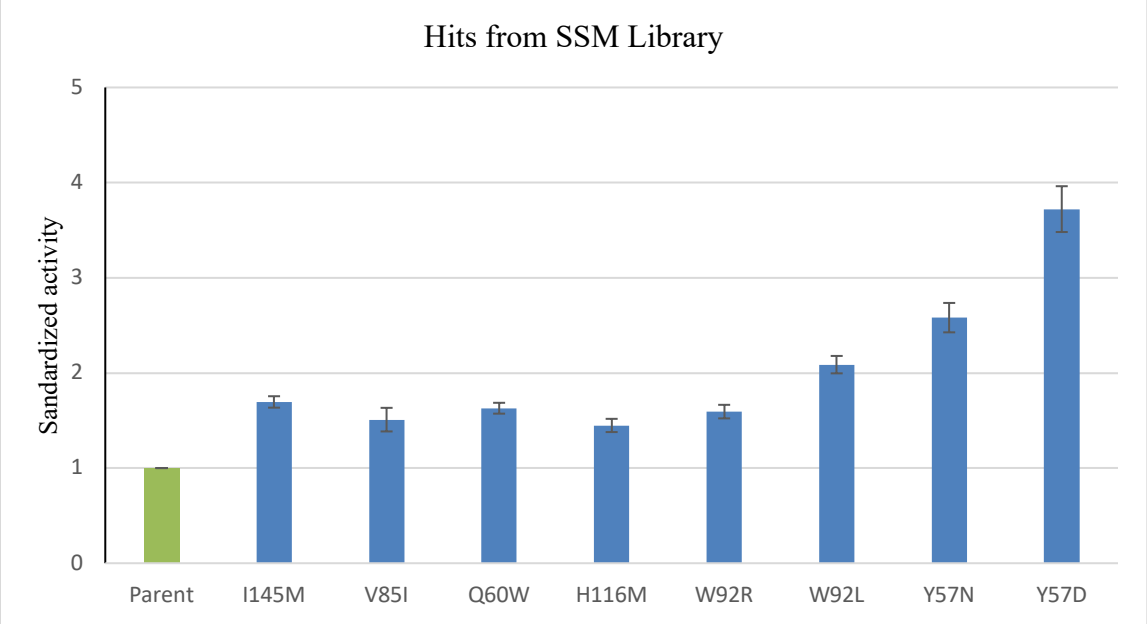


Figure C-1. Hits from the SSM libraries.

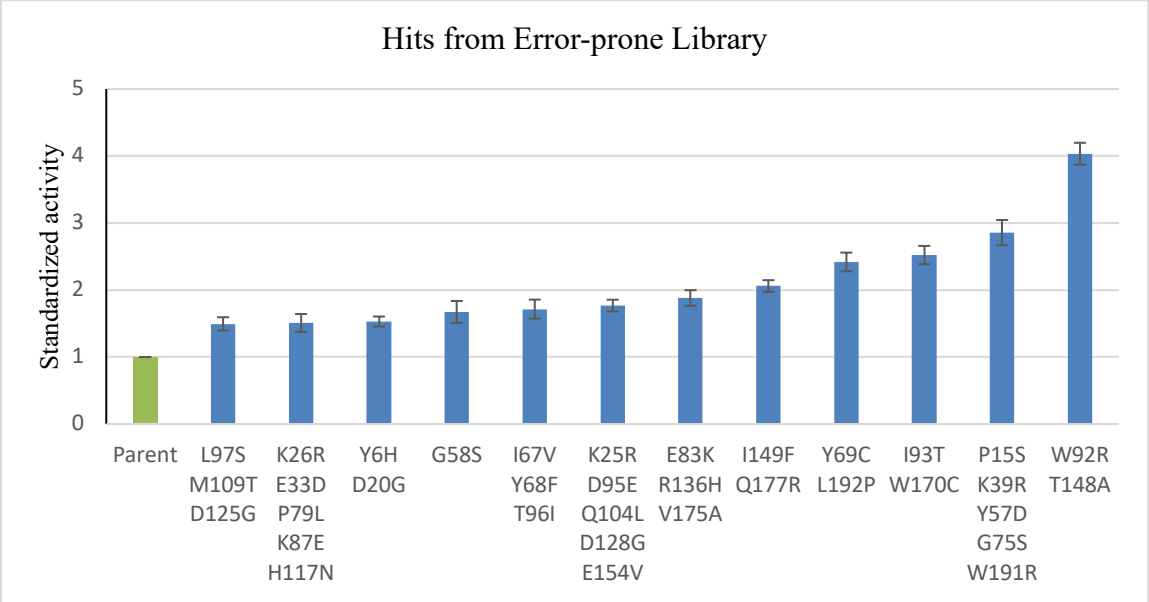


Figure C-2. Hits from the random mutagenesis library.

3.1.2. Summary of StEP Recombination

The subsequent three rounds of directed evolution were performed by StEP recombination of all variants that were more active than the parent variants. No new mutations were introduced during these three rounds. The leading variants from each round are shown below:

Round	Method	Leading Variant	Yield of **3a**	TTN of **3a**
0	--	ParPgb-HYA-5209	trace	trace
1	SSM + Error Prone	ParPgb-HYA-5210	0.27±0.03	15 ± 2.1
2	StEP Recombination of Top 20	ParPgb-HYA-5211	2.34±0.03	70 ± 1.0

3	StEP Recombination of Top 13	ParPgb-HYA-5212	4.83±0.12	180 ± 4.7
4	StEP Recombination of Top 7	ParPgb-HYA-5213	12.4±0.12	320 ± 11

3.1.3. Summary of Lineage Variants

Variant Name	Mutations Relative to Wild-type ParPgb	Mutations Relative to ParPgb-HYA-5209
ParPgb-HYA-5209	W59L V60Q	--
ParPgb-HYA-5210	W59L V60Q W92R T148A	W92R T148A
ParPgb-HYA-5211	K26R E33D Y57D W59L V60Q I149F Q177R	K26R E33D Y57D I149F Q177R
ParPgb-HYA-5212	Y57D W59L V60Q I93T H116M	Y57D I93T H116M
ParPgb-HYA-5213	Y57D W59L V60Q V85I I149F V175A Q177R	Y57D V85I I149F V175A Q177R

C.3.2. DNA and Protein Sequences

3.2.1. ParPgb-HYA-5209 DNA sequence:
ATGGCGGTTCCCGGCTACGATTTTGGCAAAGTCCCGGATGCCCCAATCTCAG
ACGCGGATTTTGAGAGTTTAAAAAAAACCGTGATGTGGGGTGAGGAAGATGA
GAAATATCGCAAAATGGCTTGCGAAGCCTTAAAGGGTCAAGTAGAAGATATT
TTAGATTTGTGGTACGGCCTGCAGGGAAGCAATCAACACCTTATCTACTACTT
CGGTGATAAGAGTGGTCGTCCAATTCCGCAATACCTGGAAGCGGTCCGCAAG
CGTTTCGGGTTGTGGATCATTGATACATTGTGTAAGCCACTGGACCGCCAGTG
GTTGAATTACATGTACGAAATTGGCCTTCGCCATCACCGTACCAAGAAAGGG
AAGACAGATGGCGTAGATACTGTTGAACATATCCCATTACGCTACATGATTG
CTTTCATCGCTCCCATCGGTCTGACTATTAAGCCGATCTTGGAAAAATCGGGA
CATCCGCCAGAGGCCGTGGAGCGTATGTGGGCAGCATGGGTTAAGTTGGTGG
TGTTACAGGTAGCTATCTGGTCGTACCCCTATGCAAAGACGGGCGAATGGCT
CGAGCACCACCATCACCACCAC

3.2.2. ParPgb-HYA-5209 Protein sequence:
MAVPGYDFGKVPDAPISDADFESLKKTVMWGEEDEKYRKMACEALKGQVEDIL
DLWYGLQGSNQHLIYYFGDKSGRPIPQYLEAVRKRFGLWIIDTLCKPLDRQWLN
YMYEIGLRHHRTKKGKTDGVDTVEHIPLRYMIAFIAPIGLTIKPILEKSGHPPEAVE
RMWAAWVKLVVLQVAIWSYPYAKTGEWLEHHHHHH

3.2.3. ParPgb-HYA-5213 DNA sequence:
ATGGCGGTTCCCGGCTACGATTTTGGCAAAGTCCCGGATGCCCCAATCTCAG
ACGCGGATTTTGAGAGTTTAAAAAAAACCGTGATGTGGGGTGAGGAAGATGA
GAAATATCGCAAAATGGCTTGCGAAGCCTTAAAGGGTCAAGTAGAAGATATT
TTAGATTTGTGGGACGGCCTGCAGGGAAGCAATCAACACCTTATCTACTACTT
CGGTGATAAGAGTGGTCGTCCAATTCCGCAATACCTGGAAGCGATTCGCAAG
CGTTTCGGGTTGTGGATCATTGATACATTGTGTAAGCCACTGGACCGCCAGTG
GTTGAATTACATGTACGAAATTGGCCTTCGCCATCACCGTACCAAGAAAGGG
AAGACAGATGGCGTAGATACTGTTGAACATATCCCATTACGCTACATGATTG
CTTTCATCGCTCCCATCGGTCTGACTTTTAAGCCGATCTTGGAAAAATCGGGA
CATCCGCCAGAGGCCGTGGAGCGTATGTGGGCAGCATGGGTTAAGTTGGTGG
CGTTACGGGTAGCTATCTGGTCGTACCCCTATGCAAAGACGGGCGAATGGCT
CGAGCACCACCACCACCACCAC

3.2.4. ParPgb-HYA-5213 Protein sequence:
MAVPGYDFGKVPDAPISDADFESLKKTVMWGEEDEKYRKMACEALKGQVEDIL
DLWDGLQGSNQHLIYYFGDKSGRPIPQYLEAIRKRFGLWIIDTLCKPLDRQWLNY
MYEIGLRHHRTKKGKTDGVDTVEHIPLRYMIAFIAPIGLTFKPILEKSGHPPEAVE
RMWAAWVKLVALRVAIWSYPYAKTGEWLEHHHHHH

C.3.3. Product Quantitation and TTNs

C.3.3.1. Analytical Scale Reaction Setup (Whole Cells)

M9-N buffer (pH 8.0) were placed in the Coy anaerobic chamber with an oxygen concentration below 20 ppm at least 24 hours. Harvested cells (see section 1.2) were resuspended with M9-N buffer (pH 8.0) in the anaerobic chamber, and the resuspended cells were aliquoted to 2-mL screw cap vials. M9-N medium were added to make the total volume to 360 µL of OD_{600} = 67. The hydrocarbon substrate (20 µL, 200 mM stock in ethanol) and hydroxylammonium chloride (20 µL, 600 mM stock in water) were added in a sequential manner. The reactions were then shaken at room temperature for 12 hours at 200 rpm.

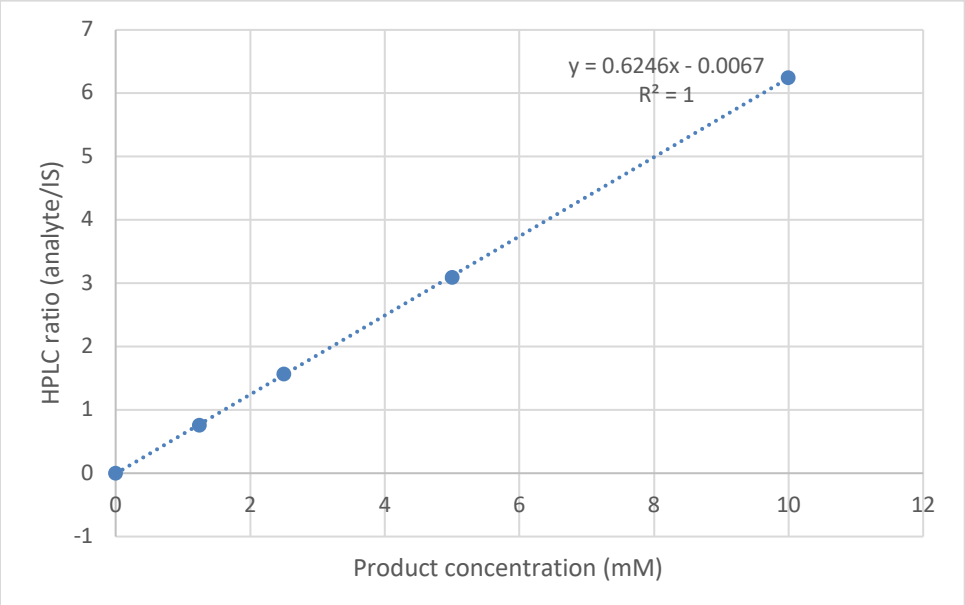
To 1.7-mL Eppendorf microcentrifuge tubes, 400 µL ethanol (10 mM 1,3,5-trimethoxybenzene as internal standard) were added, and 400 µL of the enzymatic reaction mixtures were subsequently added. After vortexing the mixtures, the microcentrifuge tubes were centrifuged (14,000*g*, 10 minutes, 4 °C) to precipitate coagulated proteins and cell debris. To identify the products and determine yields, 200 µL of the supernatants were transferred to a 2-mL screw cap vial with an insert and analyzed by reverse-phase HPLC-MS. Product formation was quantified by HPLC-MS based on the calibration curve of the corresponding racemic reference compound.

C.3.3.2. Calibration Curves for Standard Products and Reaction Product Quantitation

Enzymatic reactions on analytic scale were performed following the general procedure described in Section 3.3.1. Product formation was quantified by HPLC-MS based on the calibration curve of the corresponding racemic reference compound. TTN is defined as the concentration of product divided by the concentration of heme protein measured by the hemochrome assay (Section 1.3). Calibration curves of commercially available or synthesized reference compounds were created for the determination of yield and TTN. Harvested cells were resuspended with M9-N buffer (pH 8.0) to OD_{600} = 63, and 380 µL of resuspended cells were aliquoted to 2-mL screw cap vials. Product standard stock (20 µL; 200 mM, 100 mM, 50 mM, and 25 mM stock in ethanol) was added to the mixture. The mixture was shaken overnight at room temperature, quenched with 400 µL of ethanol (10 mM 1,3,5-trimethoxybenzene as internal standard), vortexed, and then analyzed by HPLC. The UV absorbance at selected wavelengths of product over the indicated internal standard were measured. Methods were adjusted so that none of the products eluted at the same time with any of the cell impurities. For all the analyses of whole-cells reactions, solvent A: water containing 0.1% acetic acid and solvent B: acetonitrile containing 0.1% acetic acid were used as eluents for a C18 (Agilent Poroshell 120 EC-C18, 4.6 × 50 mm, 5 µm) column. The methods used 5–80% solvent B (0.5–5.5 minutes), 80% solvent B (5.5–6.0 minutes), and 80–5% solvent B (6.0–6.5 minutes) with 5-µL sample injections. The flow rate was 1.5 mL/minute, and the column was maintained at 37 °C.

C.3.3.3. Yields and TTNs for the 4-Ethylanisole (2a) Reaction across the Lineage (Figure 4-2a)

Standard Curve based on UV absorbance at 280 nm for product **3a** and 1,3,5-trimethoxybenzene as internal standard:



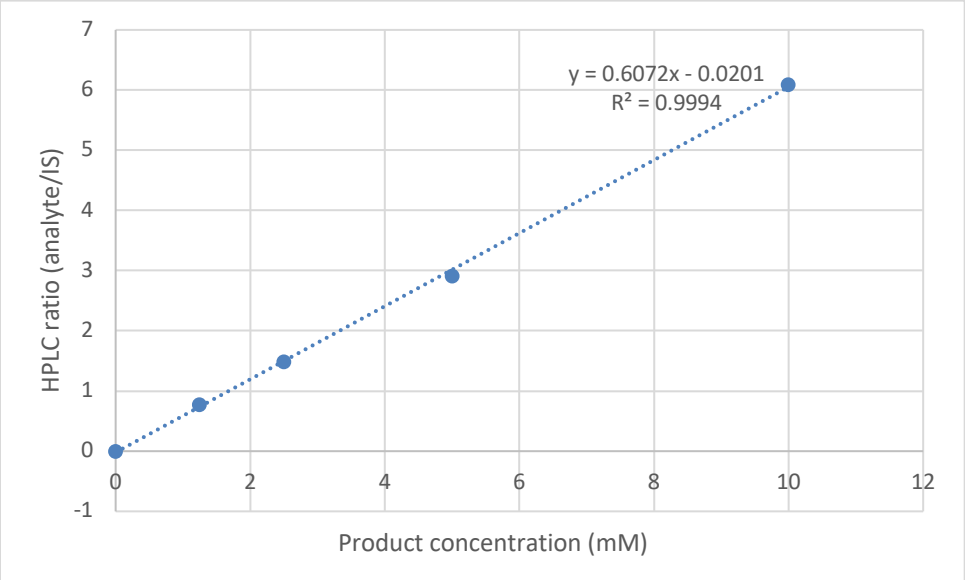
Variant	Pdt	IS	Pdt/IS	[Pdt] (mM)	yield%	[PC] (uM)	TTN	Avg yield%	SD yield %	Avg TTN	SD TTN
*Par*Pgb-HYA-5209	trace	122.45	n.d.	n.d.	n.d.	4.32	<5	n.d.	n.d.	<5	n.d.
	trace	119.43	n.d.	n.d.	n.d.	4.32	<5				
	trace	120.15	n.d.	n.d.	n.d.	4.63	<5				
	trace	121.46	n.d.	n.d.	n.d.	4.63	<5				
*Par*Pgb-HYA-5210	1.09	122.13	0.0089	0.0251	0.25	1.54	16	0.27	0.03	15	2.1
	1.58	119.72	0.0132	0.0319	0.32	1.54	21				
	1.21	123.15	0.0098	0.0265	0.26	1.59	17				
	1.05	122.25	0.0086	0.0245	0.24	1.59	15				
	16.66	120.30	0.1385	0.2324	2.32	3.26	71	2.34	0.03	70	1.0

*Par*Pgb-HYA-5211	17.19	121.43	0.1416	0.2374	2.37	3.26	73				
	16.07	116.83	0.1375	0.2309	2.31	3.30	70				
	16.86	119.42	0.1412	0.2368	2.37	3.30	72				
*Par*Pgb-HYA-5212	34.15	118.76	0.2875	0.4710	4.71	2.64	180				
	37.19	121.15	0.3070	0.5023	5.02	2.64	190				
	35.11	121.26	0.2895	0.4743	4.74	2.56	190				
	35.81	121.00	0.2959	0.4845	4.85	2.56	190	4.83	0.12	180	4.7
*Par*Pgb-HYA-5213	91.80	120.99	0.7587	1.2254	12.25	3.69	330				
	91.63	117.67	0.7787	1.2575	12.57	3.69	340				
	89.74	117.32	0.7649	1.2353	12.35	3.91	320				
	90.95	119.16	0.7633	1.2328	12.33	3.91	320	12.4	0.12	320	11

Notes: Pdt = product area, IS = internal standard area, [Pdt] = product concentration in reaction, [PC] = protein concentration in reaction, Avg yield= average of percentage yield, SD yield = standard deviation of percentage yield, Avg TTN = average of total turnover number, SD TTN = standard deviation of TTN. The yield of *Par*Pgb-HYA-5209 was too low to be accurately determined. n.d. = not determined.

C.3.3.4. Yields and TTNs for the 4-Vinylanisole (4a) Reaction across the Lineage (Figure 4-2a)

Standard Curve based on UV absorbance at 280 nm for product **5a** and 1,3,5-trimethoxybenzene as internal standard.



Variant	Pdt	IS	Pdt/IS	[Pdt] (mM)	yield%	[PC] (uM)	TTN	Avg yield%	SD yield%	Avg TTN	SD TTN
ParPgb-HYA-5209	4.48	92.72	0.0483	0.1127	1.13	4.32	26	1.09	0.03	25	1.6
	4.06	85.42	0.0475	0.1114	1.11	4.32	26				
	4.38	95.76	0.0458	0.1085	1.08	4.63	23				
	4.10	95.52	0.0430	0.1038	1.04	4.63	22				
ParPgb-HYA-5210	8.57	86.94	0.0985	0.1954	1.95	1.54	130	1.65	0.18	100	13
	6.73	88.66	0.0759	0.1581	1.58	1.54	103				
	7.15	93.52	0.0765	0.1590	1.59	1.59	100				
	6.46	93.17	0.0694	0.1473	1.47	1.59	90				
ParPgb-HYA-5211	24.51	94.74	0.2587	0.4592	4.59	3.26	140	4.93	0.25	150	7.4
	27.77	94.49	0.2939	0.5171	5.17	3.26	160				
	25.77	94.99	0.2712	0.4798	4.80	3.30	150				

	27.62	94.30	0.2929	0.5155	5.15	3.30	160				
ParPgb-HYA-5212	39.36	92.13	0.4272	0.7367	7.37	2.64	280	7.34	0.13	280	7.4
	39.36	93.43	0.4213	0.7269	7.27	2.64	280				
	38.45	92.26	0.4167	0.7194	7.19	2.56	280				
	41.16	93.90	0.4383	0.7549	7.55	2.56	300				
ParPgb-HYA-5213	62.94	94.12	0.6687	1.1345	11.34	3.69	310	11.3	0.14	300	8.1
	61.39	93.16	0.6589	1.1183	11.18	3.69	300				
	62.36	94.66	0.6587	1.1179	11.18	3.91	290				
	64.45	94.77	0.6801	1.1532	11.53	3.91	300				

Notes: Pdt = product area, IS = internal standard area, [Pdt] = product concentration in reaction, [PC] = protein concentration in reaction, Avg yield = average of percentage yield, SD yield = standard deviation of percentage yield, Avg TTN = average of total turnover number, SD TTN = standard deviation of TTN.

C.4. Product Quantitation and TTNs for the Substrate Scope.

C.4.1. Analytical-Scale Reaction Setup (Purified Enzyme)

KPi buffer (pH 8.0) was placed in the anaerobic chamber with an oxygen concentration below 20 ppm for at least 24 hours. Purified enzyme (see section 1.5) was diluted with KPi buffer (pH 8.0) to make the total volume to 340 µL of [Enzyme] = 59 µM and aliquoted to 2-mL screw cap vials in a Coy anaerobic chamber. The hydrocarbon substrate (20 µL, 200 mM stock in ethanol), fresh sodium dithionite (20 µL, 200 mM stock in water) and hydroxylammonium chloride (20 µL, 600 mM stock in water) were added in a sequential manner. The reactions were then shaken at room temperature for 12 hours at 200 rpm.

To 1.7-mL Eppendorf microcentrifuge tubes, 400 µL ethanol (10 mM 1,3,5-trimethoxybenzene as internal standard) were added, and 400 µL of the mixtures were subsequently transferred. After vortexing the mixtures, the microcentrifuge tubes were centrifuged (14,000*g*, 10 minutes, 4 °C) to precipitate coagulated proteins. To identify the products and determine yields, 200 µL of the supernatants were transferred to a 2-mL screw cap vial with an insert and analyzed by reverse-phase HPLC-MS. Product formation was quantified by HPLC-MS based on the calibration curve of the corresponding racemic reference compound.

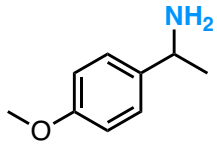
C.4.2. Calibration Curves for Standard Products and Reaction Product Quantitation

Enzymatic reactions on analytic scale were performed following the general procedure described in Section 5.1. Product formation was quantified by HPLC-MS based on the

calibration curve of the corresponding racemic reference compound. TTN is defined as the concentration of product divided by the concentration of heme protein (50 μM for all analytical scale reactions). Calibration curves of commercially available or synthesized reference compounds were created for the determination of yield and TTN. To prepare the calibration curve samples, 360 μL of enzyme solution ([Enzyme] = 55 μM) were aliquoted to 2-mL screw cap vials. Product standard stock (20 μL; 200 mM, 100 mM, 50 mM, and 25 mM stock in ethanol) was added to the mixture. The mixture was shaken overnight at room temperature, quenched with 400 μL of ethanol (10 mM 1,3,5-trimethoxybenzene as internal standard), vortexed, and then analyzed by HPLC. The integrated peak area of the UV absorbance at selected wavelengths of the products and the integrated peak area of indicated internal standard were measured. Methods were adjusted so that none of the products eluted at the same time as any of the cell impurities. For compounds **3a**–**3e**, solvent A: water containing 0.1% acetic acid and solvent B: acetonitrile containing 0.1% acetic acid were used as eluents for a C18 (Agilent Poroshell 120 EC-C18, 4.6 × 50 mm, 5 μm) column. The methods used 5–95% solvent B (0.0–3.0 minutes), 95% solvent B (3.0–4.0 minutes), and 95–5% (4.0–4.5 minutes) with 5-μL sample injections. The flow rate was 1.5 mL/minute, and the column was maintained at 37 °C. For compounds **5a**–**5e**, solvent A: 50 mM disodium phosphate buffer (pH 7.0) and solvent B: 25 mM disodium phosphate in 50% acetonitrile/water were used as eluents for a C18 (Agilent Poroshell 120 EC-C18, 4.6 × 100 mm, 5 μm) column. The methods used 0–80% solvent B (0.0–5.0 minutes), 80-95% solvent B (5.0–5.5 minutes), 95% solvent B (5.5–6.5 minutes) and 95%–0% solvent

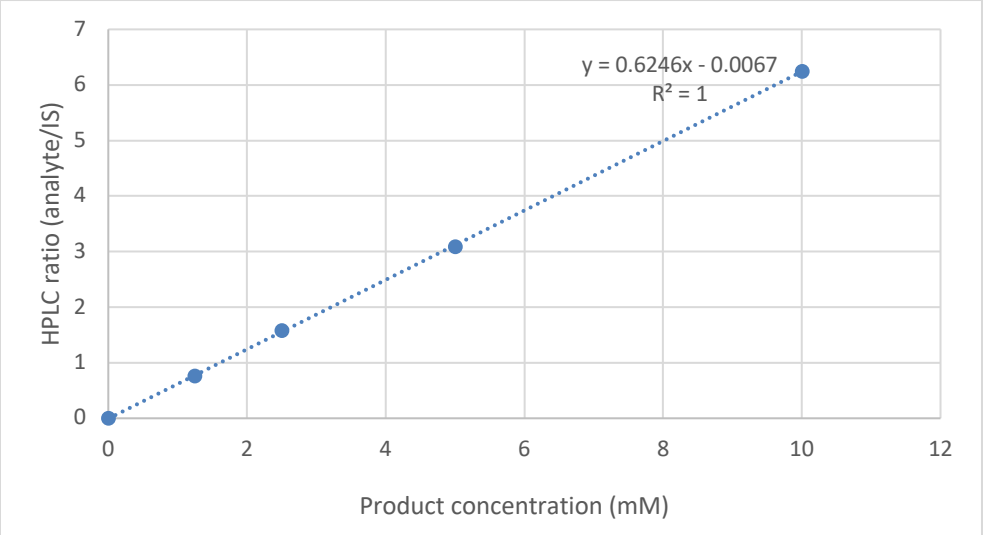
B (6.5–7.0 minutes) with 5-µL sample injections. The flow rate was 2.0 mL/minute, and the column was maintained at 22 °C.

C.4.3 Yields and TTNs for the Substrate Scope (Figure 4-3a,b)



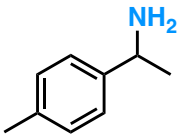
*1-(4-Methoxyphenyl)ethan-1-amine (**3a**)*

Standard curve based on UV absorbance at 280 nm for product **3a** and 1,3,5-trimethoxybenzene as internal standard.



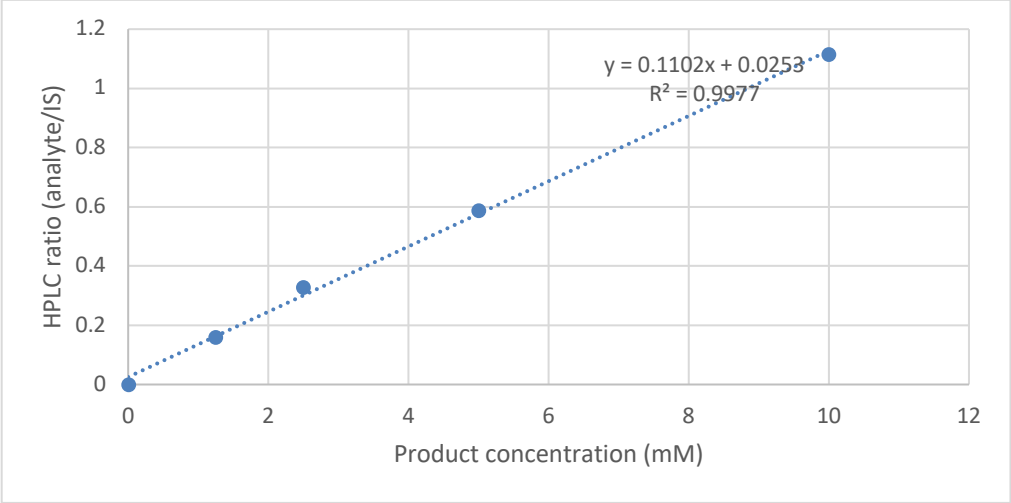
Pdt	IS	Pdt/IS	[Pdt] (mM)	yield %	TTN	Average yield	SD yield	Average TTN	SD TTN
717.86	119.63	6.001	9.62	96.2	190				
711.46	120.12	5.923	9.49	94.9	190				
714.82	120.57	5.929	9.50	95.0	190				
711.93	119.83	5.941	9.52	95.2	190	95.3	0.49	190	0.99

Notes: Pdt = product area, IS = internal standard area, [Pdt] = product concentration in reaction, [PC] = protein concentration in reaction, Avg yield = average of percentage yield, SD yield = standard deviation of percentage yield, Avg TTN = average of total turnover number, SD TTN = standard deviation of TTN.



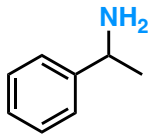
*1-(p-Tolyl)ethan-1-amine (**3b**)*

Standard curve based on UV absorbance at 210 nm for product **3b** and 1,3,5-trimethoxybenzene as internal standard.



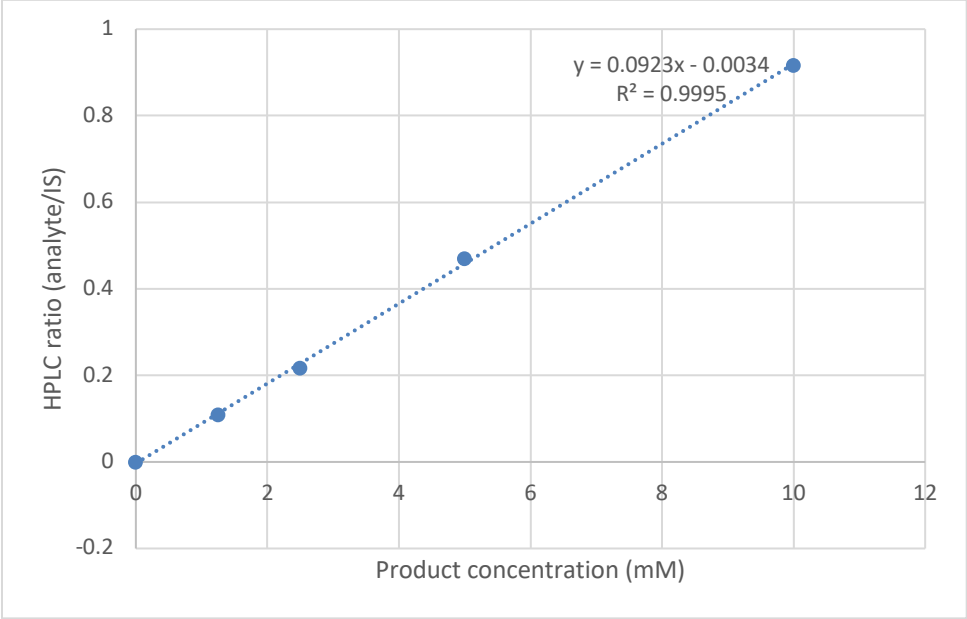
Pdt	IS	Pdt/IS	[Pdt] (mM)	yield %	TTN	Average yield	SD yield	Average TTN	SD TTN
3040.45	6344.53	0.4792	4.12	41.1	82				
3025.13	6421.63	0.4711	4.05	40.4	81				
2877.76	6340.01	0.4539	3.89	38.9	78				
2833.63	6315.33	0.4487	3.84	38.4	77	39.7	1.1	79	2.3

Notes: Pdt = product area, IS = internal standard area, [Pdt] = product concentration in reaction, [PC] = protein concentration in reaction, Avg yield = average of percentage yield, SD yield = standard deviation of percentage yield, Avg TTN = average of total turnover number, SD TTN = standard deviation of TTN.



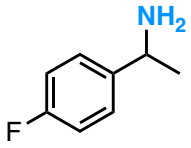
*1-Phenylethan-1-amine (**3c**)*

Standard curve based on UV absorbance at 210 nm for product **3c** and 1,3,5-trimethoxybenzene as internal standard.



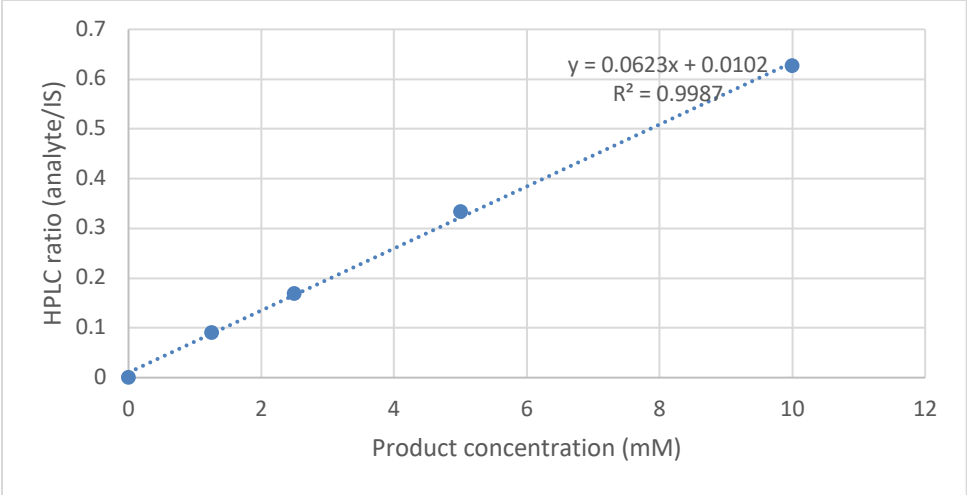
Pdt	IS	Pdt/IS	[Pdt] (mM)	yield %	TTN	Average yield	SD yield	Average TTN	SD TTN
3724.13	7707.51	0.4832	5.27	52.7	110				
3609.22	7621.44	0.4736	5.17	51.7	100				
3689.88	7378.33	0.5001	5.45	54.6	110				
3666.53	7632.72	0.4804	5.24	52.4	110	52.8	1.06	100	2.1

Notes: Pdt = product area, IS = internal standard area, [Pdt] = product concentration in reaction, [PC] = protein concentration in reaction, Avg yield = average of percentage yield, SD yield = standard deviation of percentage yield, Avg TTN = average of total turnover number, SD TTN = standard deviation of TTN.



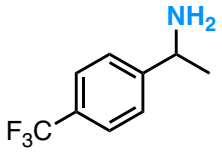
*1-(4-Fluorophenyl)ethan-1-amine (**3d**)*

Standard curve based on UV absorbance at 210 nm for product **3d** and 1,3,5-trimethoxybenzene as internal standard.



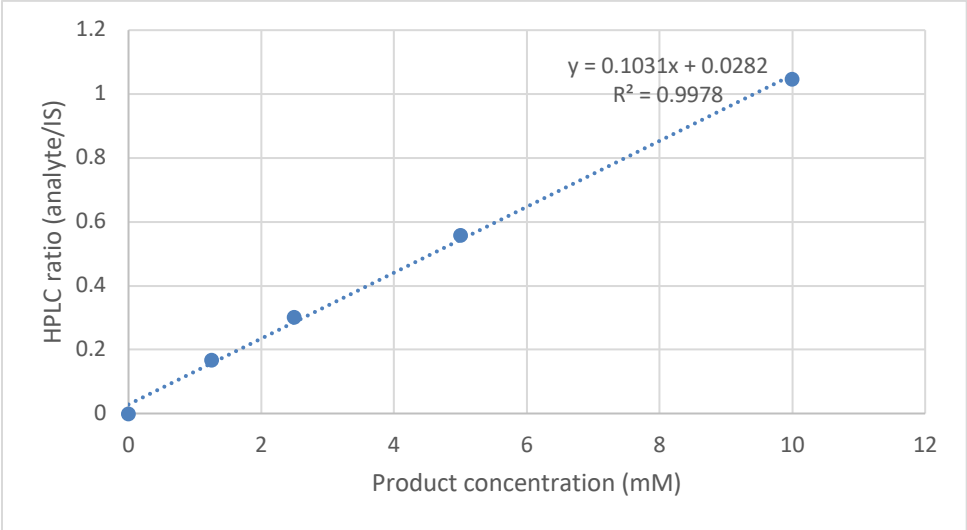
Pdt	IS	Pdt/IS	[Pdt] (mM)	yield %	TTN	Average yield	SD yield	Average TTN	SD TTN
3502.26	7946.71	0.4407	6.91	69.1	140				
3489.44	8093.24	0.4312	6.76	67.6	140				
3489.54	7933.97	0.4398	6.90	69.0	140				
3561.17	7989.39	0.4457	6.99	69.9	140	68.9	0.84	140	1.7

Notes: Pdt = product area, IS = internal standard area, [Pdt] = product concentration in reaction, [PC] = protein concentration in reaction, Avg yield = average of percentage yield, SD yield = standard deviation of percentage yield, Avg TTN = average of total turnover number, SD TTN = standard deviation of TTN.



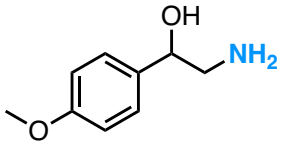
*1-(4-(Trifluoromethyl)phenyl)ethan-1-amine (**3e**)*

Standard curve based on UV absorbance at 210 nm for product **3e** and 1,3,5-trimethoxybenzene as internal standard.



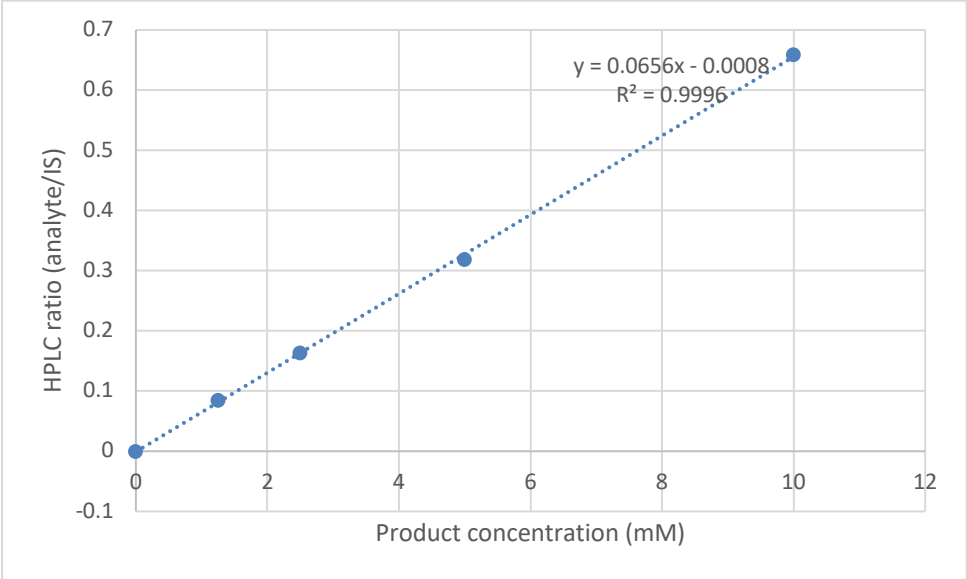
Pdt	IS	Pdt/IS	[Pdt] (mM)	yield %	TTN	Average yield	SD yield	Average TTN	SD TTN
1170.33	7243.17	0.1616	1.29	12.9	26				
1147.29	7344.25	0.1562	1.24	12.4	25				
862.70	6966.50	0.1238	0.93	9.28	19				
850.50	6979.73	0.1219	0.91	9.08	18	10.9	1.76	22	3.5

Notes: Pdt = product area, IS = internal standard area, [Pdt] = product concentration in reaction, [PC] = protein concentration in reaction, Avg yield = average of percentage yield, SD yield = standard deviation of percentage yield, Avg TTN = average of total turnover number, SD TTN = standard deviation of TTN.



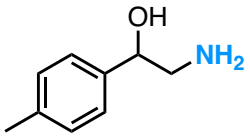
*2-Amino-1-(4-methoxyphenyl)ethan-1-ol (**5a**)*

Standard curve based on UV absorbance at 280 nm for product **5a** and 1,3,5-trimethoxybenzene as internal standard.



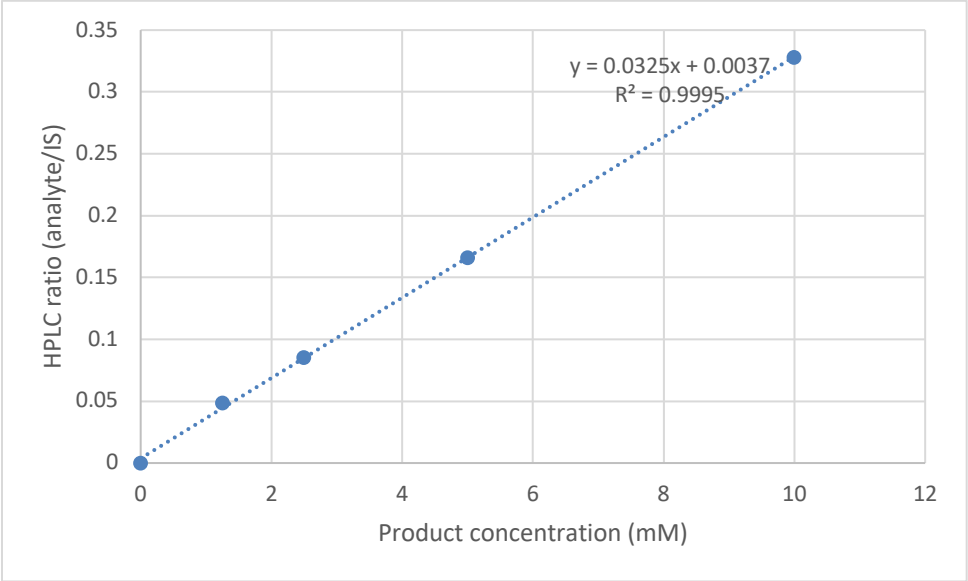
Pdt	IS	Pdt/IS	[Pdt] (mM)	yield %	TTN	Average yield	SD yield	Average TTN	SD TTN
3049.16	5412.31	0.5634	8.60	86.0	170				
3055.70	5409.49	0.5649	8.62	86.2	170				
3089.26	5422.14	0.5697	8.70	87.0	170				
3066.93	5426.01	0.5652	8.63	86.3	170	86.4	0.36	170	0.73

Notes: Pdt = product area, IS = internal standard area, [Pdt] = product concentration in reaction, [PC] = protein concentration in reaction, Avg yield = average of percentage yield, SD yield = standard deviation of percentage yield, Avg TTN = average of total turnover number, SD TTN = standard deviation of TTN.



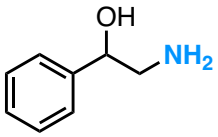
*2-Amino-1-(p-tolyl)ethan-1-ol (**5b**)*

Standard curve based on UV absorbance at 210 nm for product **5b** and 1,3,5-trimethoxybenzene as internal standard.



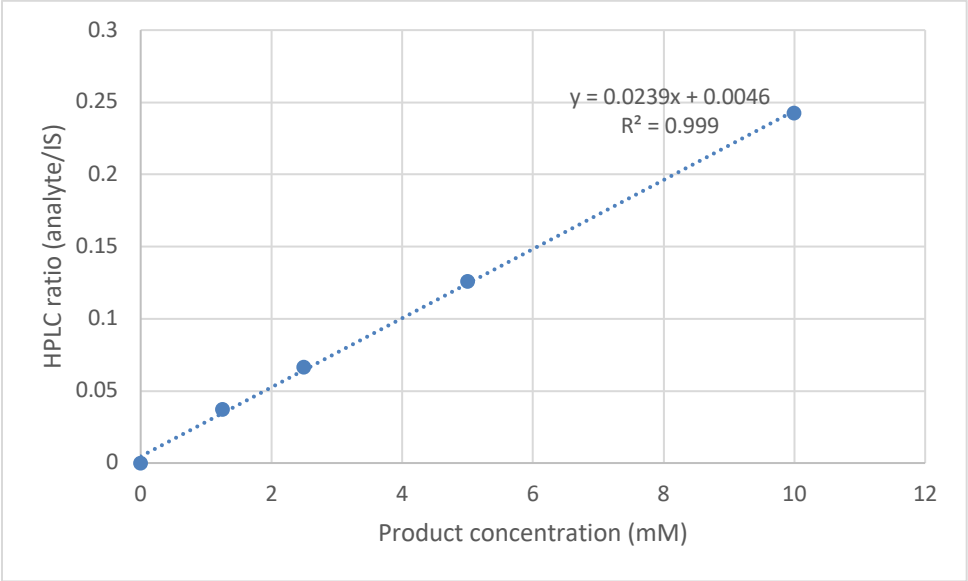
Pdt	IS	Pdt/IS	[Pdt] (mM)	yield %	TTN	Average yield	SD yield	Average TTN	SD TTN
2054.75	15671.27	0.1311	3.92	39.2	78				
2050.15	15501.23	0.1323	3.96	39.6	79				
2059.46	15416.50	0.1336	4.00	40.0	80				
2089.11	15348.85	0.1361	4.07	40.7	81	39.9	0.57	80	1.1

Notes: Pdt = product area, IS = internal standard area, [Pdt] = product concentration in reaction, [PC] = protein concentration in reaction, Avg yield = average of percentage yield, SD yield = standard deviation of percentage yield, Avg TTN = average of total turnover number, SD TTN = standard deviation of TTN.



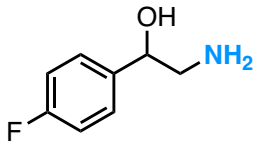
*2-Amino-1-phenylethan-1-ol (**5c**)*

Standard curve based on UV absorbance at 210 nm for product **5c** and 1,3,5-trimethoxybenzene as internal standard.



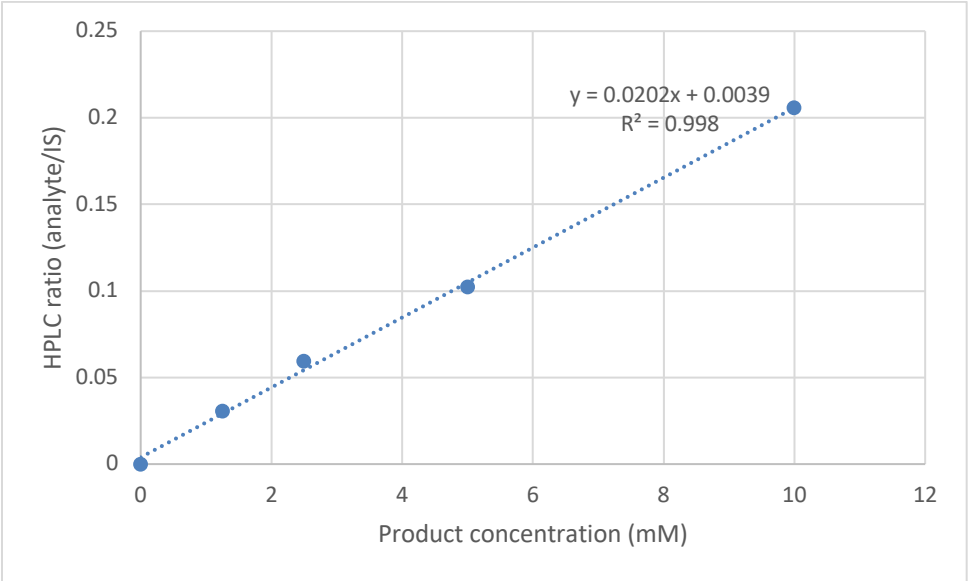
Pdt	IS	Pdt/IS	[Pdt] (mM)	yield %	TTN	Average yield	SD yield	Average TTN	SD TTN
870.04	15484.83	0.05619	2.16	21.6	43				
868.20	15466.38	0.05613	2.16	21.6	43				
860.61	15410.30	0.05585	2.14	21.4	43				
866.57	15374.60	0.05636	2.17	21.7	43	21.6	0.08	43	0.16

Notes: Pdt = product area, IS = internal standard area, [Pdt] = product concentration in reaction, [PC] = protein concentration in reaction, Avg yield = average of percentage yield, SD yield = standard deviation of percentage yield, Avg TTN = average of total turnover number, SD TTN = standard deviation of TTN.



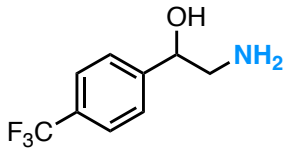
*2-Amino-1-(4-fluorophenyl)ethan-1-ol (**5d**)*

Standard curve based on UV absorbance at 210 nm for product **5d** and 1,3,5-trimethoxybenzene as internal standard.



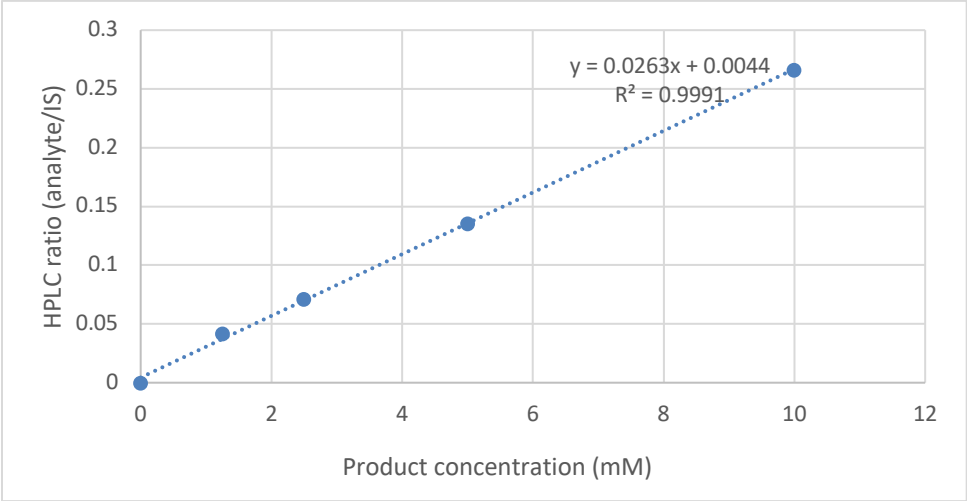
Pdt	IS	Pdt/IS	[Pdt] (mM)	yield %	TTN	Average yield	SD yield	Average TTN	SD TTN
1008.86	15512.63	0.06503	3.03	30.3	61				
1027.51	15308.27	0.06712	3.13	31.3	63				
1015.79	15407.34	0.06593	3.07	30.7	61				
1009.88	15501.20	0.06515	3.03	30.3	61	30.7	0.41	61	0.82

Notes: Pdt = product area, IS = internal standard area, [Pdt] = product concentration in reaction, [PC] = protein concentration in reaction, Avg yield = average of percentage yield, SD yield = standard deviation of percentage yield, Avg TTN = average of total turnover number, SD TTN = standard deviation of TTN.



*2-Amino-1-(4-(trifluoromethyl)phenyl)ethan-1-ol (**5e**)*

Standard curve based on UV absorbance at 210 nm for product **5e** and 1,3,5-trimethoxybenzene as internal standard.



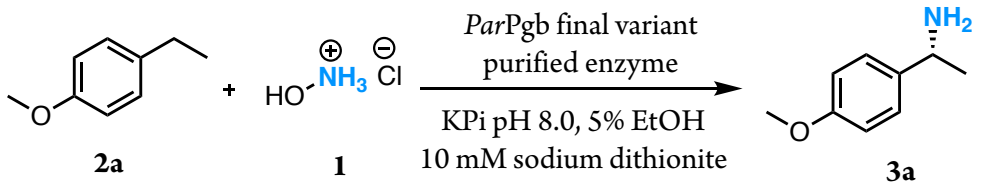
Pdt	IS	Pdt/IS	[Pdt] (mM)	yield %	TTN	Average yield	SD yield	Average TTN	SD TTN
440.59	15301.18	0.02879	0.93	9.3	19				
440.00	15423.48	0.02853	0.92	9.2	18				
436.52	15497.09	0.02817	0.90	9.0	18				
437.75	15311.42	0.02859	0.92	9.2	18	9.2	0.09	18	0.17

Notes: Pdt = product area, IS = internal standard area, [Pdt] = product concentration in reaction, [PC] = protein concentration in reaction, Avg yield = average of percentage yield, SD yield = standard deviation of percentage yield, Avg TTN = average of total turnover number, SD TTN = standard deviation of TTN.

C.4.4. Preparative-Scale Enzymatic Reaction

To a 40-mL drum vial, a solution of purified enzyme (final variant ParPgb-HYA-5213, 17 mL, 59 μM), hydrocarbon substrate (0.2 mmol, 1 mL, 200 mM in ethanol), fresh sodium dithionite (0.2 mmol, 1 mL, 200 mM in water) and hydroxylammonium chloride (0.6 mmol, 1 mL, 600 mM in water) were added sequentially under anaerobic conditions. The mixture was shaken at 100 rpm in a shaker inside the anaerobic chamber for 20 h.

The product was further purified via an acid-base workup.[7] After the reaction was completed, 40 mL ethanol were added and mixed vigorously by shaking. The debris was removed by centrifugation (5500 g, 10 minutes, 4 °C). Then ethanol and most of the water in the supernatant were removed by rotary evaporation until a final volume of about 20 mL was achieved, and the remaining aqueous layer was treated with HCl solution (1 M) to pH < 2 in an ice bath. After washing two times with diethyl ether (5 mL each time), the aqueous phase was treated with NaOH solution (1 M) to pH > 12 on an ice bath and extracted with ethyl acetate for six times (10 mL each time). During the extraction, centrifugation was used for phase separation (5,000 g, 5 minutes, 4 °C). The combined organic layer was washed with brine (10 mL), dried by Na_2SO_4, and concentrated *in vacuo* to give the final products without further purification.



2a **1** **3a**

(*R*)-1-(4-methoxyphenyl)ethan-1-amine

3a: 21 mg, 69% yield, >99% *ee*. ^{1}H NMR (400 MHz, CDCl3) δ 7.29 – 7.27 (m, 2H), 7.88 – 7.86 (m, 2H), 4.14 – 4.09 (m, 1H), 3.80 (s, 3H), 1.40 (d, J = 8.0 Hz, 3H); ^{13}C NMR (101 MHz, CDCl3) δ 158.55, 140.03, 126.85, 113.92, 55.40, 50.80, 25.84.

C.5. Determination of Enantioselectivity.

C.5.1. Sample Preparation

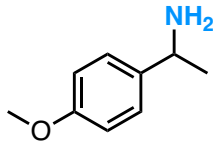
For derivatization by benzoyl chloride:

To a 2-mL screw cap vial, the solution of the amine product (400 μL, enzymatic product or commercial standard from Section 4.1. and Section 4.2.) was added. Sodium bicarbonate (50 μL, saturated in water) and benzoyl chloride (10 μL, neat) were added. The reaction was left to shake overnight and extracted with hexane:ethyl acetate (1:1), centrifuged (14,000*g*, 10 minutes, 4 °C). The organic layer was analyzed by normal-phase HPLC (Abs at 235 nm).

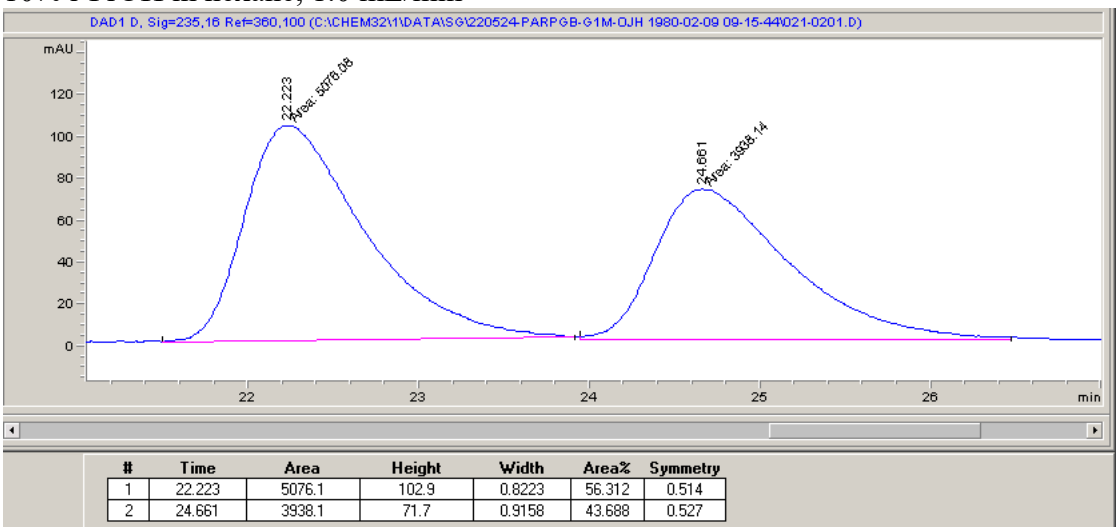
For derivatization by Marfey's reagent:

To a 1.7-mL Eppendorf tube, the solution of the amine product (enzymatic rection or commercial standard from Section 4.1. and Section 4.2., 200 μL of enzymatic reaction quenched with 400 μL of acetonitrile) was added. Sequentially, sodium bicarbonate (100 μL, 40 mM in water) and Marfey's reagent (100 μL, 15 mM in acetonitrile) were added. The reaction was left to shake overnight and quenched with HCl (400 μL, 40 mM in acetonitrile), centrifuged (14,000*g*, 10 minutes, 4 °C), and analyzed by reverse-phase HPLC (absorbance at 340 nm) and LCMS (single-ion channel).

C.5.2. HPLC Traces

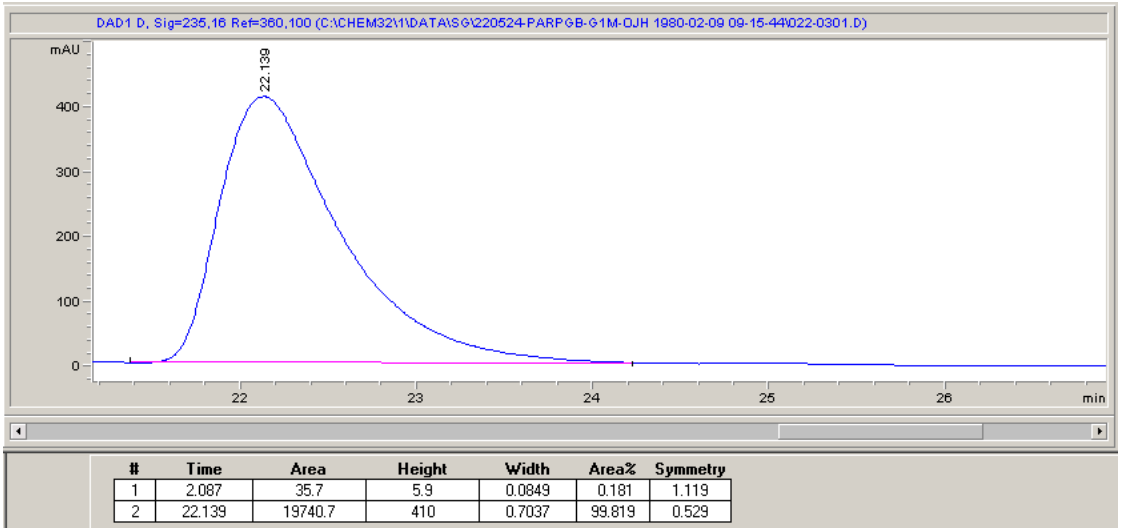


3a-rac – normal phase – racemic standard
Chiral-phase HPLC conditions: Daicel Chiralcel OJ-H (4.6 × 250 mm, 5 µm) column, 10% i-PrOH in hexane, 1.0 mL/min



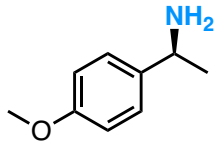
#	Time	Area	Height	Width	Area%	Symmetry
1	22.223	5076.1	102.9	0.8223	56.312	0.514
2	24.661	3938.1	71.7	0.9158	43.688	0.527

3a-enz – normal phase – enzymatic reaction: >99% *ee*
Chiral-phase HPLC conditions: Daicel Chiralcel OJ-H (4.6 × 250 mm, 5 µm) column, 10% i-PrOH in hexane, 1.0 mL/min

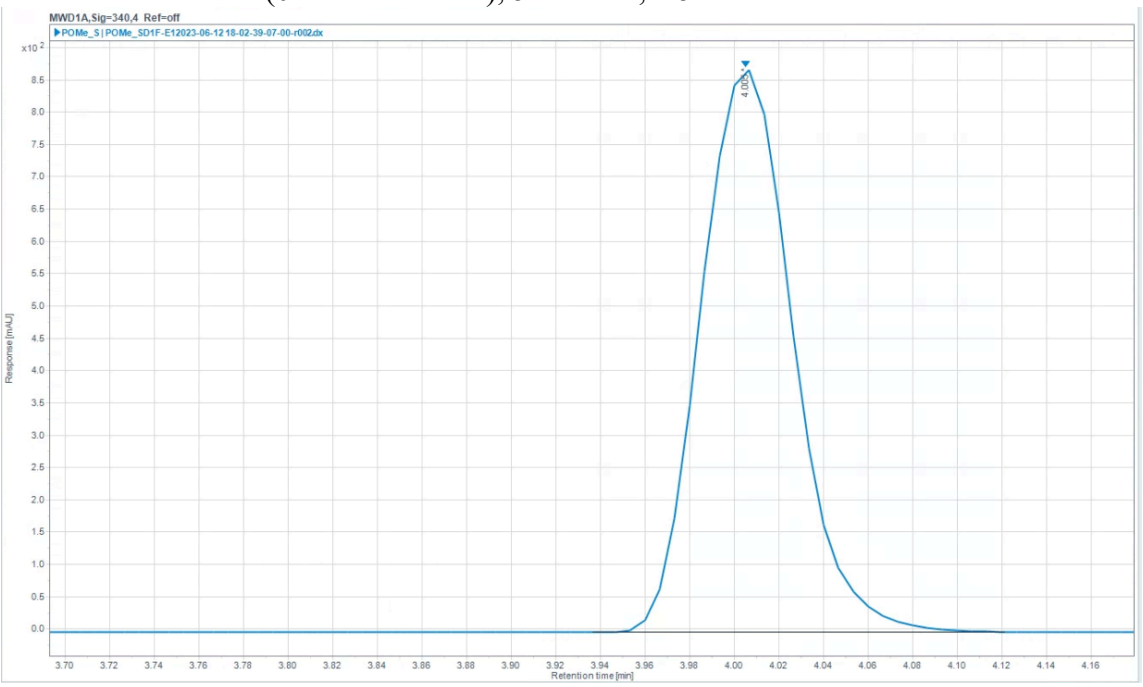


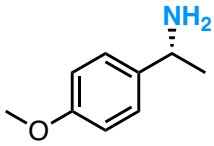
#	Time	Area	Height	Width	Area%	Symmetry
1	2.087	35.7	5.9	0.0849	0.181	1.119
2	22.139	19740.7	410	0.7037	99.819	0.529

The absolute stereochemistry for enzymatic products **3a** were assigned to be R by comparing to their commercial standard:

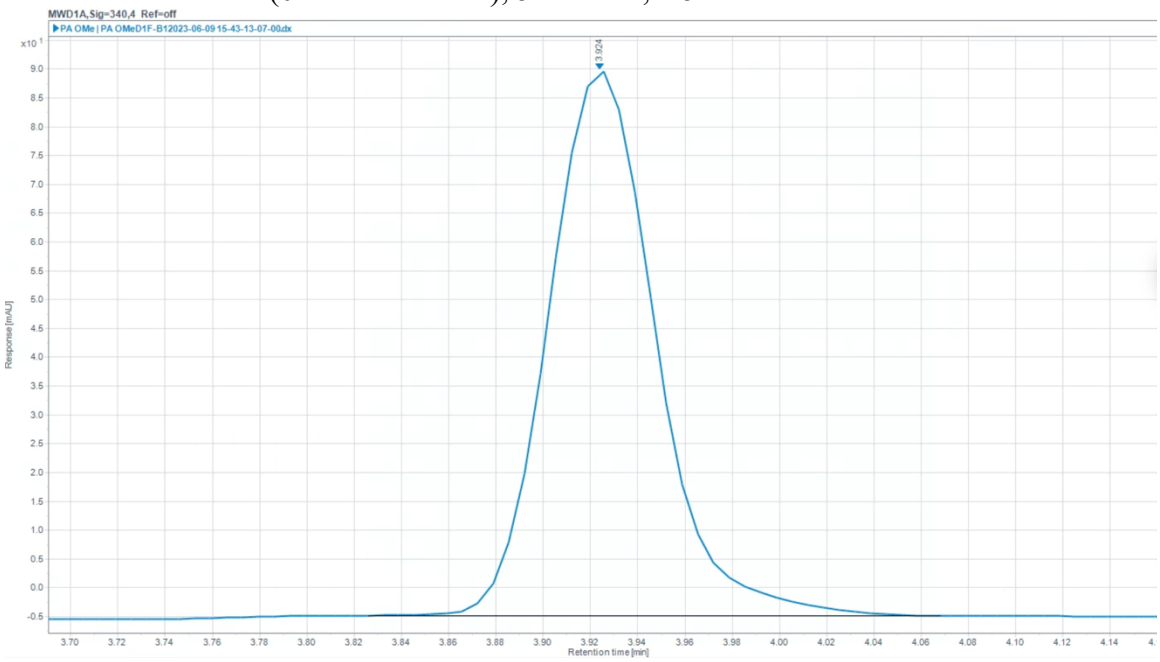


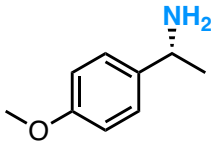
3a-(S) – reverse phase – (S)-standard
HPLC conditions: C18 (Poroshell, EC-C18, 4.6 × 50 mm, 5 µm) column, 10–90% acetonitrile in water (0.1% acetic acid), 5 minutes, 1.5 mL/min



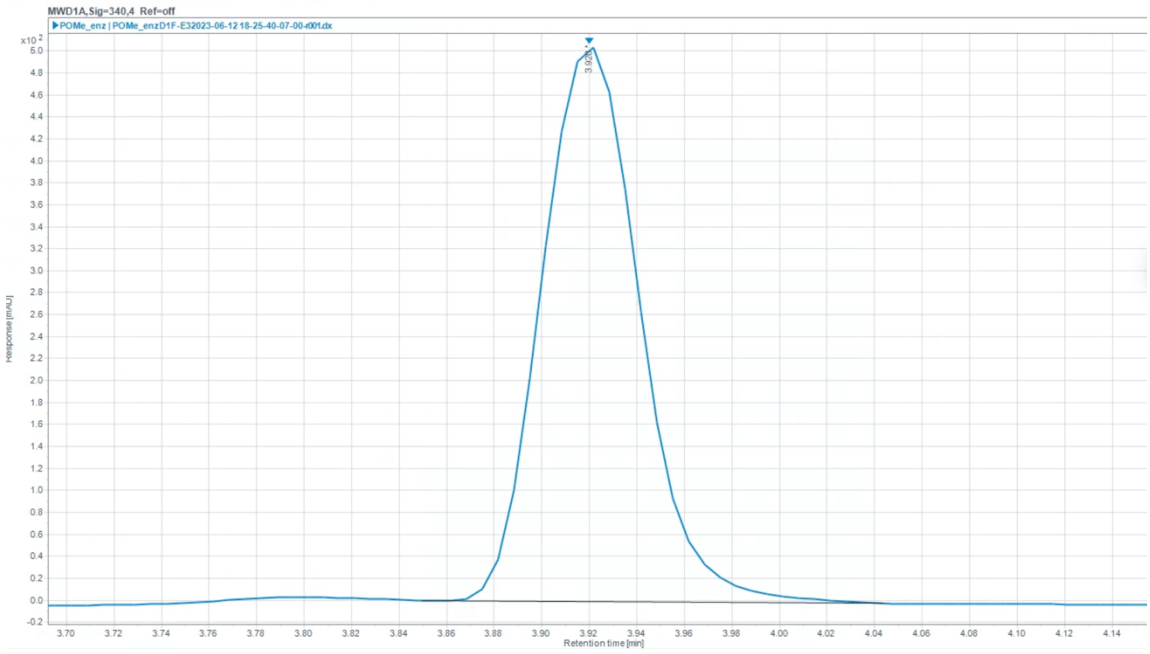


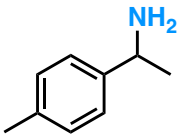
3a-(R) – reverse phase – (R)-standard
HPLC conditions: C18 (Poroshell, EC-C18, 4.6 × 50 mm, 5 µm) column, 10–90% acetonitrile in water (0.1% acetic acid), 5 minutes, 1.5 mL/min



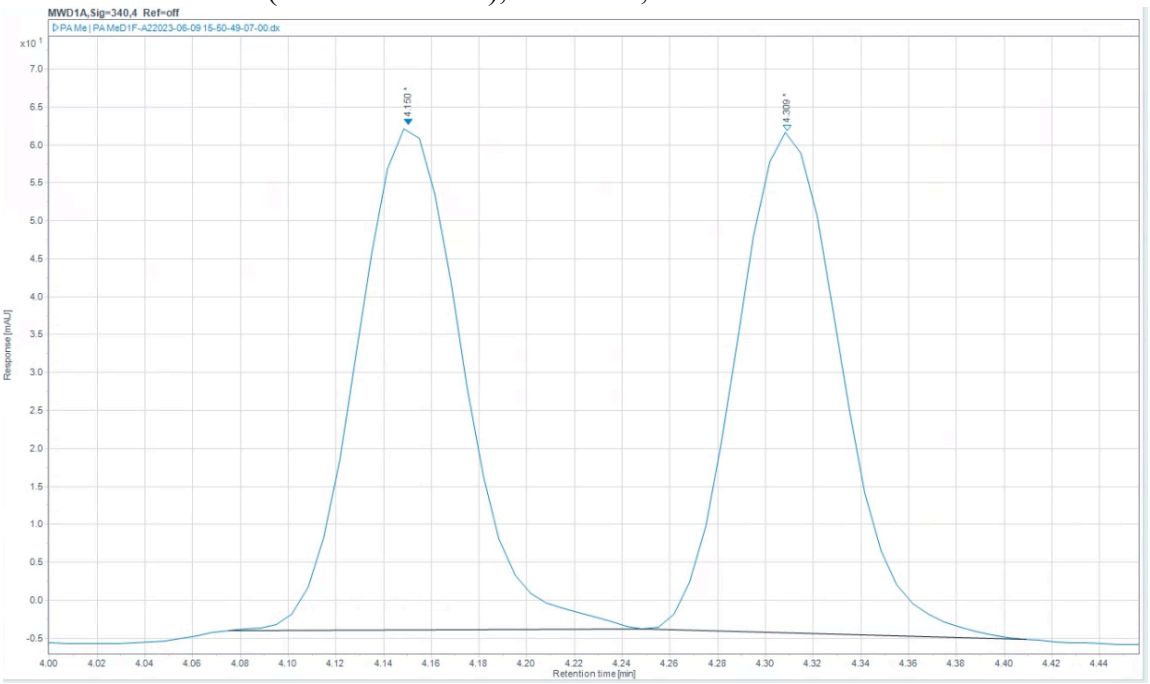


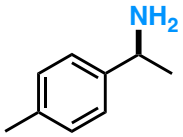
3a-enz – reverse phase – enzymatic reaction
HPLC conditions: C18 (Poroshell, EC-C18, 4.6 × 50 mm, 5 µm) column, 10–90% acetonitrile in water (0.1% acetic acid), 5 minutes, 1.5 mL/min



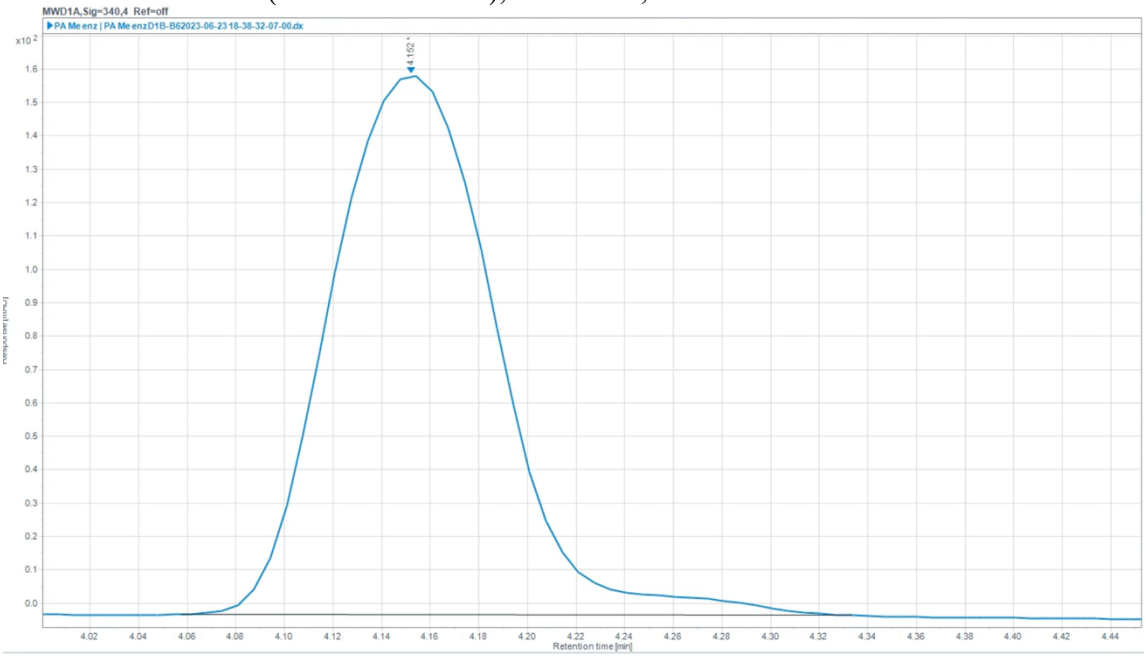


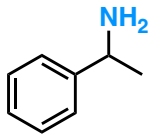
3b-rac – reverse phase – racemic standard
HPLC conditions: C18 (Poroshell, EC-C18, 4.6 × 50 mm, 5 µm) column, 10–90% acetonitrile in water (0.1% acetic acid), 5 minutes, 1.5 mL/min



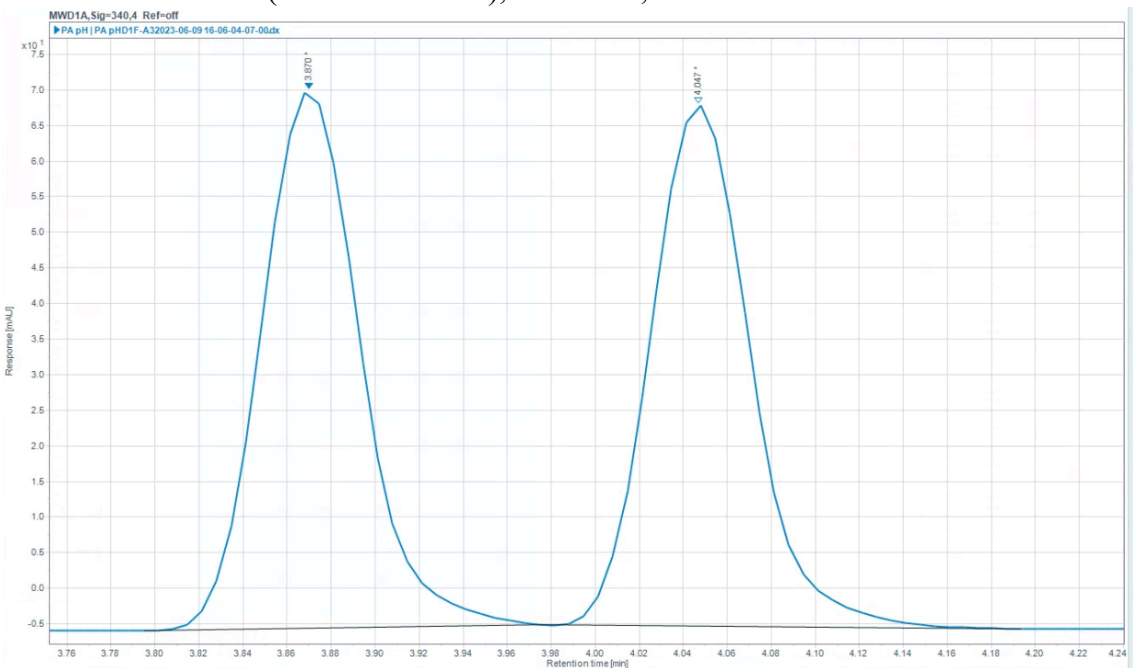


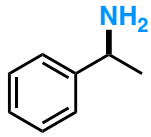
3b-enz – reverse phase – enzymatic reaction: >99% *ee*
HPLC conditions: C18 (Poroshell, EC-C18, 4.6 × 50 mm, 5 μm) column, 10–90% acetonitrile in water (0.1% acetic acid), 5 minutes, 1.5 mL/min



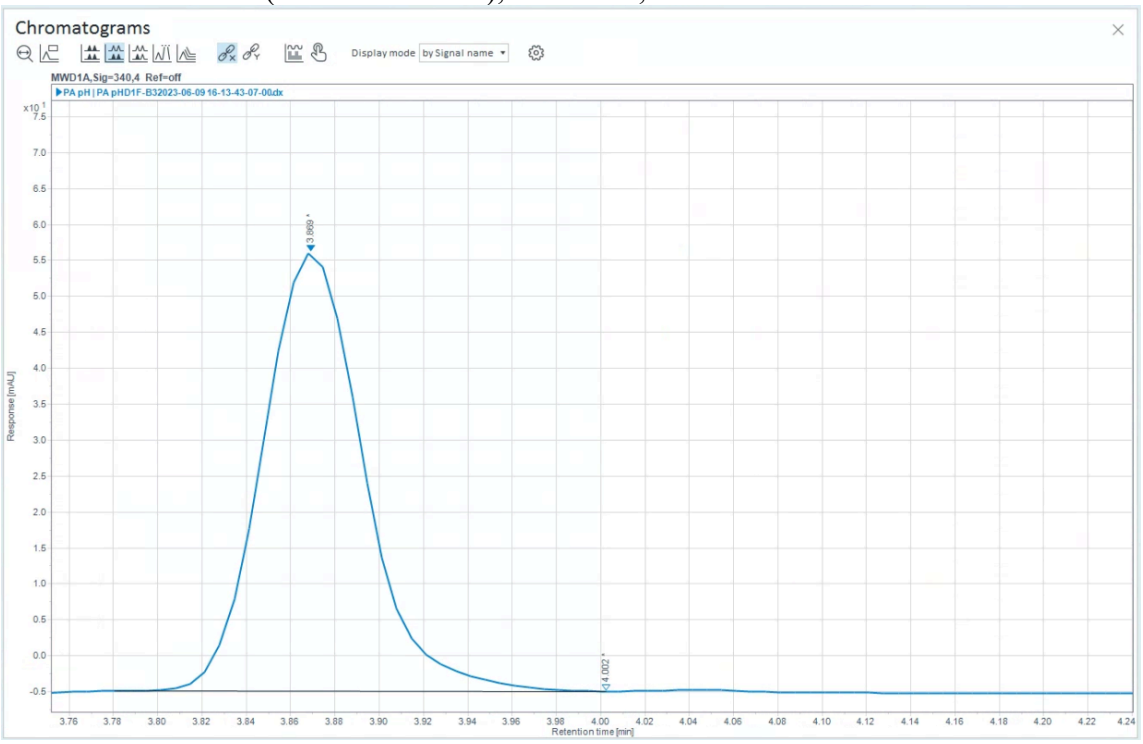


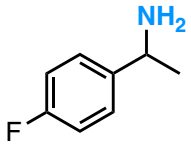
3c-rac – reverse phase – racemic standard
HPLC conditions: C18 (Poroshell, EC-C18, 4.6 × 50 mm, 5 µm) column, 10–90% acetonitrile in water (0.1% acetic acid), 5 minutes, 1.5 mL/min



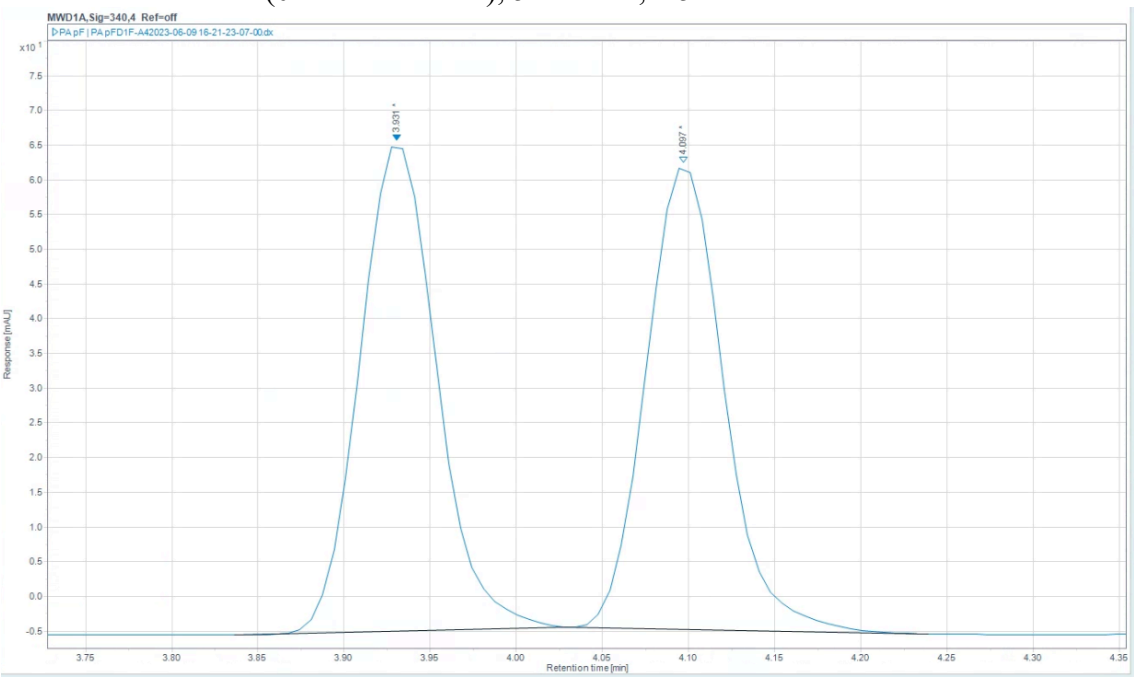


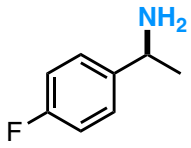
3c-enz – reverse phase – enzymatic reaction: >99% *ee*
HPLC conditions: C18 (Poroshell, EC-C18, 4.6 × 50 mm, 5 µm) column, 10–90% acetonitrile in water (0.1% acetic acid), 5 minutes, 1.5 mL/min



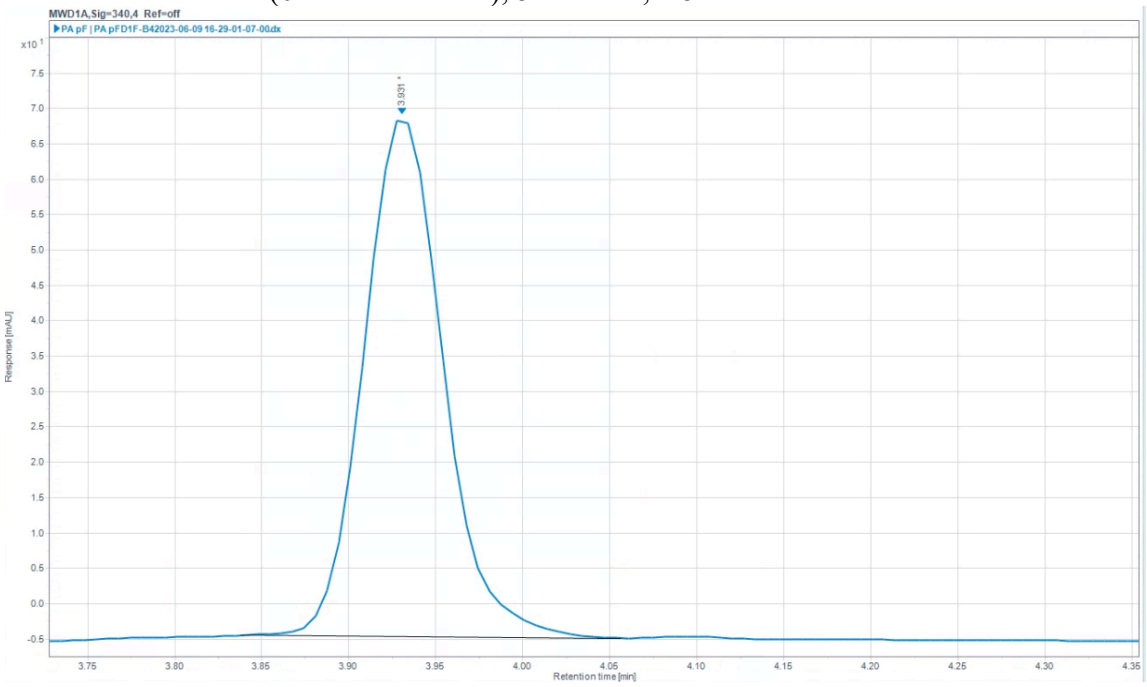


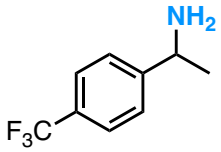
3d-rac – reverse phase – racemic standard
HPLC conditions: C18 (Poroshell, EC-C18, 4.6 × 50 mm, 5 μm) column, 10–90% acetonitrile in water (0.1% acetic acid), 5 minutes, 1.5 mL/min



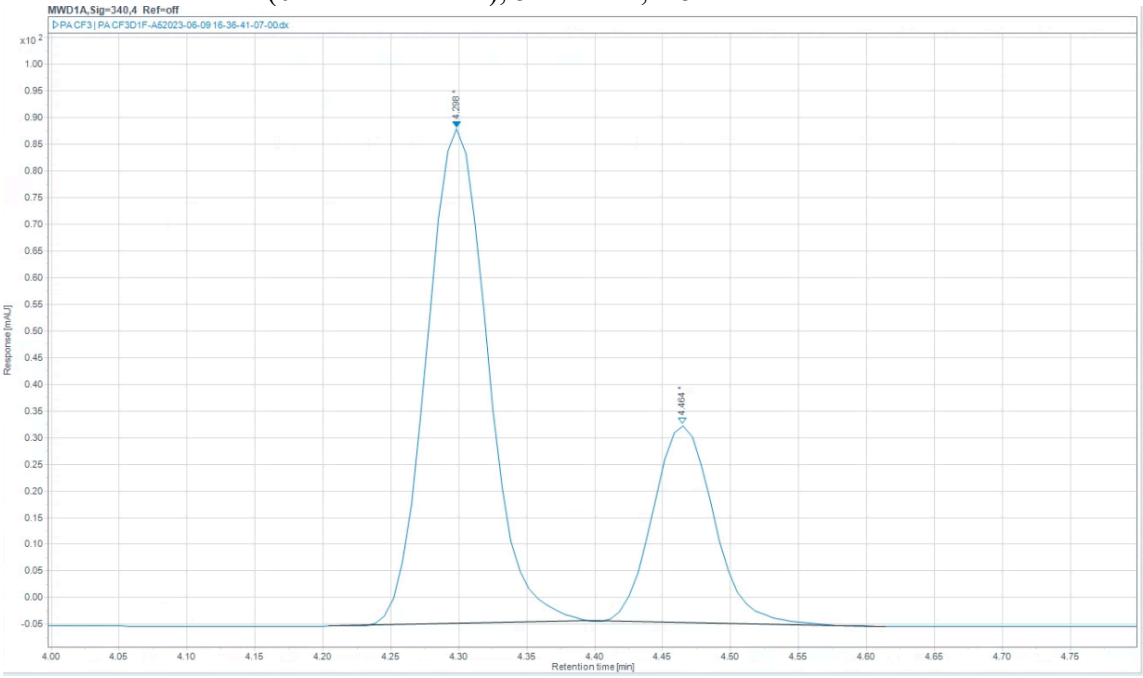


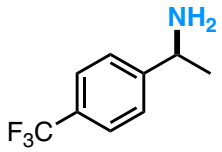
3d-enz – reverse phase – enzymatic reaction: >99% *ee*
HPLC conditions: C18 (Poroshell, EC-C18, 4.6 × 50 mm, 5 µm) column, 10–90% acetonitrile in water (0.1% acetic acid), 5 minutes, 1.5 mL/min



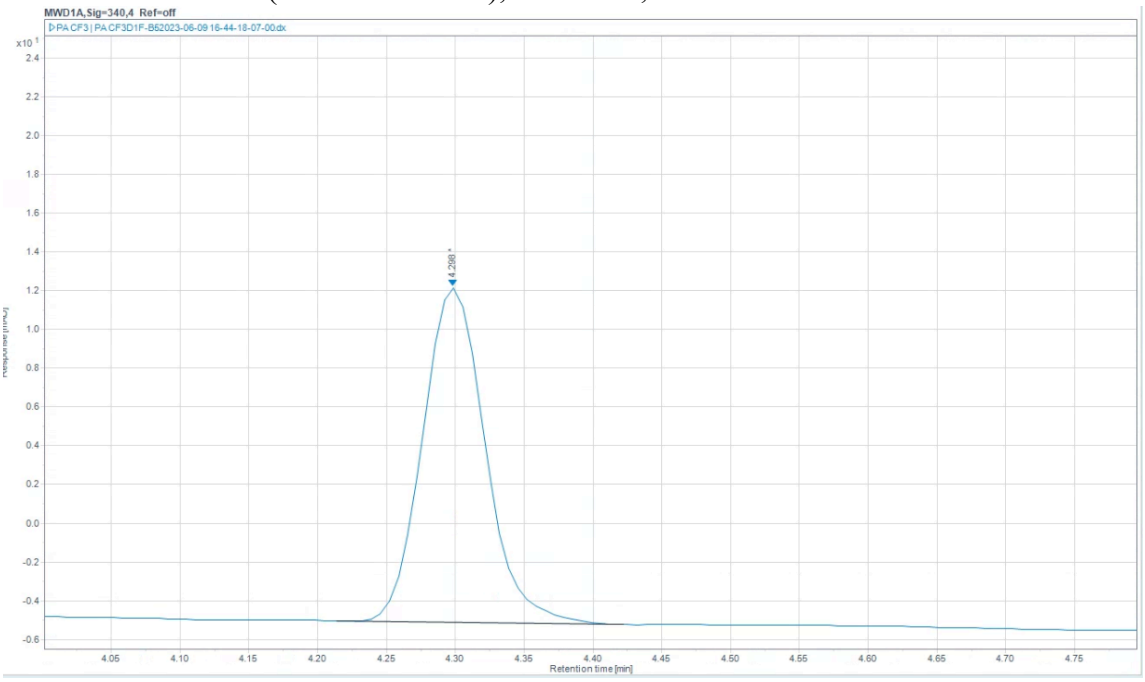


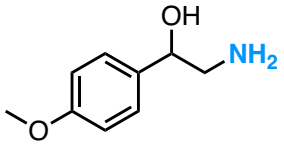
3e-rac – reverse phase – the starting precursor for authentic **3e** was purchased as a scalemic mixture.
HPLC conditions: C18 (Poroshell, EC-C18, 4.6 × 50 mm, 5 µm) column, 10–90% acetonitrile in water (0.1% acetic acid), 5 minutes, 1.5 mL/min



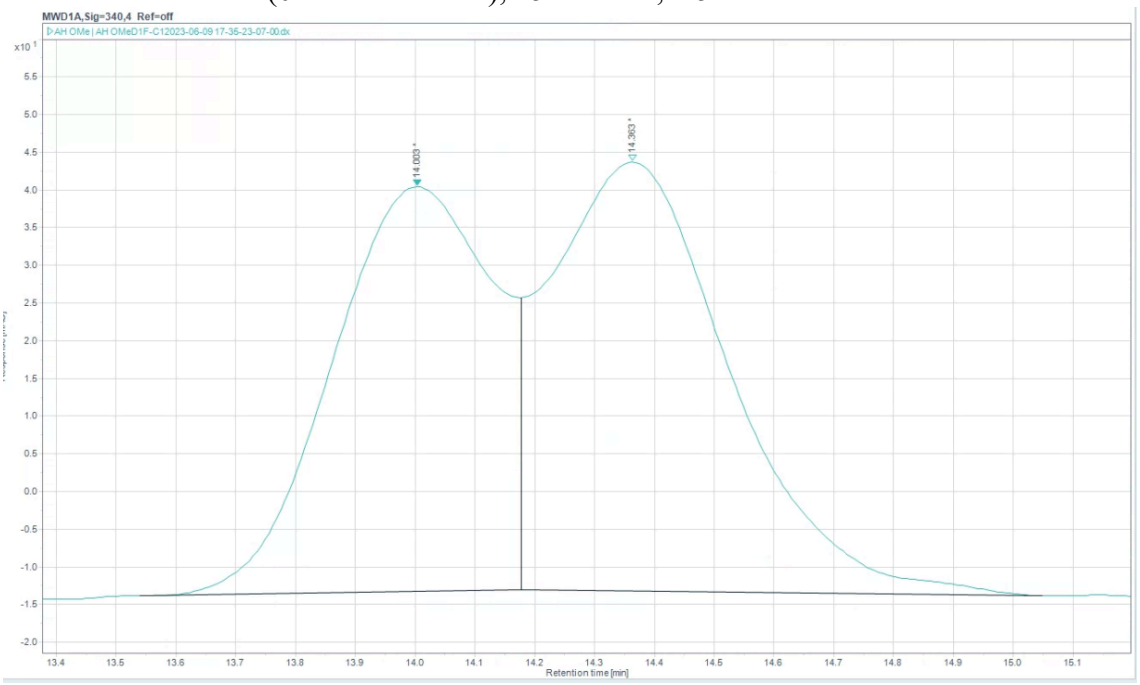


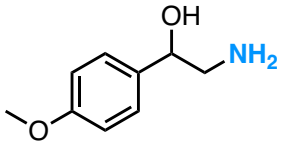
3e-enz – reverse phase – enzymatic reaction: >99% *ee*
HPLC conditions: C18 (Poroshell, EC-C18, 4.6 × 50 mm, 5 μm) column, 10–90% acetonitrile in water (0.1% acetic acid), 5 minutes, 1.5 mL/min



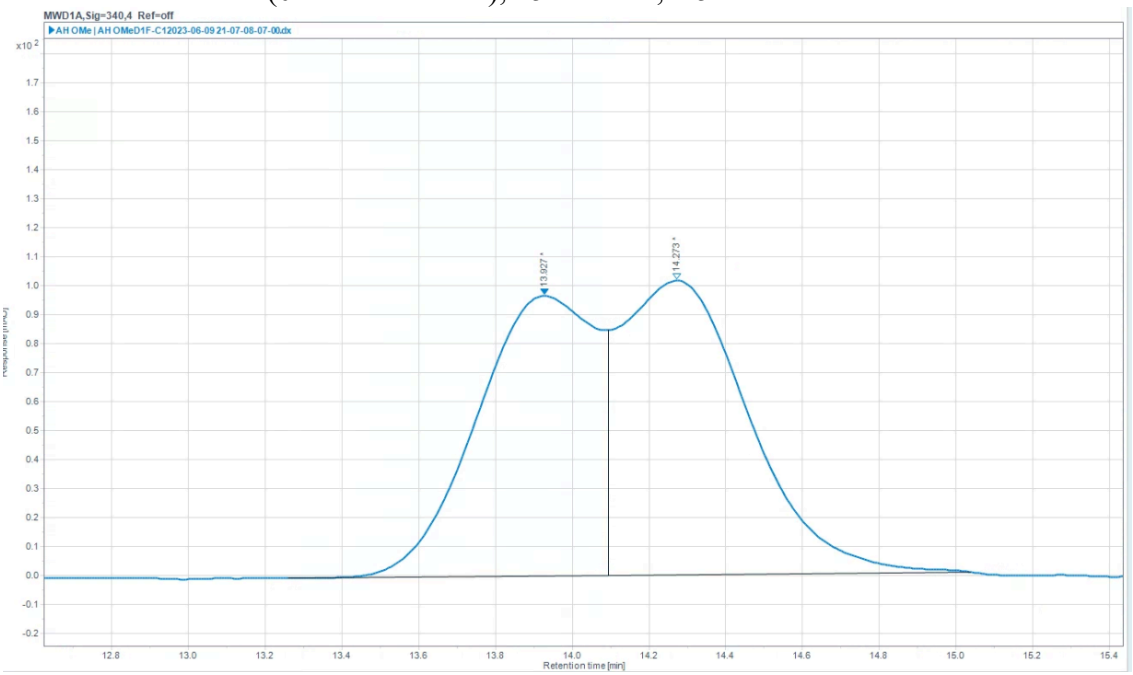


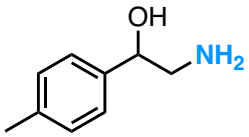
5a-rac – reverse phase – racemic standard
HPLC conditions: C18 (Poroshell, EC-C18, 4.6 × 150 mm, 5 µm) column, 25–95% acetonitrile in water (0.1% acetic acid), 15 minutes, 1.5 mL/min



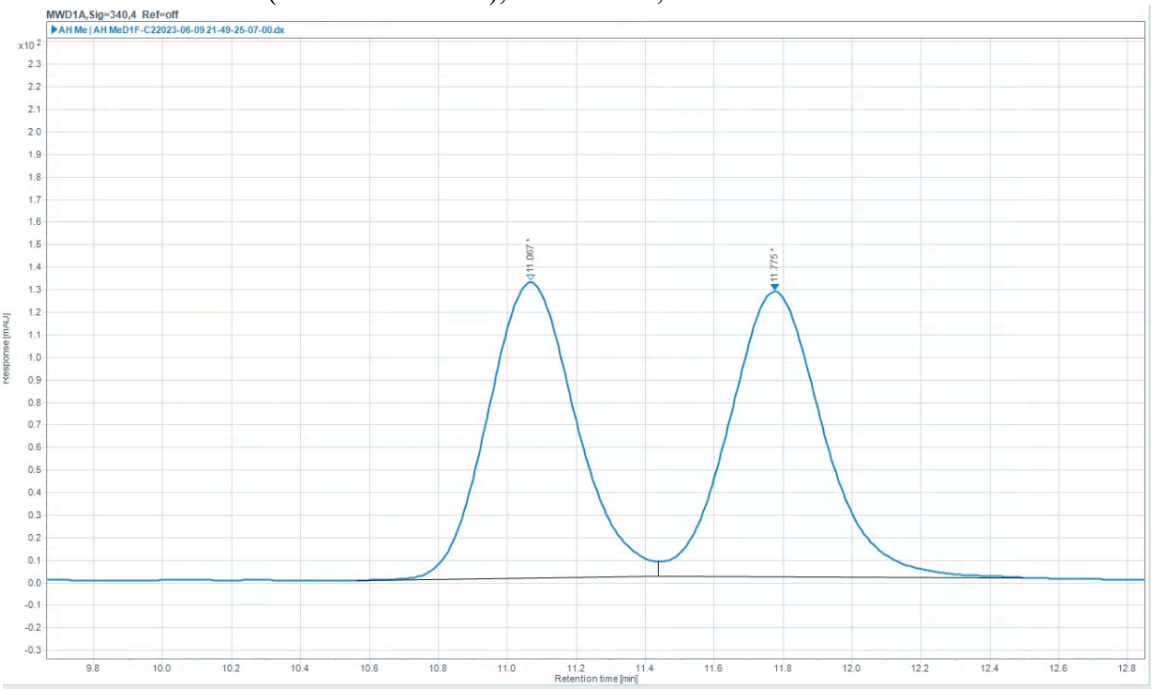


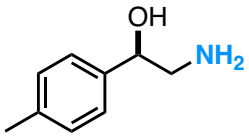
5a-enz – reverse phase – enzymatic reaction: racemic
HPLC conditions: C18 (Poroshell, EC-C18, 4.6 × 150 mm, 5 µm) column, 25–95% acetonitrile in water (0.1% acetic acid), 15 minutes, 1.5 mL/min



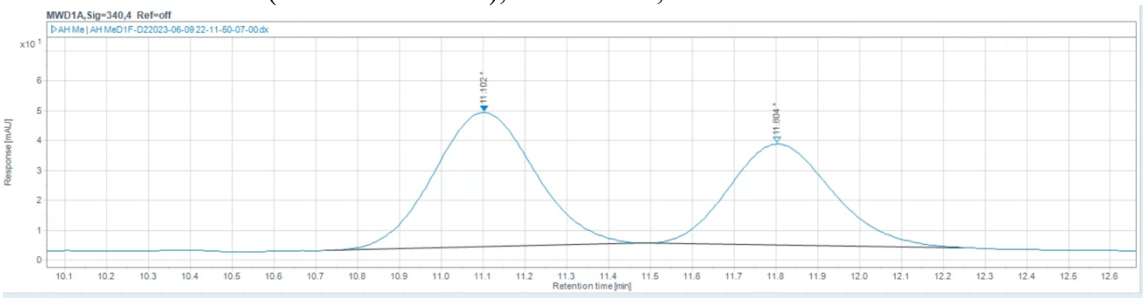


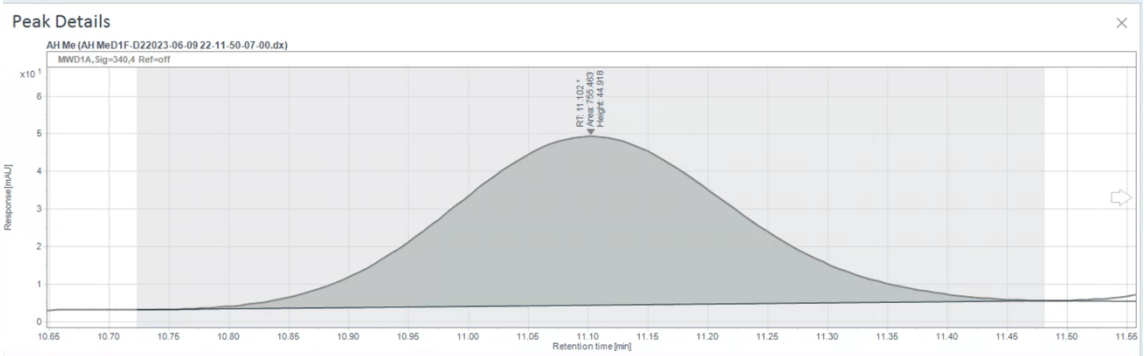
5b-rac – reverse phase – racemic standard
HPLC conditions: C18 (Poroshell, EC-C18, 4.6 × 150 mm, 5 µm) column, 30–95% acetonitrile in water (0.1% acetic acid), 15 minutes, 1.5 mL/min

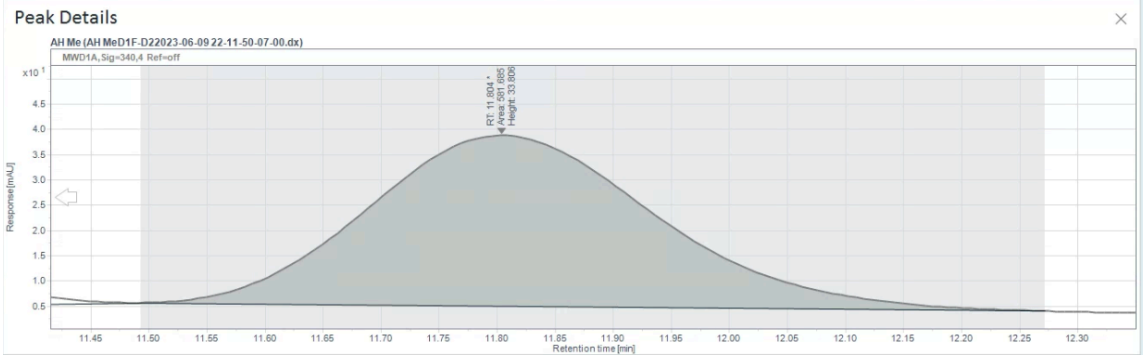


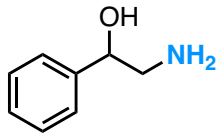


5b-enz – reverse phase – enzymatic reaction: 12% *ee*
HPLC conditions: C18 (Poroshell, EC-C18, 4.6 × 150 mm, 5 µm) column, 30–95% acetonitrile in water (0.1% acetic acid), 15 minutes, 1.5 mL/min

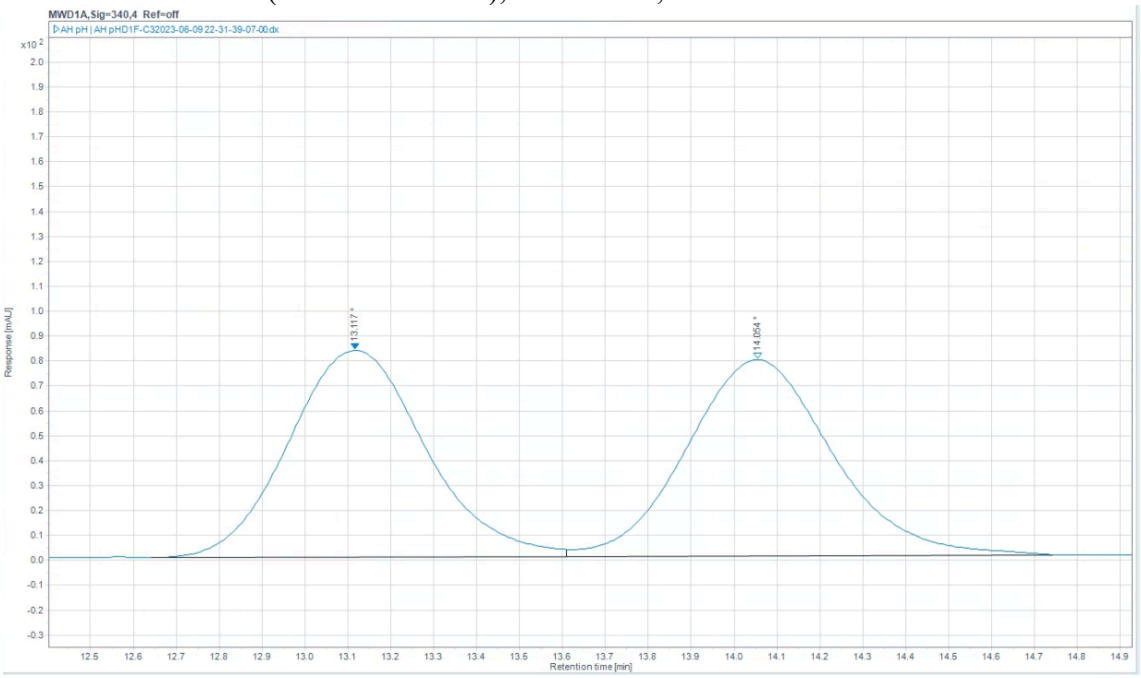


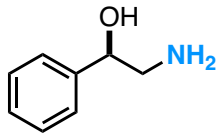




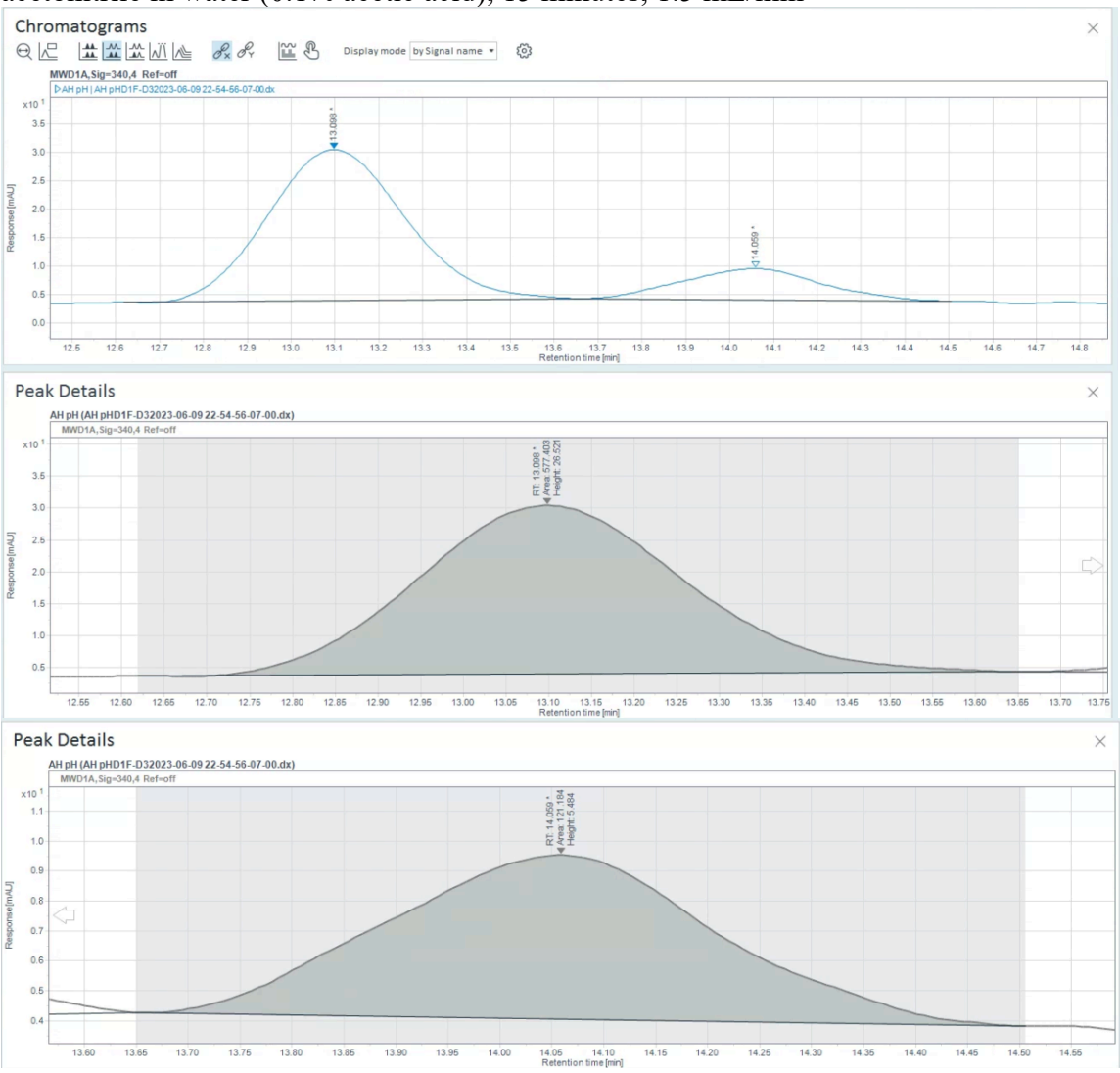


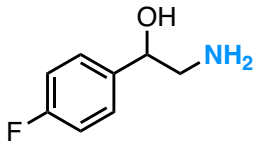
5c-rac – reverse phase – racemic standard
HPLC conditions: C18 (Poroshell, EC-C18, 4.6 × 150 mm, 5 µm) column, 35–95% acetonitrile in water (0.1% acetic acid), 15 minutes, 1.5 mL/min



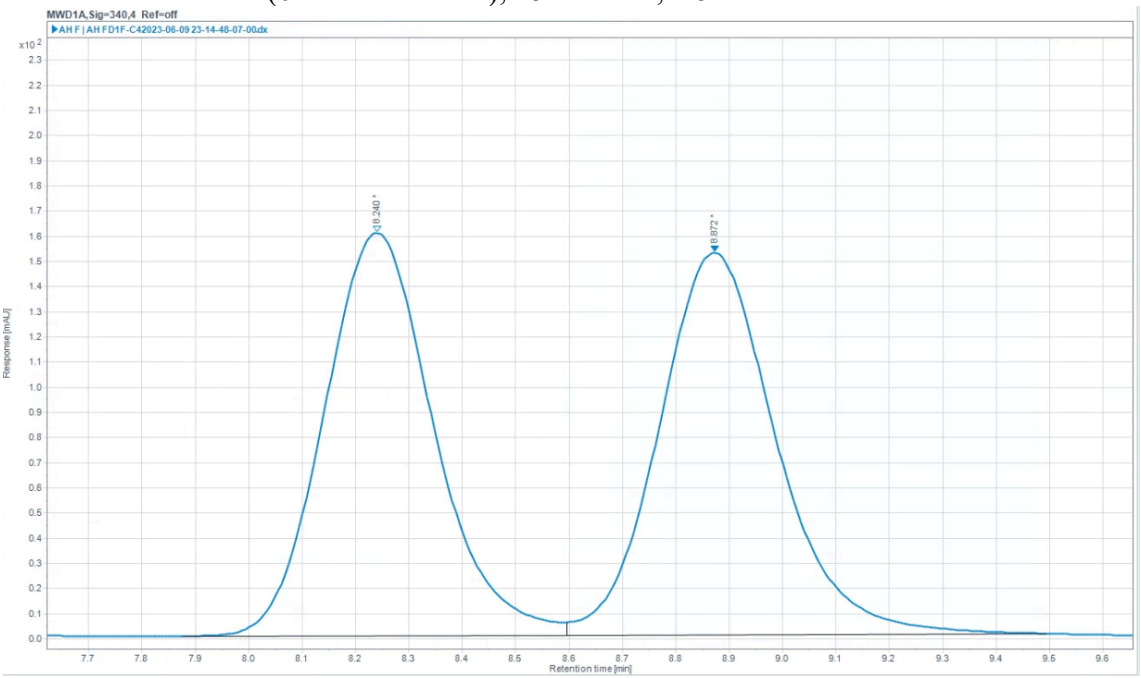


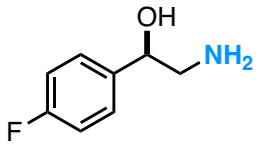
5c-enz – reverse phase – enzymatic reaction: 65% *ee*
HPLC conditions: C18 (Poroshell, EC-C18, 4.6 × 150 mm, 5 µm) column, 35–95% acetonitrile in water (0.1% acetic acid), 15 minutes, 1.5 mL/min



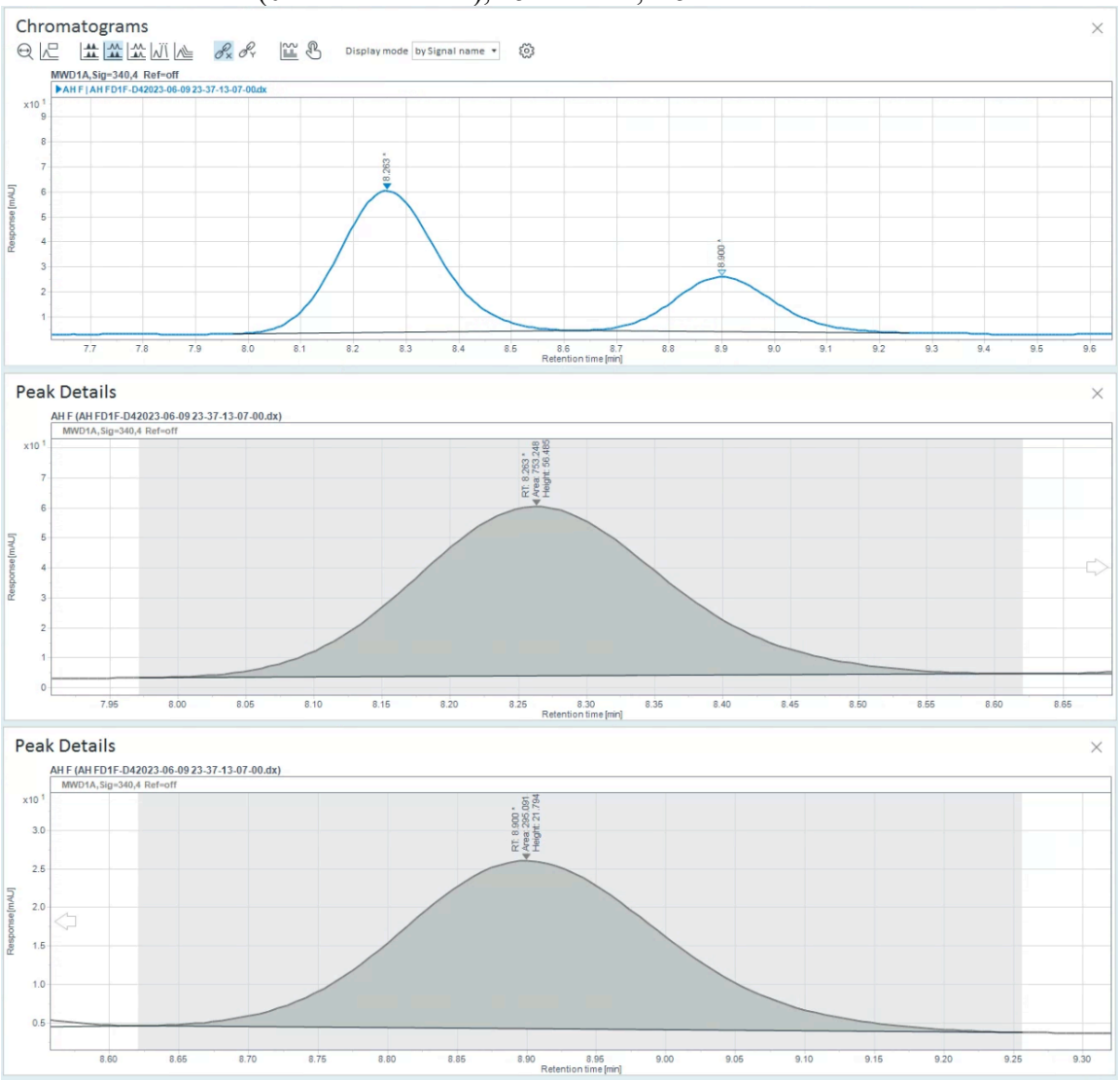


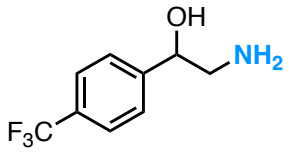
5d-rac – reverse phase – racemic standard
HPLC conditions: C18 (Poroshell, EC-C18, 4.6 × 150 mm, 5 µm) column, 35–95% acetonitrile in water (0.1% acetic acid), 15 minutes, 1.5 mL/min



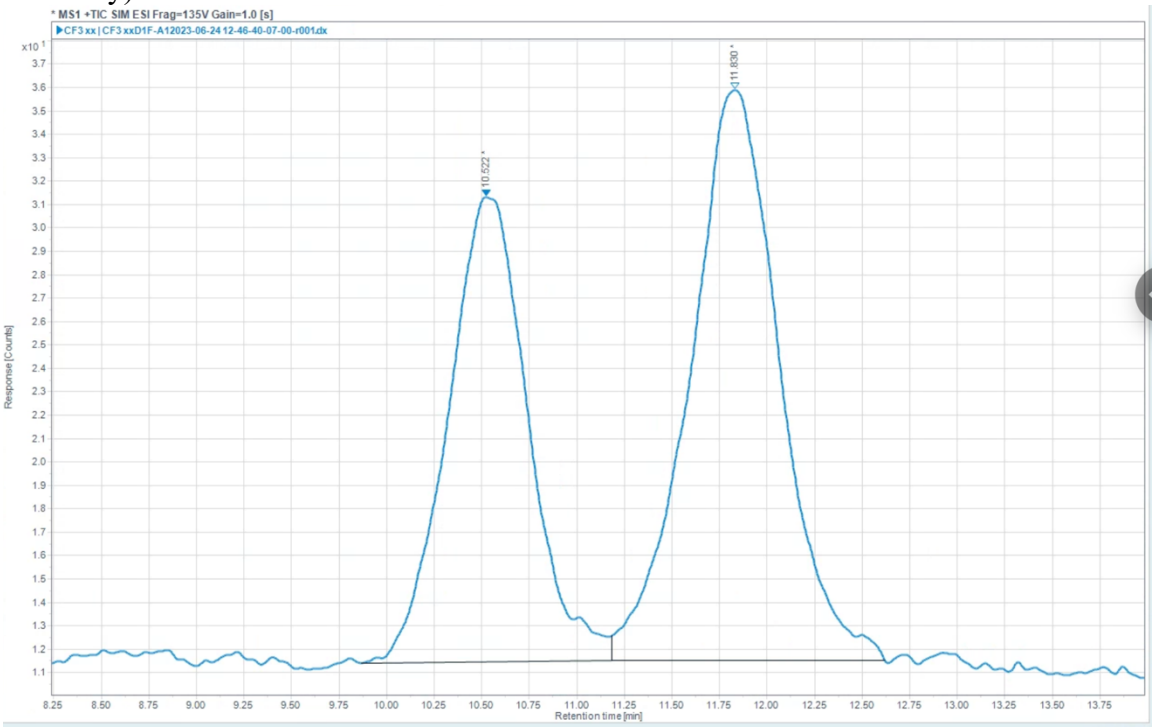


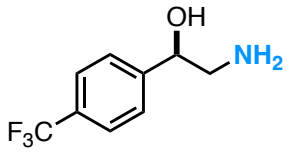
5d-enz – reverse phase – enzymatic reaction: 44% *ee*
HPLC conditions: C18 (Poroshell, EC-C18, 4.6 × 150 mm, 5 µm) column, 35–95% acetonitrile in water (0.1% acetic acid), 15 minutes, 1.5 mL/min



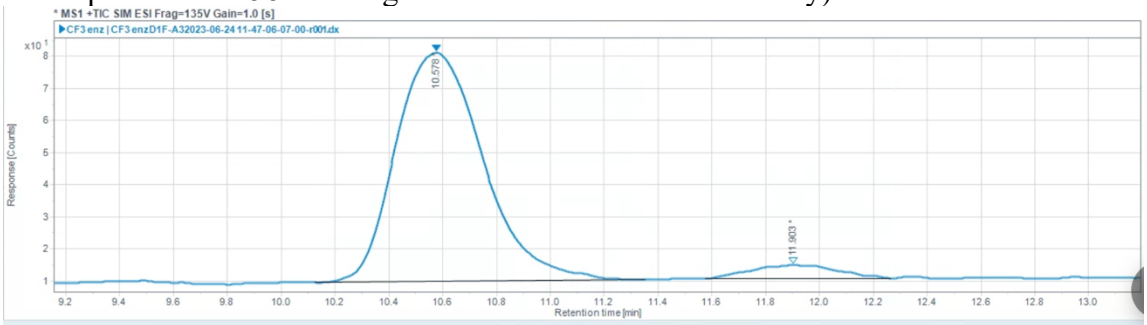


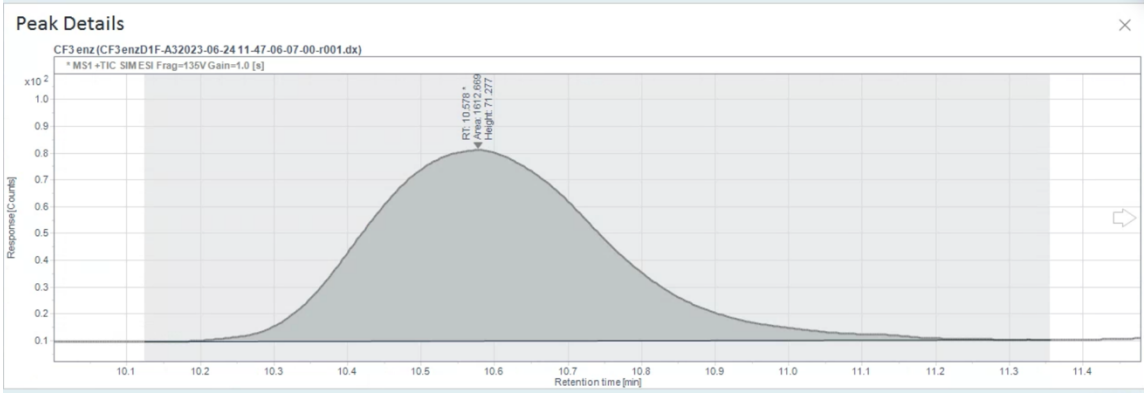
5e-rac – reverse phase racemic standard
HPLC conditions: C18 (Poroshell, EC-C18, 4.6 × 150 mm, 5 µm) column, 40–95% acetonitrile in water (0.1% acetic acid), 15 minutes, 1.5 mL/min (**5e** was detected by mass spectrum in 458 m/z single ion channel instead of UV spectrum due to low reactivity)

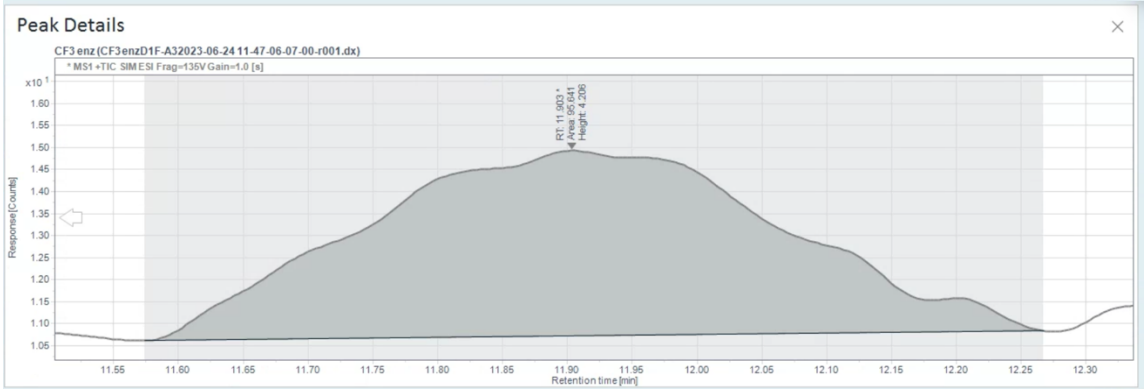




5e-enz – reverse phase enzymatic reaction: 88% *ee*
HPLC conditions: C18 (Poroshell, EC-C18, 4.6 × 150 mm, 5 µm) column, 30–95% acetonitrile in water (0.1% acetic acid), 15 minutes, 1.5 mL/min (**5e** was detected by mass spectrum in 458 m/z single ion mode due to low reactivity)







C.6. Mutation Transfer Experiment.

a

```
Aeropyrum_pernix_protoglobin_wildtype                MTPSDIPGYDYG--RVEKSPITDLEFDLLKKTVMLGEKDVMYLKKACDVLKDQVDEILDL  58
Pyrobaculum_arsenaticum_protoglobin_wildtype         ---MAVPGYDFG--KVPDAPISDADFESLKKTVMWGEEDEKYRKMACEALKGQVEDILDL  55
Pyrobaculum_ferrireducens_protoglobin_wildtype       ---MREIPGYEFG--KVPDAPISDEEFELLKKSVMWTEEDEKYRKLACEVLKGQVEQILDL  56
Crenotalea_thermophila_protoglobin_wildtype          MANTFIPGYTYG--QVPEAPFSMHDFELLKQTVMFTEKDAEYLRKAGEILHDQVNEILDL  58
Thermorudis_peleae_protoglobin_wildtype              -MMSQIPGYTYGTTAVPRSPISLEEFELLKKTVLFTDEDIAALRQAHDVLADQVDAILDV  59
Pyrinomonas_methylaliphatogenes_protoglobin_wildtype MKENKITGYTYGTDEVARSPLSEQEFDLLKQTVLFTDEDVHYLRMAGEVFADQVEDVLDV  60
Methanosarcina_acetivorans_protoglobin_wildtype      MSVEKIPGYTYGETE-NRAPFNLEDLKLLKEAVMFTAEDEEYIQKAGEVLEDQVEEILDT  59
Thermanaerothrix_daxensis_protoglobin_wildtype       ---MSAIPGYTYGLTE-ARSPLSPTDLEHLKQAVLFGPEDEHYLRLAGEVLADQIEEILDV  57
Thermoflexus_hugenholtziin_protoglobin_wildtype      MSVQTIPGYTYGTAEAARSPVSMEEFADLQKATMFTEEDVRYLRMAGEVLADQVEDILDL  60
Acetothermus_autotrophicum_protoglobin_wildtype      -MAVAIPGYTYS---SVPQSPVTMKELNELKQAVLFSAEDEKYLKMAGEVLKDQIDAVLDL  57
Thermus_arciformis_protoglobin_wildtype              ----MVLPGYTYGDPGLPPSPVSLEALHQLKAALLWSEEDEKALKEAGEVLKDQVEEVLDL  57
Thermus_amyloliquefaciens_protoglobin_wildtype       ---MPQIPGYTYGDPALPPSPVSLEELERLKACLLWTEEDDKALEQAGKVLEDQVEEVLDL  58
                                                          : ** :.     :*..   :  *:   :   :*    . * . : .*:: :**

Aeropyrum_pernix_protoglobin_wildtype                WYGWVASNEHLIYYFSNPDTGEPIKEYLERVRARFGAWILDTTCRDYNREWLDYQYEVGL  118
Pyrobaculum_arsenaticum_protoglobin_wildtype         WYGWVGSNQHLIYYFGD-KSGRPIPQYLEAVRKRFGLWIIDTLCKPLDRQWLNYMYEIGL  114
Pyrobaculum_ferrireducens_protoglobin_wildtype       WYGWVGSNPHLVYFGD-RSGRPIPQYLEAVRKRFGQWILDTVCRSYDRQWLNYVYEIGL  115
Crenotalea_thermophila_protoglobin_wildtype          WYGFVGQHPHLLAYFAK--DGKVDSSYLQAVRKRFGQWIHDLCEKPFDQDWLNYQYEIGL  116
Thermorudis_peleae_protoglobin_wildtype              WYGFVGANPHLLYYFTRRSDGQPDTTYLDRVRQRFGQWIRDLTAANYNQAWLDYQYEIGL  119
Pyrinomonas_methylaliphatogenes_protoglobin_wildtype WYGFVASHPHLVYYFSD-RDGKPIPEYLTSVRARFGQWIRDTCAARYDRAWLDYQQEIAL  119
Methanosarcina_acetivorans_protoglobin_wildtype      WYGFVGSHPHLLYYFTS-PDGTPNEKYLAAVRKRFSRWILDTCNRSYDQAWLDYQYEIGL  118
Thermanaerothrix_daxensis_protoglobin_wildtype       WYGFVGSHPHLVYYFTD-GHGNPNAEYLAAVRRRFGQWILDTCTRPYDQDWLNYQHEIGL  116
Thermoflexus_hugenholtziin_protoglobin_wildtype      WYSWVGSQPHLVYYFSG-PDGKPDMNYLQAVRRRFGQWILDTCNRPYDQDWLNYQQEIGL  119
Acetothermus_autotrophicum_protoglobin_wildtype      WYGFVGSHPHLVYYFSG-PDGKPDANYLAAVRKRFGQWILDTCTRPYDQTWLNYQHEIGL  116
Thermus_arciformis_protoglobin_wildtype              WYSFVGSHPHLLHYFTD-GQGNPIPEYLERVRKRFGQWVLDTCFRPHDEAWLAYQHEIGL  116
Thermus_amyloliquefaciens_protoglobin_wildtype       WYGFVGSHPHLLHYFTD-PQGNPIPDYLERVRRRFGQWILDTCFRPKDETWLRYQHEIGL  117
                                                     **.:*. : **: **     *     **  ** **. *: *      :. ** *  *:.*

Aeropyrum_pernix_protoglobin_wildtype                RHHRSKKGVTDGVRTV-PHIPLRYLIAFIYPITATIKPFLAKKGGSPEDIEGMYNAWFKS  177
Pyrobaculum_arsenaticum_protoglobin_wildtype         RHHRTKKGKTDGVDTV-EHIPLRYMIAFIAPIGLTIKPILEKSGHPPEAVERMWAAWVKL  173
Pyrobaculum_ferrireducens_protoglobin_wildtype       RHHRTKKGKTDGVETV-EHIPLRYMVAFIAPIGLTIKPFLEKGGHPPDVVEKMWAAWIKS  174
Crenotalea_thermophila_protoglobin_wildtype          RHHATKKNLTDHAQAA-PLIHYRYMVAFIYPITATIRAFLARKGHPAELVEQMHQAWFKA  175
Thermorudis_peleae_protoglobin_wildtype              RHHRSKKNQTDHVEAP-EHIPLRYVLALTYPIYATIKPFLAKKGHSAEEVERMHQAWLKA  178
Pyrinomonas_methylaliphatogenes_protoglobin_wildtype RHTRAKKNRTDRVESIADHIPLRYMIAFIYPITATIKPFLAKKGHAEEDVEKMHQAWFKS  179
Methanosarcina_acetivorans_protoglobin_wildtype      RHHRTKKNQTDNVESV-PNIGYRYLVAFIYPITATMKPFLARKGHTPEEVEKMYQAWFKA  177
Thermanaerothrix_daxensis_protoglobin_wildtype       RHHRSKKNRTDGVEAV-PIVHHRYLVAFIYPITATIKPFLARKGHSPEEVEKMHQAWFKS  175
Thermoflexus_hugenholtziin_protoglobin_wildtype      RHHRTKKNQTDGVRSV-PHIPMRYLIAFIYPITSTIRPFLAKKGHSPEDVEKMHQAWTKS  178
Acetothermus_autotrophicum_protoglobin_wildtype      RHHRTKKNKTDNVRSV-PNIPLRYLIAFIYPITATIKPFLAKKGHSPDEVDKMHHAWFKA  175
Thermus_arciformis_protoglobin_wildtype              RHHRTKKNQTDGVASV-PHIPLRYLIAFIYPITATIRPFLAKKGHPPEAVERMHQAWFKA  175
Thermus_amyloliquefaciens_protoglobin_wildtype       RHHRTKKNQTDGVTSV-PHIPLRYLISFIYPITATIKPFLTKKGHNPEEVERMYQAWFKA  176
                                                     **  :**. ** . :    :  **:::  **  *:: :* : *   : :: *  ** *

Aeropyrum_pernix_protoglobin_wildtype                VVLQVAIWSHPYTKENDW    195
Pyrobaculum_arsenaticum_protoglobin_wildtype         VVLQVAIWSYPYAKTGEW    191
Pyrobaculum_ferrireducens_protoglobin_wildtype       VVLQVAIWSHPYAKPGEW    192
Crenotalea_thermophila_protoglobin_wildtype          VVLTVVLWTYPYVKQGEF    193
Thermorudis_peleae_protoglobin_wildtype              VLLSVILWSQPYVRDGDF    196
Pyrinomonas_methylaliphatogenes_protoglobin_wildtype VVMQVALWSAPYARPGDF    197
Methanosarcina_acetivorans_protoglobin_wildtype      TTLQVALWSYPYVKYGDF    195
Thermanaerothrix_daxensis_protoglobin_wildtype       VVMQVALWSAPYVHPEDF    193
Thermoflexus_hugenholtziin_protoglobin_wildtype      VIMQVALWCYPYTREGDW    196
Acetothermus_autotrophicum_protoglobin_wildtype      VTLQVALWSYPYAKDGDF    193
Thermus_arciformis_protoglobin_wildtype              VVLQVALWSYPYAKEGDF    193
Thermus_amyloliquefaciens_protoglobin_wildtype       VVLQVALWSYPYTQPGDF    194
                                                     . : * :*  **.:  ::
```

b

```
final_variant_5      MAVPGYDFGKVPDAPISDADFESLKKTAMWGEEDEKYRKMACEALKGQVEDILDLWDGLQ       60
ParPgb_HYA_5213      MAVPGYDFGKVPDAPISDADFESLKKTVMWGEEDEKYRKMACEALKGQVEDILDLWDGLQ       60
final_variant_7      MAVPGYDFGKVPDAPISDADFESLKRTVMWGEDDEKYRKMACEALKGQVEDILDLWDGLQ       60
final_variant_2      MAVPGYDFGKVPDAPISDADFESLKRTVMWGEDDEKYRKMACEALKGQVEDILDLWDGLQ       60
final_variant_6      MAVPGYDFGKVPDAPISDADFESLKRTVMWGEEDEKYRKMACEALKGQVEDILDLWDGLQ       60
ParPgb_wildtype      MAVPGYDFGKVPDAPISDADFESLKKTVMWGEEDEKYRKMACEALKGQVEDILDLWYGWV       60
final_variant_3      MAVPGYDFGKVPDAPISDADFESLKKTVMWGEEDEKYRKMACEALKGQVEDILDLWDGLQ       60
final_variant_4      MAVPGYDFGKVPDAPISDADFESLKKTVMWGEEDEKYRKMACEALKGQVEDILDLWDGLQ       60
final_variant_8      MAVPGYDFGKVPDAPISDADFESLKRTVMWGEDDEKYRKMACEALKGQVEDILDLWDGLQ       60
                     *************************:*.****:********************** *

final_variant_5      GSNQHLIYYFGDKSSRPIPQYLEAVRKRFGLWIIDTLCKPLDRQWLNYMYEIGLRHHRTK       120
ParPgb_HYA_5213      GSNQHLIYYFGDKSGRPIPQYLEAIRKRFGLWIIDTLCKPLDRQWLNYMYEIGLRHHRTK       120
final_variant_7      GSNQHLIYYFGDKSGRPIPQYLEAVRKRFGLWIIDTLCKPLDRQWLNYMYEIGLRHHRTK       120
final_variant_2      GSNQHLIYYFGDKSGRPIPQYLEAVRKRFGLWIIDTLCKPLDRQWLNYMYEIGLRHHRTK       120
final_variant_6      GSNQHLIYYFGDKSGRPIPQYLEAVRKRFGLWIIDTLCKPLDRQWLNYMYEIGLRHHRTK       120
ParPgb_wildtype      GSNQHLIYYFGDKSGRPIPQYLEAVRKRFGLWIIDTLCKPLDRQWLNYMYEIGLRHHRTK       120
final_variant_3      GSNQHLIYYFGDKSGRPIPQYLEAVRKRFGLWTIDTLCKPLDRQWLNYMYEIGLRMHRTK       120
final_variant_4      GSNQHLIYYFGDKSGRPIPQYLEAVRKRFGLWTIDTLCKPLDRQWLNYMYEIGLRMHRTK       120
final_variant_8      GSNQHLIYYFGDKSGRPIPQYLEAVRKRFGLWTIDTLCKPLDRQWLNYMYEIGLRMHRTK       120
                     **************.*********:******* ********************** ****

final_variant_5      KGKTDGVDTVEHIPLRYMIAFIAPIGLTFKPILEKSGHPPEAVERMWAAWVKLVVLRVAI       180
ParPgb_HYA_5213      KGKTDGVDTVEHIPLRYMIAFIAPIGLTFKPILEKSGHPPEAVERMWAAWVKLVALRVAI       180
final_variant_7      KGKTDGVDTVEHIPLRYMIAFIAPIGLTIKPILVKSGHPPEAVERMWAAWVKLVVLRVAI       180
final_variant_2      KGKTDGVDTVEHIPLRYMIAFIAPIGLTFKPILEKSGHPPEAVERMWAAWVKLVVLRVAI       180
final_variant_6      KGKTDGVDTVEHIPLRYMIAFIAPIGLTFKPILEKSGHPPEAVERMWAAWVKLVVLRVAI       180
ParPgb_wildtype      KGKTDGVDTVEHIPLRYMIAFIAPIGLTIKPILEKSGHPPEAVERMWAAWVKLVVLQVAI       180
final_variant_3      KGKTDGVDTVEHIPLRYMIAFIAPIGLTIKPILEKSGHPPEAVERMWAAWVKLVALQVAI       180
final_variant_4      KGKTDGVDTVEHIPLRYMIAFIAPIGLAIKPILVKSGHPPEAVERMWAAWVKLVVLQVAI       180
final_variant_8      KGKTDGVDTVEHIPLRYMIAFIAPIGLAIKPILVKSGHPPEAVERMWAAWVKLVVLQVAI       180
                     **************************::**** ********************.*:***

final_variant_5      WSYPYAKTGEWLEHHHHHH        199
ParPgb_HYA_5213      WSYPYAKTGEWLEHHHHHH        199
final_variant_7      WSYPYAKTGEWLEHHHHHH        199
final_variant_2      WSYPYAKTGEWLEHHHHHH        199
final_variant_6      WSYPYAKTGEWLEHHHHHH        199
ParPgb_wildtype      WSYPYAKTGEWLEHHHHHH        199
final_variant_3      WSYPYAKTGEWLEHHHHHH        199
final_variant_4      WSYPYAKTGEWLEHHHHHH        199
final_variant_8      WSYPYAKTGEWLEHHHHHH        199
                     ******************
```

Figure C-3. a. Multiple sequence alignment (MSA) of wild-type protoglobin homologs. **b.** MSA of final variants, along with wild-type *Par*Pgb.

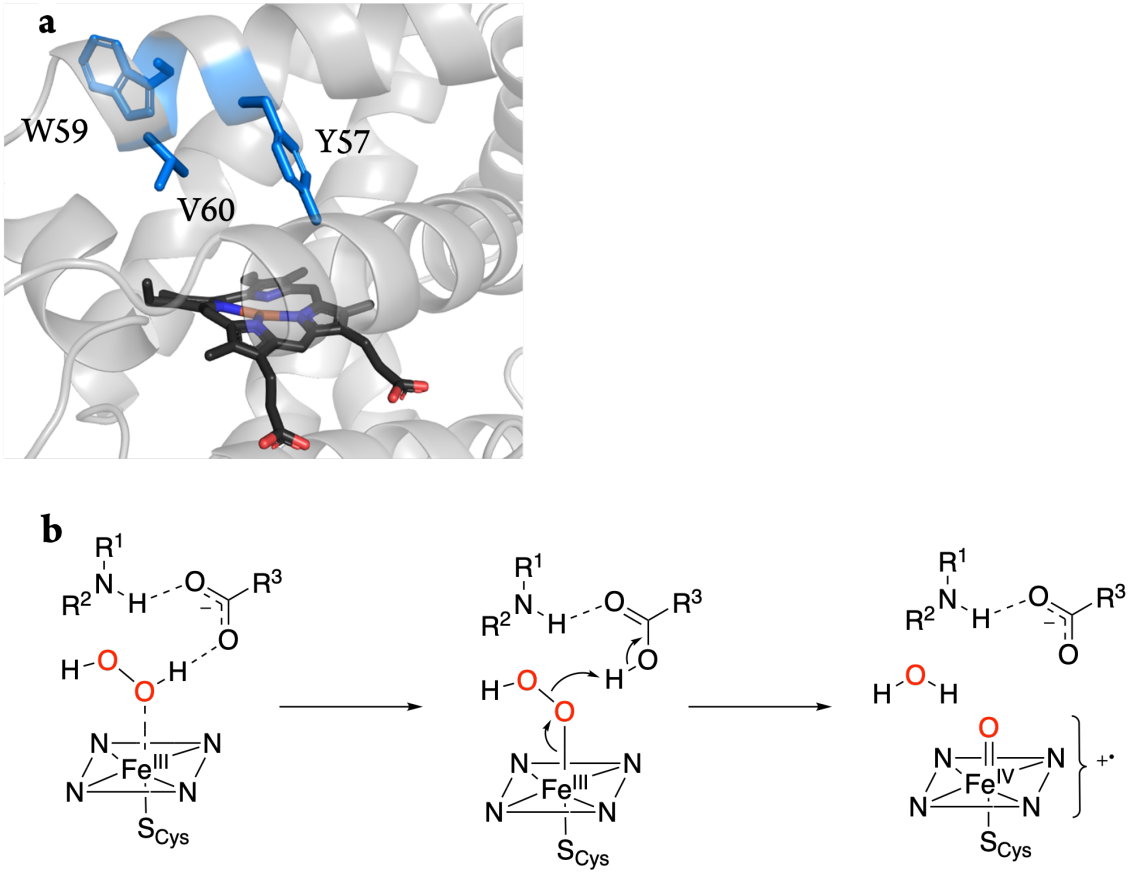


Figure C-4. **a.** The active site and residues Y57, W59, V60 in the homology model of *Par*Pgb wild type based on the crystal structure of *Mac*Pgb (PDB: 2VEE). **b.** Proposed mechanism in literature for the formation of compound I in the presence of acid–base catalytic residues in heme-thiolate enzymes.[10]

Summary of Mutated Variants

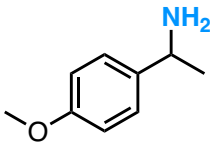
Variant Name	Mutations Relative to Wild-type Protein
AauPgb-HYA-5214	Y59D W61L V62Q
ApePgb-HYA-5215	Y60D W62L V63Q
CthPgb-HYA-5216	Y60D W62L V63Q
MacPgb-HYA-5217	Y61D W63L V64Q
ParPgb-HYA-5218	Y57D W59L V60Q
PfePgb-HYA-5219	Y58D W60L V61Q

PmePgb-HYA-5220	Y62D W64L V65Q
TamPgb-HYA-5221	Y60D W62L V63Q
TarPgb-HYA-5222	Y59D W61L V62Q
TdaPgb-HYA-5223	Y59D W61L V62Q
ThuPgb-HYA-5224	Y62D W64L V65Q
TpePgb-HYA-5225	Y61D W63L V64Q

Experimental setup and data analysis:

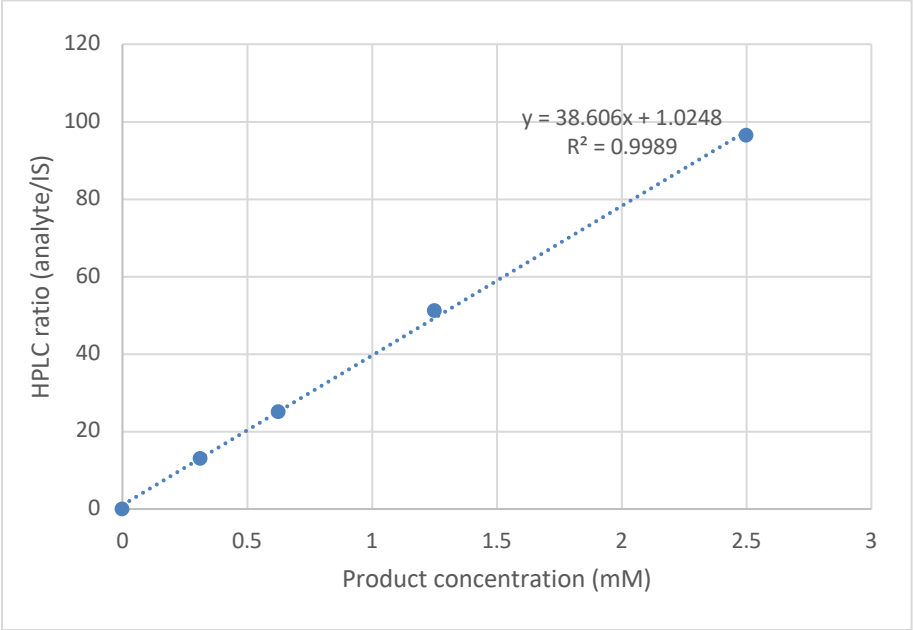
Site-directed mutagenesis was used to introduce 57D, 59L, and 60Q mutations in each variant following the procedure in Section 1.1. Each variant was expressed in 50 mL of HB medium following the procedure in Section 1.2. After centrifugation and removal of the supernatant, 6 mL of KPi buffer were used to resuspend the cell pellets. Then, the cells were lyzed by sonication and centrifuged (14,000*g*, 10 minutes, 4 °C). The supernatant was applied to a 1-mL Ni^{2+}-NTA column equilibrated with 50 mM KPi (pH 8.0) at 4 °C. Once all lysate had been applied, the column was washed with 10 to 20 mL of buffer A (KPi buffer with 20 mM imidazole, pH 8.0) until the flow-through was clear. The protein-of-interest was eluted with 1 to 2 mL buffer B (KPi buffer with 500 mM imidazole, pH 8.0). The protein solution was dialyzed in 4 L of KPi buffer at 4 °C overnight. The protein concentration was measured by NanoDrop, and the heme concentration was measured by hemochrome assay. Next, exogenous heme *b* cofactor (20 mM stock solution in 0.1 M NaOH solution) was supplemented according to the difference between protein concentration and heme concentration. The protein solution was chilled on ice for 30 min and centrifuged (5000*g*, 10 minutes, 4 °C). The final protein concentrations were measured

following the procedure in Section 1.3. The reactions were setup following the procedure in Section 4.1. and analyzed following the procedure in Section 4.2.



*1-(4-Methoxyphenyl)ethan-1-amine (**3a**)*

Standard curve based on cation account in 135 m/z single ion mode for product **3a** and 1,3,5-trimethoxybenzene on UV absorbance at 280 nm as internal standard.



Variant	Pdt	IS	Pdt/IS	[Pdt] (µM)	yield %	[PC] (µM)	TTN	Average TTN	STD TTN
AauPgb-HYA-5214	845.98	127.87	6.62	144.76	1.45	52.63	2.8		
	869.43	128.02	6.79	149.31	1.49	52.63	2.8		
	792.29	122.32	6.48	141.17	1.41	52.63	2.7	2.8	0.06
ApePgb-HYA-5215	0	120.40	0	0	0	4.52	0	0	0
CthPgb-HYA-5216	297.01	126.36	2.35	34.27	0.34	48.13	0.71		
	302.60	125.29	2.42	35.95	0.36	48.13	0.75		
	351.24	124.47	2.82	46.48	0.46	48.13	0.97	0.73	0.11
MacPgb-HYA-5217	0	124.65	0	0	0	9.68	0	0	0
ParPgb-HYA-5218	8055.94	121.48	66.31	1691.14	16.9	59.03	29		
	8350.89	123.95	67.37	1718.55	17.2	59.03	29		
	7465.50	121.12	61.64	1570.00	15.7	59.03	27	30	1.1
	4153.89	119.66	34.72	872.63	8.73	35.62	25	25	0.84

PfePgb-HYA-5219	3898.93	121.78	32.02	802.70	8.03	35.62	23		
	4114.01	126.20	32.60	817.80	8.18	35.62	23		
PmePgb-HYA-5220	0	123.41	0	0	0	2.42	0	0	0
TamPgb-HYA-5221	761.69	124.94	6.10	131.30	1.31	46.05	2.9		
	803.67	125.64	6.40	139.08	1.39	46.05	3.0		
	742.36	118.15	6.28	136.14	1.36	46.05	3.0	2.9	0.07
TarPgb-HYA-5222	1167.55	127.59	9.15	210.42	2.10	51.91	4.1		
	1327.37	124.11	10.69	250.42	2.50	51.91	4.8		
	1209.32	125.51	9.64	222.97	2.23	51.91	4.3	4.4	0.32
TdaPgb-HYA-5223	534.33	133.55	4.00	77.03	0.77	50.00	1.5		
	483.08	130.39	3.70	69.35	0.69	50.00	1.4		
	367.28	128.13	2.87	47.64	0.48	50.00	1.0	1.5	0.25
ThuPgb-HYA-5224	259.37	118.74	2.18	29.97	0.30	49.50	0.61		
	310.65	120.63	2.58	40.09	0.40	49.50	0.81		
	251.57	112.84	2.23	31.13	0.31	49.50	0.63	0.68	0.09
TpePgb-HYA-5225	279.51	128.46	2.18	29.74	0.30	44.96	0.66		
	284.49	128.90	2.21	30.55	0.31	44.96	0.68		
	237.51	126.55	1.88	22.00	0.22	44.96	0.49	0.68	0.09

Notes: Pdt = product area, IS = internal standard area, [Pdt] = product concentration in reaction, [PC] = protein concentration in reaction, Avg TTN = average of total turnover number, SD TTN = standard deviation of TTN.

C.7. Kinetic Experiment.

C.7.1. Experimental Details

Experimental setup:

In each 2-mL screw cap vial, the hydrocarbon substrate (10 mM final concentration, 36 μL of 200 mM), sodium dithionite (10 mM final concentration, 36 μL of 200 mM), and purified enzymes (1 μM final concentration for the final variant, 10 μM final concentration for the initial variant) were added. Hydroxylamine (0.025–50 mM various final concentrations, 36 μL of 0.5–1000 mM) was added (720 μL total volume) to start the reaction. At 60, 120, and 180 seconds, 200 μL of the reaction mixture from each vial were added to a 1.7-mL Eppendorf tube containing 400 μL of ethanol to quench the reactions. The reaction mixture was vortexed and centrifuged (14,000g, 15 minutes, 4 °C). The initial rates of several C–H substrates were tested and ethylbenzene **2c** was chosen as the reaction partner to accurately capture the initial rate. All time points were collected in triplicates.

Data analysis:

The reaction mixture was analyzed by LCMS, and the product concentrations were calculated with a calibration curve within the linear region. The data points that cannot be integrated accurately were discarded. The initial rate was calculated by plotting the slope of each concentration set. The initial rates were then used to fit the kinetic model in a Python notebook.

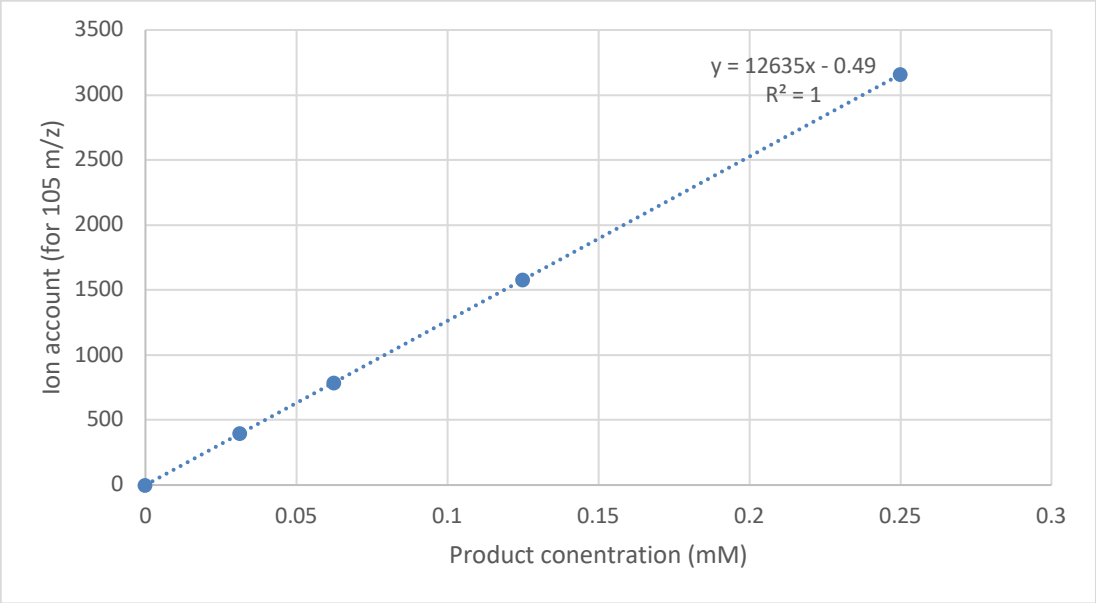
Kinetic model:

$$v = Vmax\frac{[S]}{Km + [S]}$$

C.7.2. Michaelis-Menten Plots

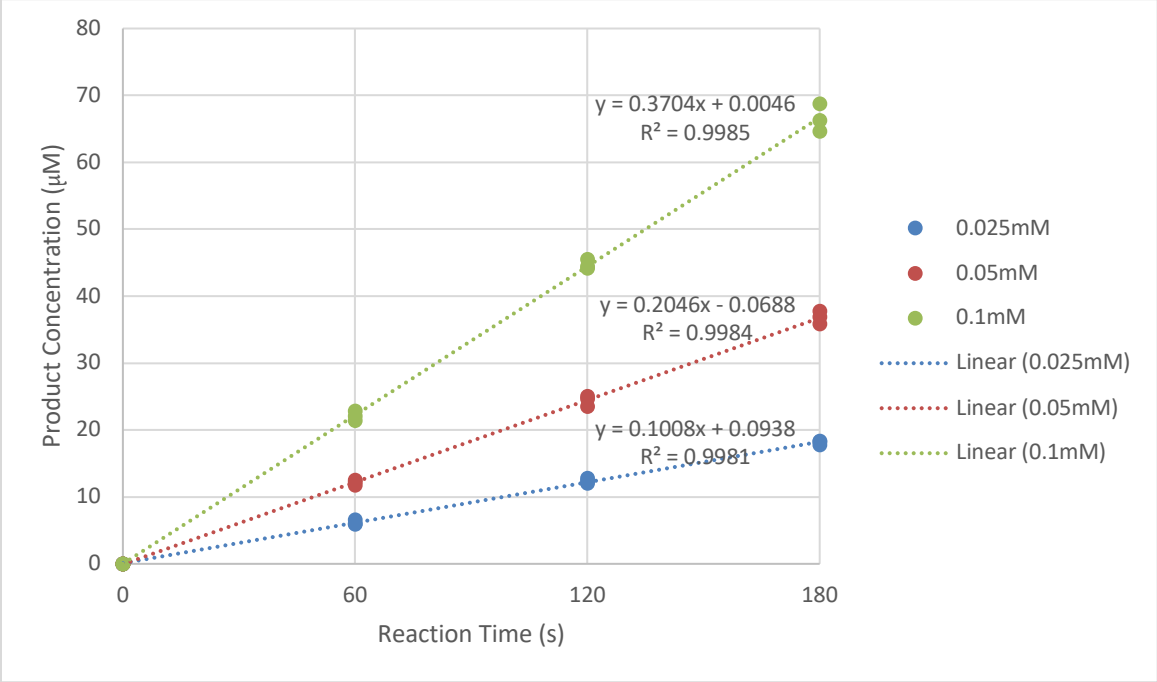
Standard Curve of **3c** for KIE experiments:

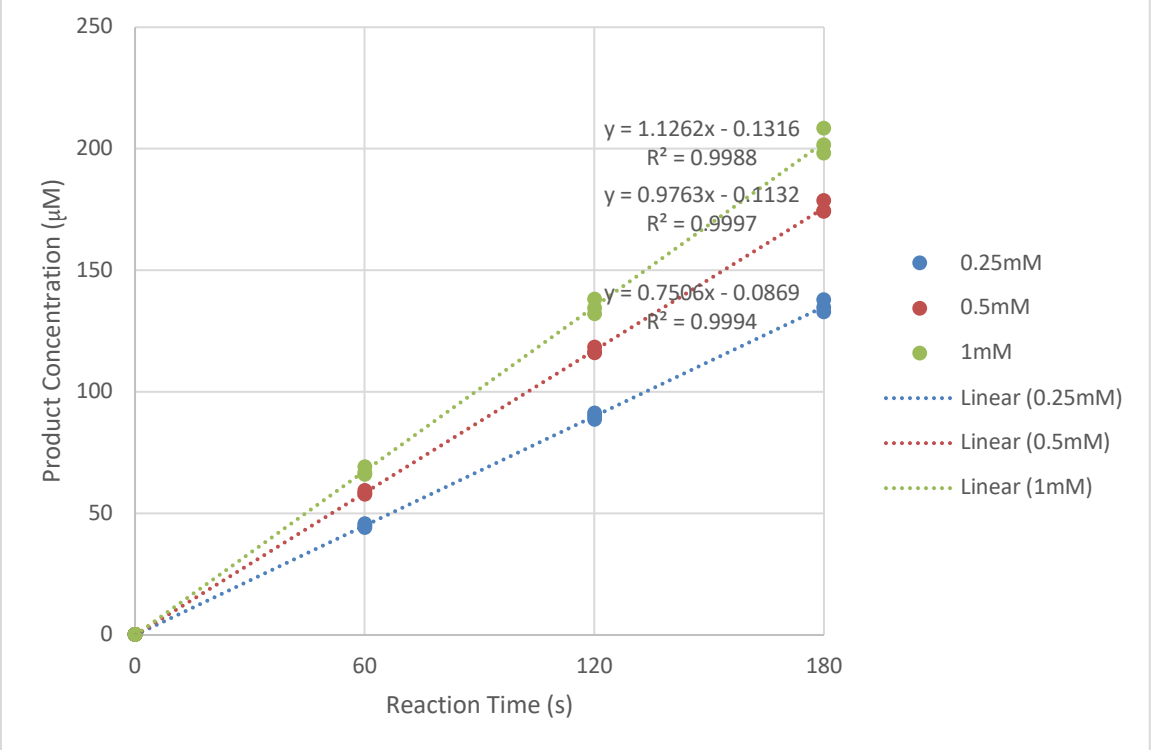
Samples for the standard curve were prepared following the procedure in Section 4.2., but quenching with 800 µL ethanol instead. Standard curve based on cation account in 105 m/z single ion mode for product **3c**.

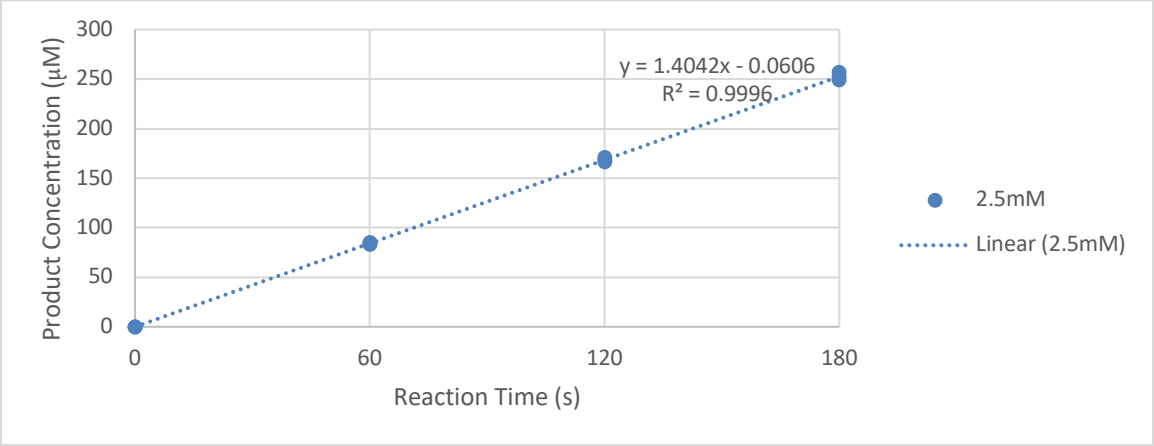


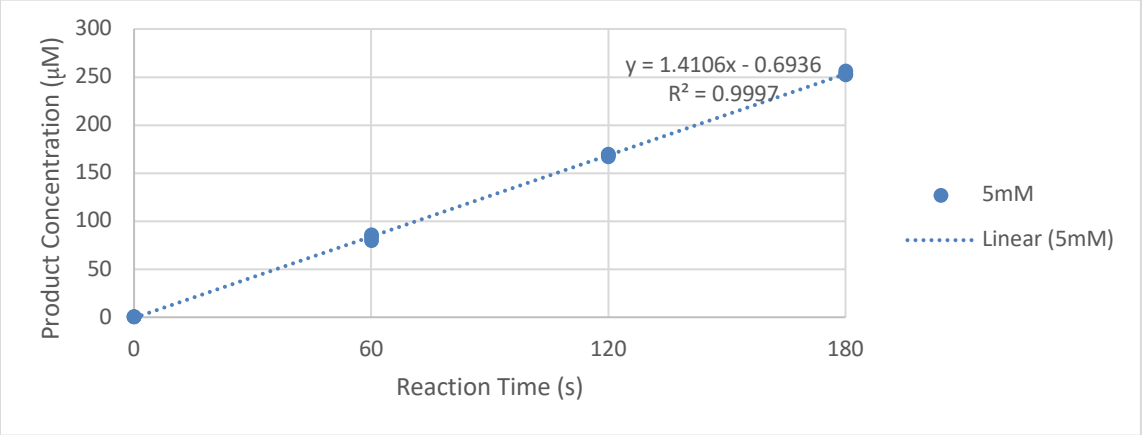
Final Variant (ParPgb-HYA-5213):

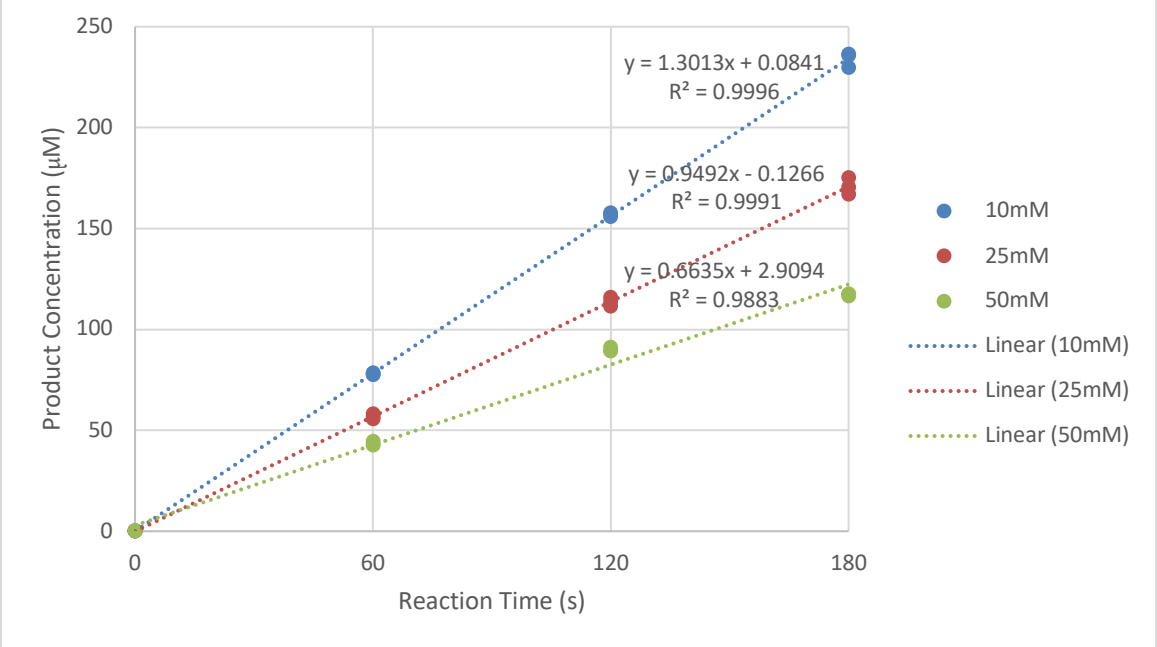
Initial-rate plots: (trend lines showing one of the three data sets)





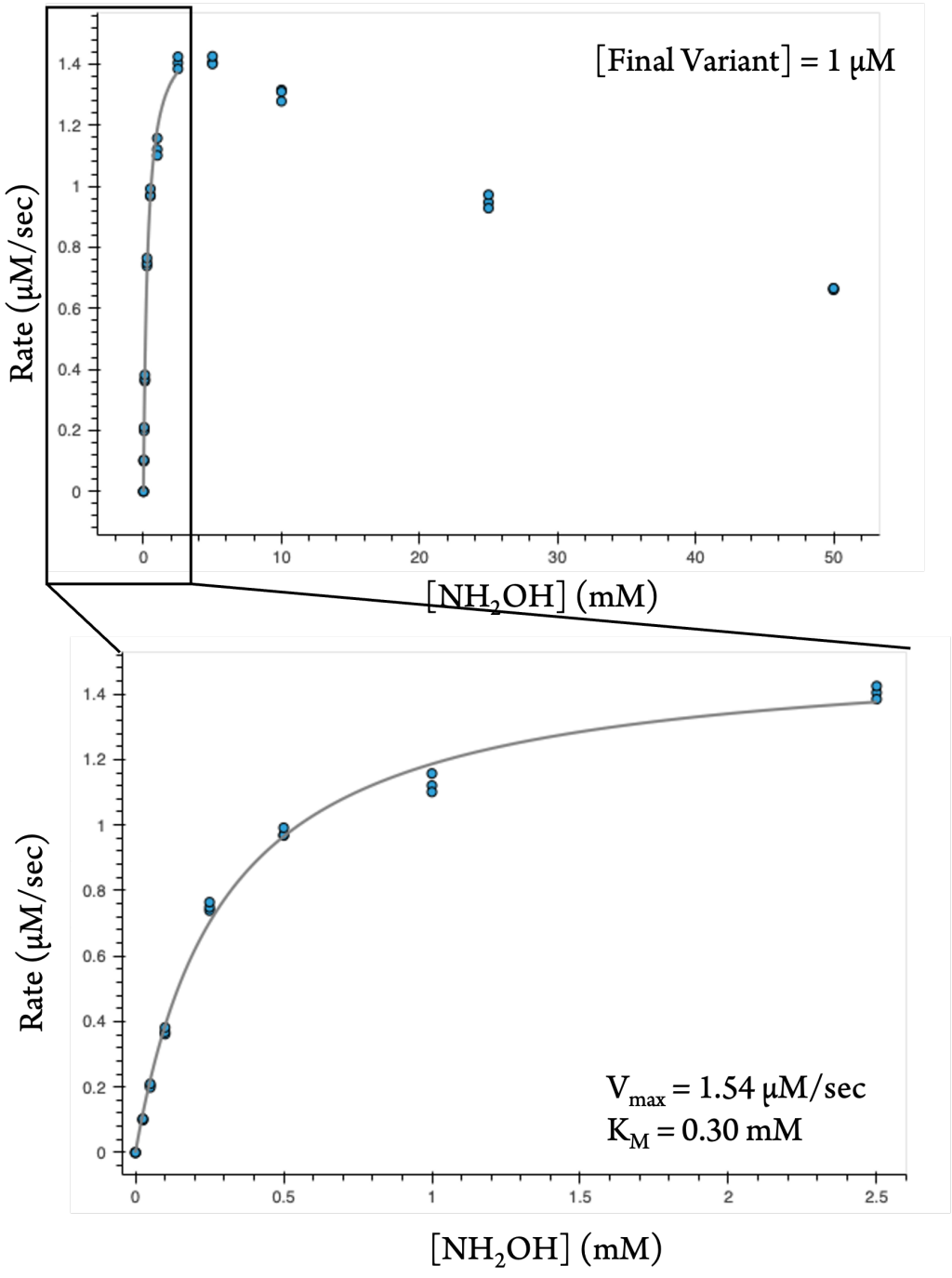

Product Concentration (µM)
300
250
200
150
100
50
0
0
60
120
180
Reaction Time (s)
y = 1.4042x - 0.0606
R² = 0.9996
2.5mM
Linear (2.5mM)


Product Concentration (µM)
300
250
200
150
100
50
0
0
60
120
180
Reaction Time (s)
y = 1.4106x - 0.6936
R² = 0.9997
5mM
Linear (5mM)


Product Concentration (µM)
250
200
150
100
50
0
0
60
120
180
Reaction Time (s)
y = 1.3013x + 0.0841
R² = 0.9996
y = 0.9492x - 0.1266
R² = 0.9991
y = 0.6635x + 2.9094
R² = 0.9883
10mM
25mM
50mM
Linear (10mM)
Linear (25mM)
Linear (50mM)

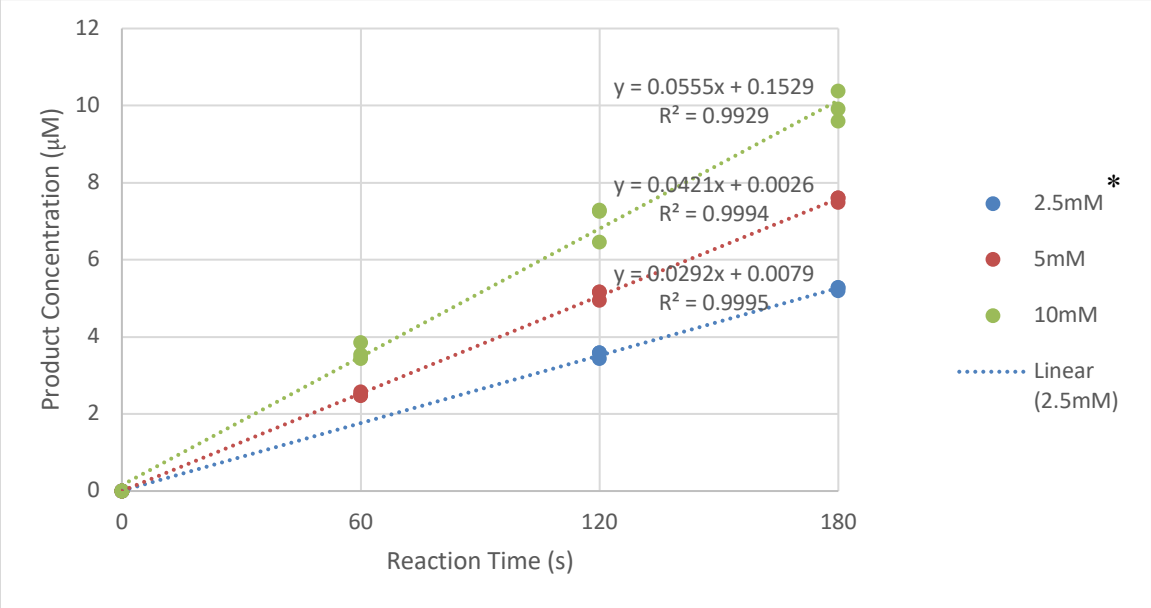
Michaelis-Menten Plots:

Decreased rates of reaction were observed after 5 mM of hydroxylamine. The first eight data points from 0.025 to 2.5 mM were used to fit the curve.

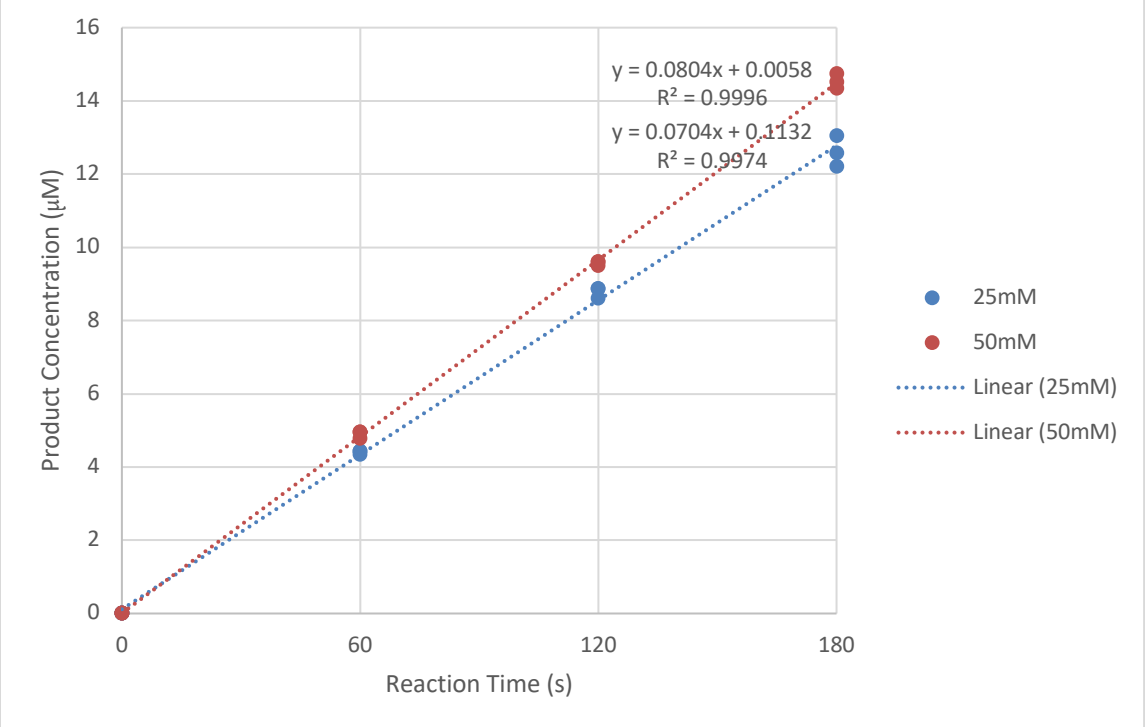


Initial Variant (ParPgb-HYA-5209):

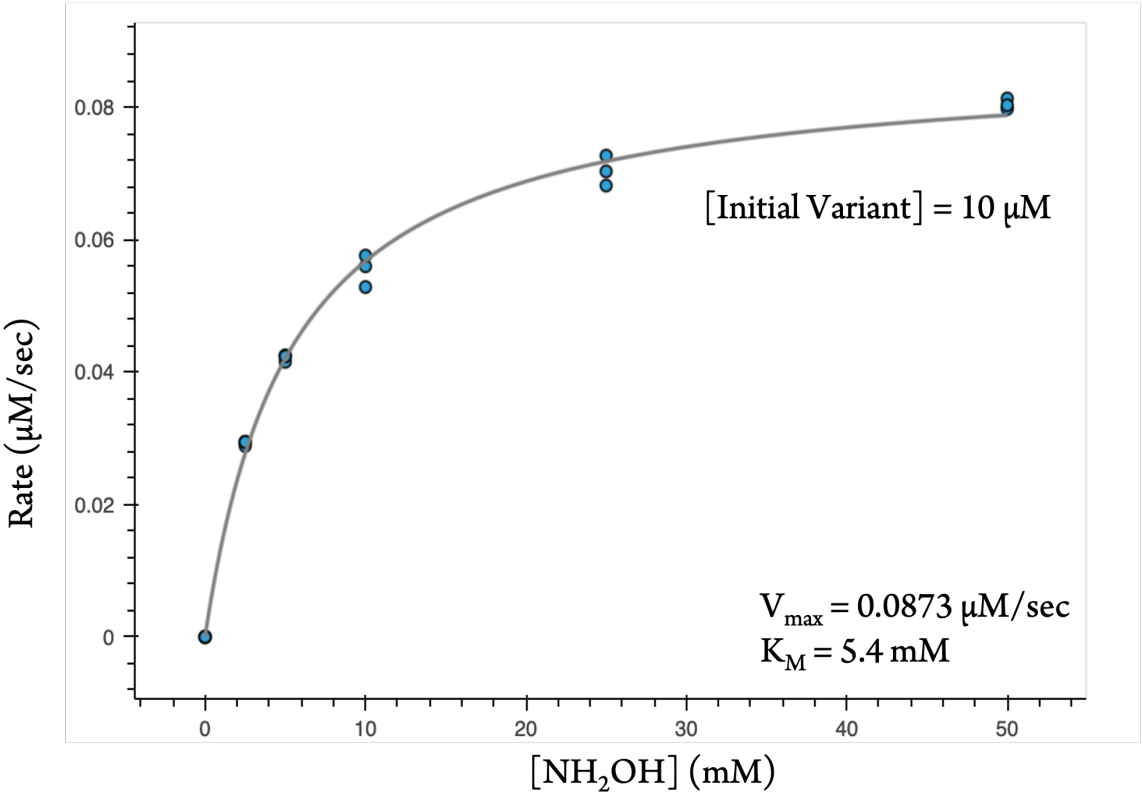
Initial-rate Plots: (trend lines showing one of the three data sets)



*Product formation with 2.5 mM can only be measured accurately at 120 and 180 seconds.



Michaelis-Menten Plots:



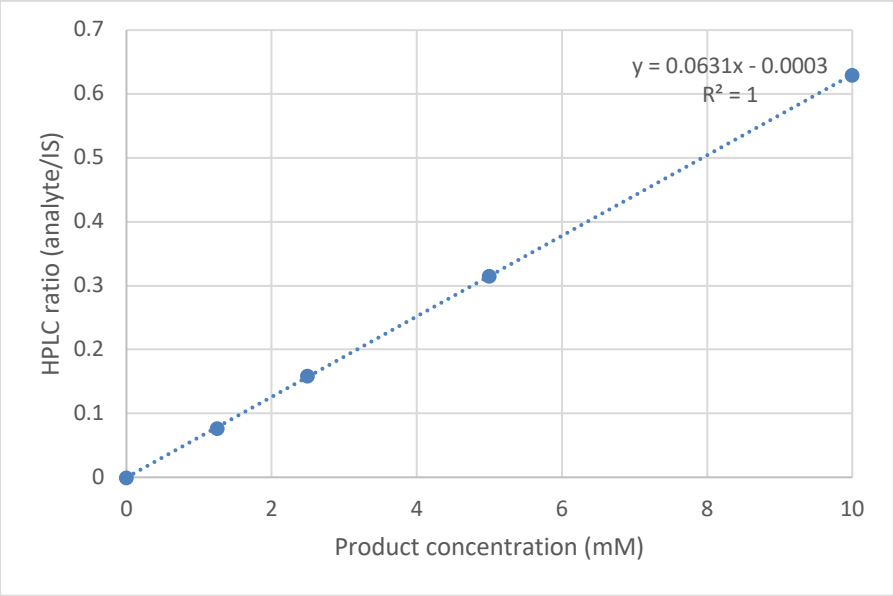
C.8. Kinetic Isotope Experiments.

C.8.1. Non-Competitive KIE Experiment

Experimental setup:

To each 2-mL screw cap vial, the substrate ethylbenzene **2c** or ethylbenzene-d_{10} **2c-*d*$_{10}$** (10 mM final concentration, 72 μL of 200 mM), sodium dithionite (10 mM final concentration, 72 μL of 200 mM), purified enzymes (10 μM final concentration for the final variant ParPgb-HYA-5213) were added. Hydroxylamine (30 mM final concentrations, 72 μL of 600 mM) was added (1440 μL total volume) to start the reaction. At 60, 120, 180, 300, and 600 seconds, 200 μL of the reaction mixture from each vial were added to a 1.7-mL Eppendorf tube containing 400 μL of ethanol with internal standard to quench the reactions. The reaction mixture was vortexed and centrifuged (14,000g, 15 minutes, 4 °C). The reaction mixture was analyzed by HPLC, and the product concentrations were calculated by their UV absorbance at 280 nm with a calibration curve. The initial rate was calculated by plotting the slope of the linear region.

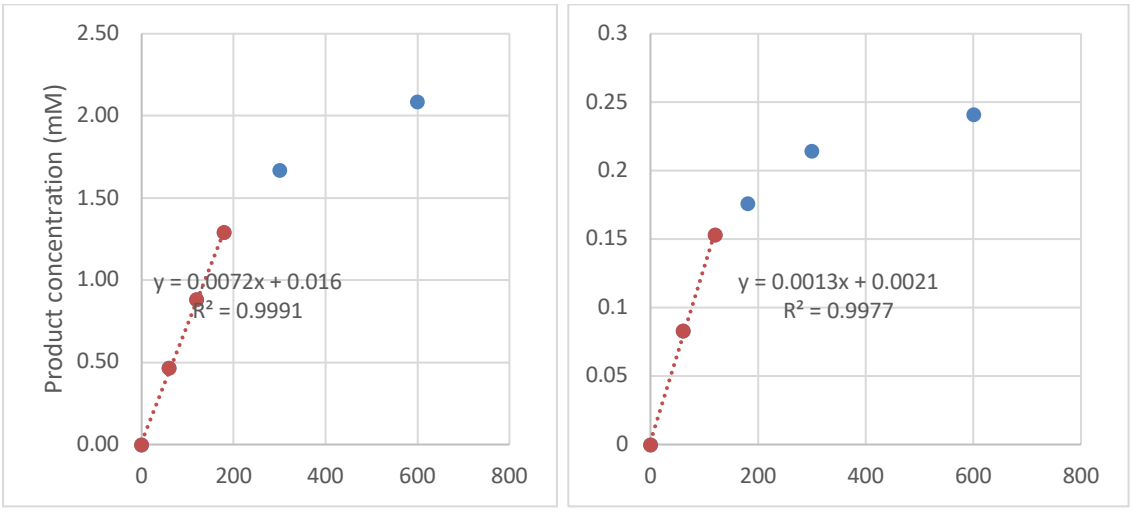
The samples for the standard curve were prepared following the procedure in Section 4.2., but quenching with 800 μL ethanol instead. Standard Curve based on UV absorbance at 280 nm for product **3c** and 1,3,5-trimethoxybenzene as internal standard.


y = 0.0631x - 0.0003
R² = 1
HPLC ratio (analyte/IS)
Product concentration (mM)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
2
4
6
8
10

Experiment 1:

2c				
Time (s)	Pdt	IS	Pdt/IS	[Pdt] (mM)
0	0			0
60	213.279	7340.457	0.029	0.465
120	402.516	7273.116	0.055	0.882
180	604.932	7450.661	0.081	1.291
300	759.832	7240.741	0.105	1.668
600	959.032	7300.152	0.131	2.087
2c-d_{10}				
Sec	Pdt	IS	Pdt/IS	[Pdt] (mM)
0	0			0
60	36.919	7339.516	0.005	0.083
120	69.450	7341.073	0.009	0.153
180	78.514	7209.663	0.011	0.176
300	95.558	7176.984	0.013	0.214
600	110.441	7362.640	0.015	0.241

k_H k_D

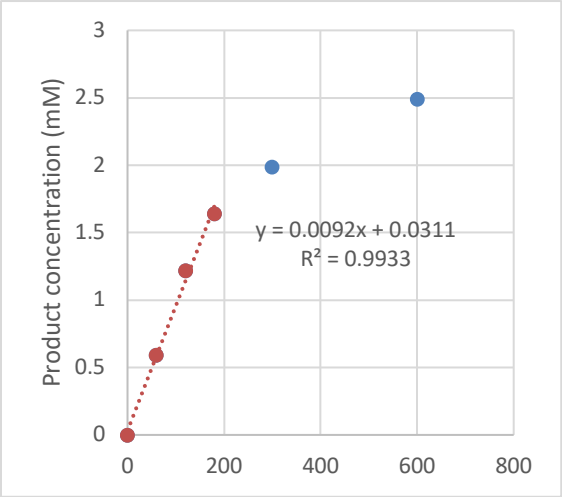


k_H/k_D = 5.61

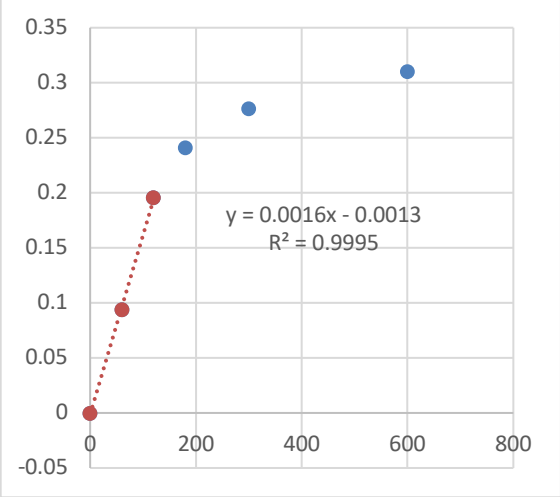
Experiment 2:

2c				
Time (s)	Pdt	IS	Pdt/IS	[Pdt] (mM)
0	0			0
60	265.355	7149.951	0.037	0.593
120	549.335	7161.32	0.077	1.220
180	750.471	7270.722	0.103	1.641
300	911.445	7286.688	0.125	1.987
600	1163.401	7418.855	0.157	2.490
2c-d_{10}				
Sec	Pdt	IS	Pdt/IS	[Pdt] (mM)
0	0			0
60	41.985	7469.188	0.006	0.094
120	89.738	7458.920	0.012	0.195
180	112.274	7522.618	0.015	0.241
300	128.686	7502.765	0.017	0.277
600	144.906	7517.056	0.019	0.310

k_H



k_D

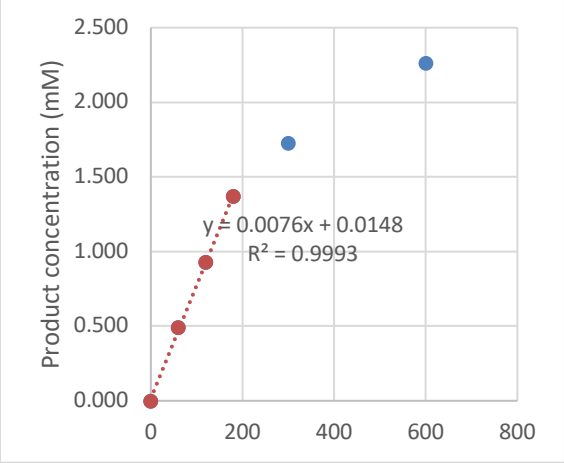


k_H/k_D = 5.68

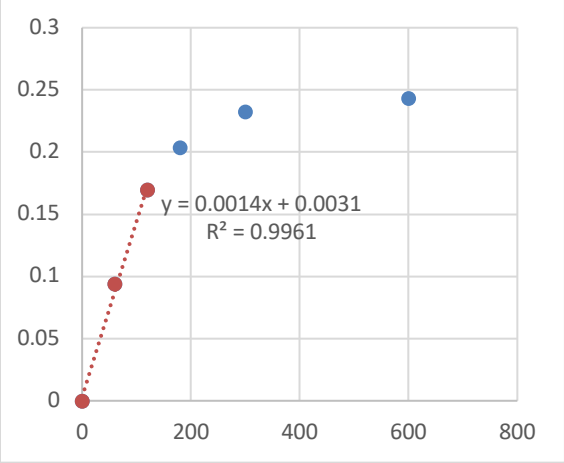
Experiment 3:

2c				
Time (s)	Pdt	IS	Pdt/IS	[Pdt] (mM)
0	0			0
60	235.484	7692.58	0.031	0.490
120	459.298	7871.999	0.058	0.929
180	687.72	7981.114	0.086	1.370
300	883.752	8140.77	0.109	1.725
600	1171.742	8228.37	0.142	2.262
2c-d_{10}				
Sec	Pdt	IS	Pdt/IS	[Pdt] (mM)
0	0			0
60	52.712	9371.191	0.006	0.094
120	96.964	9336.252	0.010	0.169
180	117.094	9335.572	0.013	0.204
300	134.463	9371.853	0.014	0.232
600	139.453	9272.778	0.015	0.243

k_H



k_D



k_H/k_D = 5.49

Average k_H/k_D = 5.6 ± 0.10

C.8.2. Competitive KIE Experiment

Experimental setup:

To each 2-mL screw cap vial, purified enzymes (50 μM final concentration for the final variant ParPgb-HYA-5213), the substrate ethylbenzene **2c** and ethylbenzene-d_{10} **2c-*d*$_{10}$** (10 mM final concentration for each in 1:1 ratio, 10 μL of 400 mM each), and sodium dithionite (10 mM final concentration, 20 μL of 200 mM) were added. Finally, hydroxylamine (2.5 mM final concentrations, 20 μL of 50 mM) as the limiting reagent was added (400 μL total reaction volume). After the reaction was shaken at room temperature for 12 hours at 200 rpm, 400 μL of the reaction mixture from each vial were added to a 1.7-mL Eppendorf tube containing 800 μL of acetonitrile to quench the reactions. The reaction mixture was vortexed, centrifuged (14,000g, 15 minutes, 4 °C), and concentrated. The supernatant was analyzed by LCMS with two separate single ion channels 105 m/z and 114 m/z. The experiment was repeated three times and the average product ratio (P_H/P_D) by ion account was calculated.

Experiment data:

Experiment	P_H	P_D	P_H/P_D
1	7162.807	800.275	8.95
2	7396.508	844.091	8.76
3	7375.189	855.696	8.62

Average $P_H/P_D = 8.8 \pm 0.14$

C.9. Miscellaneous Experiments.

C.9.1. Reaction Condition Optimization

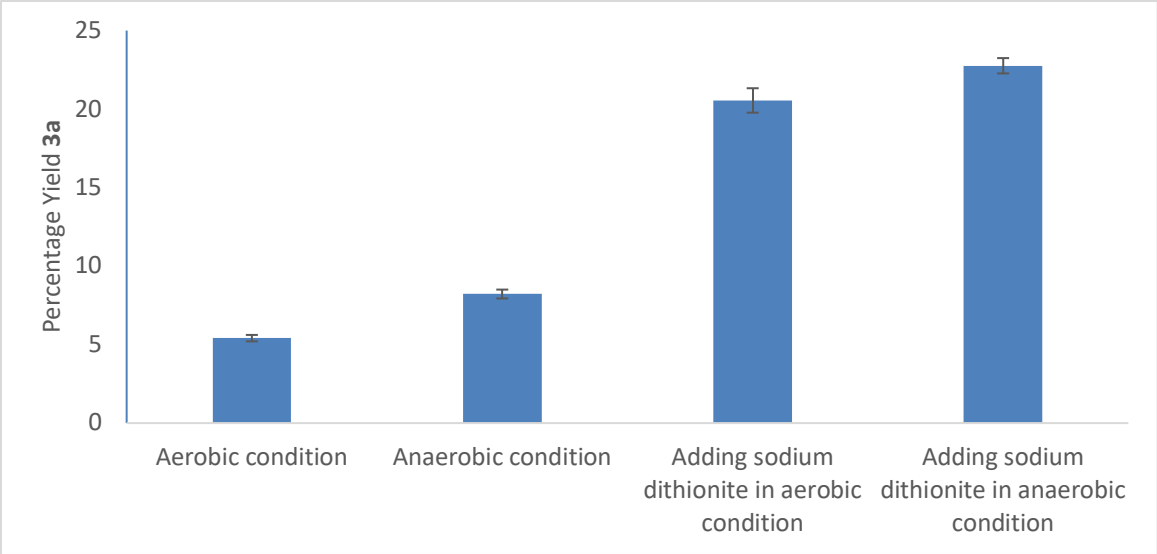


Figure C-5. Condition optimization with additional reductant and oxygen sensitivity using 10 μM purified enzyme, 10 mM of **2a**, and 30 mM of hydroxylamine in Kp buffer pH 8.0.

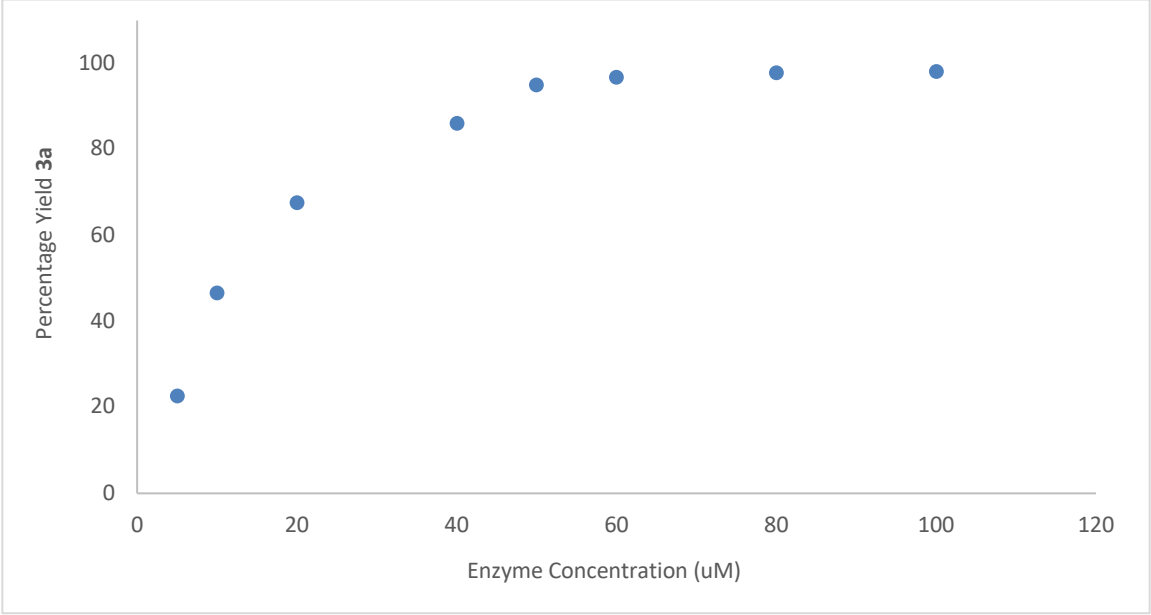


Figure C-6. Enzyme concentration vs. percentage yield of **3a**. Reactions were performed using 10 mM of sodium dithionite, 10 mM of **2a**, and 30 of mM hydroxylamine in Kpi buffer pH 8.0.

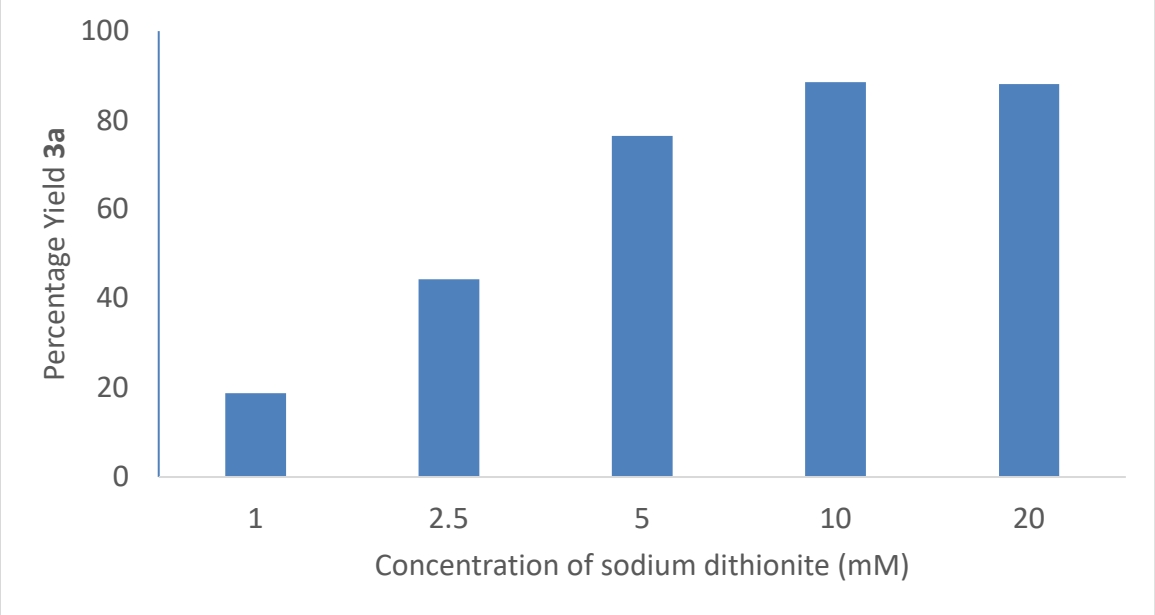


Figure C-7. Concentration of sodium dithionite vs. percentage yield of **3a**. Reactions were performed using 50 µM purified enzyme, 10 mM of **2a**, and 30 mM of hydroxylamine in Kpi buffer pH 8.0.

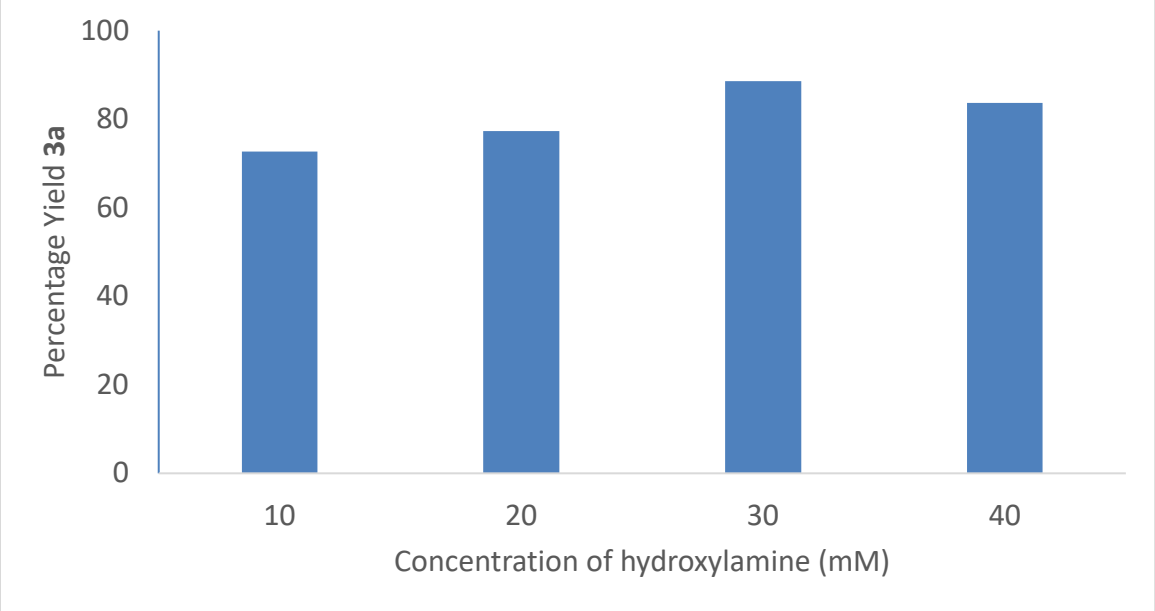


Figure C-8. Concentration of hydroxylamine vs. percentage yield of **3a**. Reactions were performed using 50 µM purified enzyme, 10 mM of **2a**, and 30 mM of sodium dithionite in Kpi buffer pH 8.0.

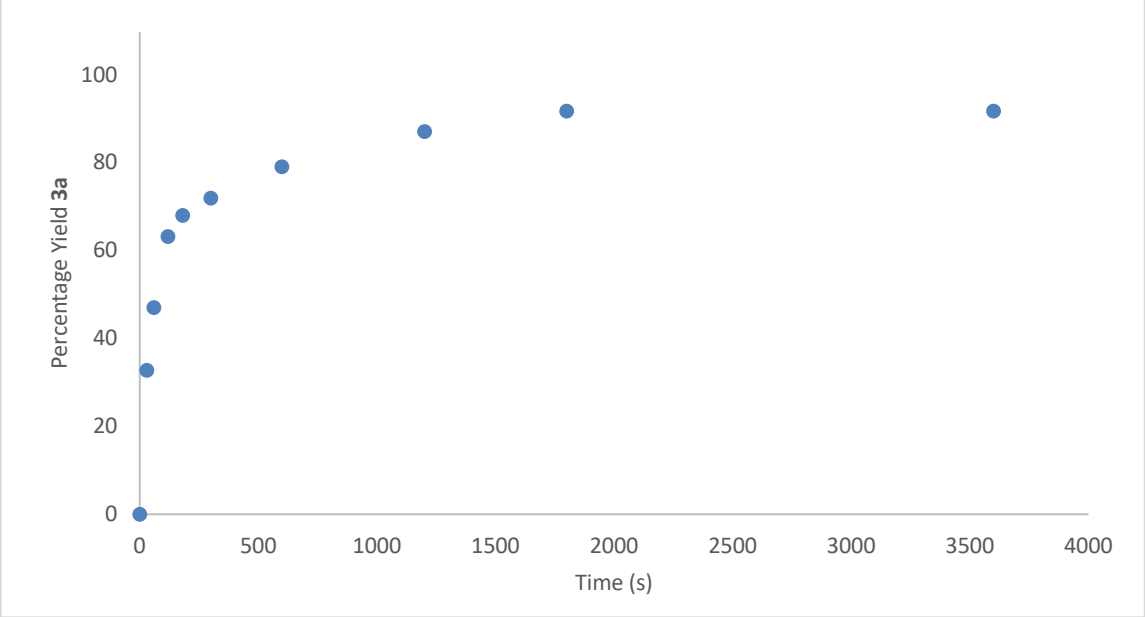


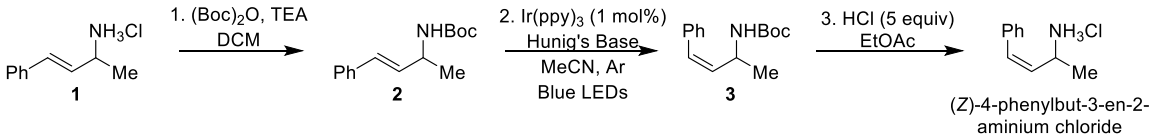
Figure C-9. Time course of enzymatic reaction using 50 μM purified enzyme, 10 M sodium dithionite, 10 mM **2a**, and 20 mM **1** in Kpi buffer pH 8.0.

C.9.2. Reaction with (*E/Z*)-1-phenyl-1-butene

Preparation of E/Z Substrates and E/Z Standard Product:

The substrates (*E*)-but-1-en-1-ylbenzene **(*E*)-6** and (*Z*)-but-1-en-1-ylbenzene **(*Z*)-6** were synthesized according to reported literature procedures.[11] ^{1}H NMR analysis agreed with known literature values. Standard product (*E*)-4-phenylbut-3-en-2-aminium chloride **(*E*)-7** was obtained from commercial suppliers and used without further purification.

Synthesis of (Z)-4-phenylbut-3-en-2-aminium chloride **(*Z*)-7**:



A 10-mL vial with a magnetic stirring bar was charged with compound **1** (223 mg, 1.21 mmol, 1 equiv.) and DCM (6 mL, 0.2 M). Thereafter, triethylamine (340 µL, 2.42 mmol, 2 equiv) and Boc anhydride (400 mg, 1.81 mmol, 1.5 equiv.) were added in that order. The reaction was left to stir overnight or until full consumption of the starting material was observed via TLC analysis. The reaction was reduced *in vacuo* and purified via flash chromatography (Hex to 10% EtOAc), affording compound **2** as a white solid (285 mg, 95% Y). ^{1}H NMR analysis agreed with known literature values.[12]

A 10-mL culture-tube with a magnetic stirring bar was charged with compound **2** (100 mg, 0.4 mmol, 1 equiv.), $Ir(ppy)_3$ (3 mg, 0.004 mmol, 0.01 equiv.), MeCN (8 mL, 0.05 M), and Hunig's base (7 µL, 0.004 mmol, 0.1 equiv.). The resultant mixture was sealed, degassed, by sparging it with argon for 5 minutes, and left adapted with an argon balloon. The

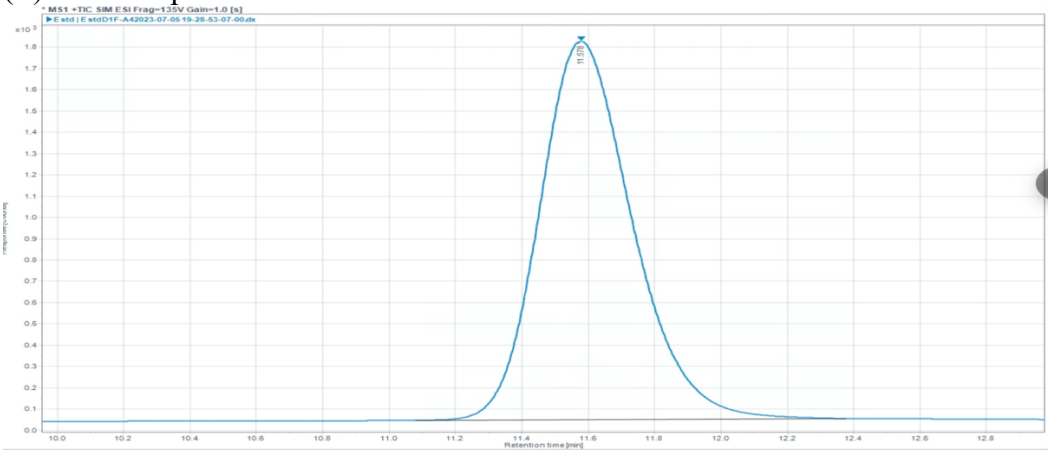
reaction was irradiated with a Kessil PR160L 440 nm lamp for 24 hours. The reaction was reduced *in vacuo* and purified via flash chromatography (Hex to 10% EtOAc), affording compound **3** as a white solid (~90 mg, 90% Y, 10:1 Z/E ratio). ^{1}H NMR analysis agreed with known literature values.[11]

A 10-mL vial with a magnetic stirring was charged with compound **3** (40 mg, 0.15 mmol, 1 equiv.) and dissolved in ethyl acetate (2 mL, 0.075 M). Thereafter, HCl (63 µL of 12 M HCl, 5 equiv.) was added and the reaction was left to stir until full consumption of the starting material was observed via TLC. The reaction was reduced *in vacuo*, and the resultant residue was dissolved in 3 mL of toluene. The resultant solution was reduced *in vacuo* affording a white solid. The obtained solid was washed with two portions of diethyl ether (2 x 1 mL) and dried under high vacuum, affording a white solid (~40 mg, 80% Y, 8:1 Z/E ratio). ^{1}H NMR analysis agreed with known literature values.[13] **^{1}H NMR** (400 MHz, DMSO) δ 8.30 (s, 3H), 7.49 – 7.21 (m, 5H), 6.63 (d, J = 11.6 Hz, 1H), 5.68 (dd, J = 11.7, 9.9 Hz, 1H), 4.12 (dq, J = 9.5, 6.5 Hz, 1H), 1.33 (d, J = 6.6 Hz, 3H).

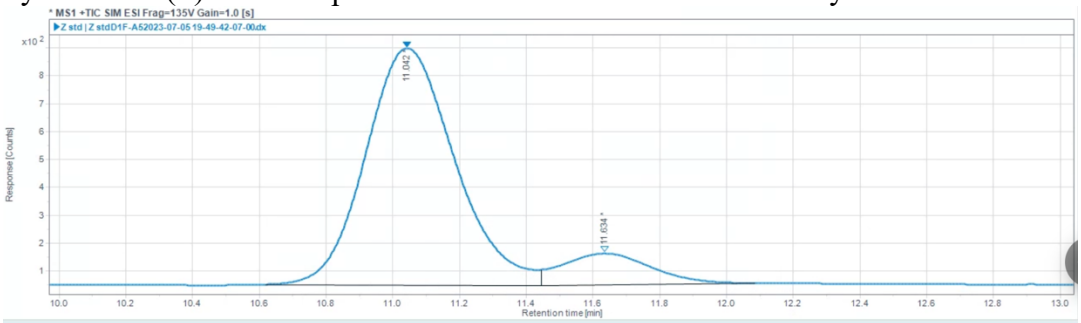
LCMS single ion channel trace:

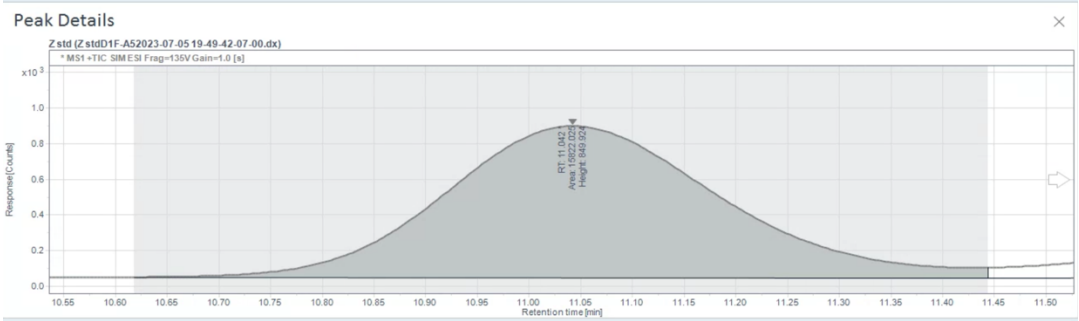
HPLC conditions: C18 (Poroshell, EC-C18, 4.6 × 100 mm, 5 µm) column, 5–90% acetonitrile in water (0.1% acetic acid), 15 minutes, 1.5 mL/min. The products were detected by mass spectrum in 131 m/z single ion channel.

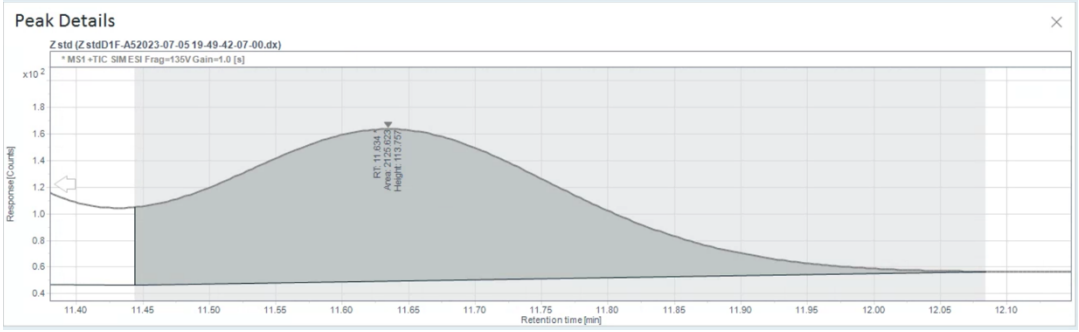
(*E*)-standard product from commercial source



Synthesized (*Z*)-standard product with 8:1 Z/E ratio confirmed by ^{1}H NMR.

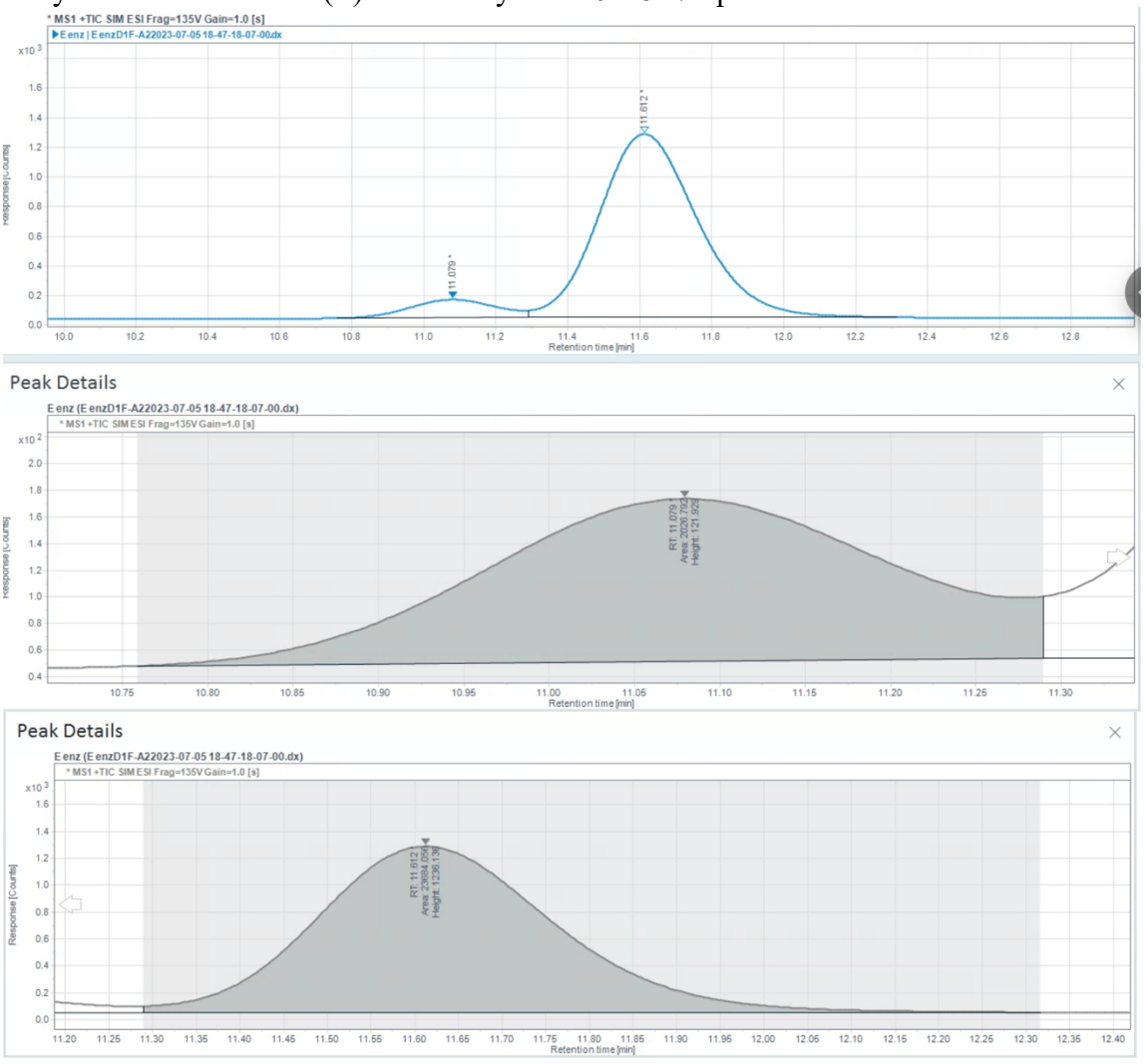




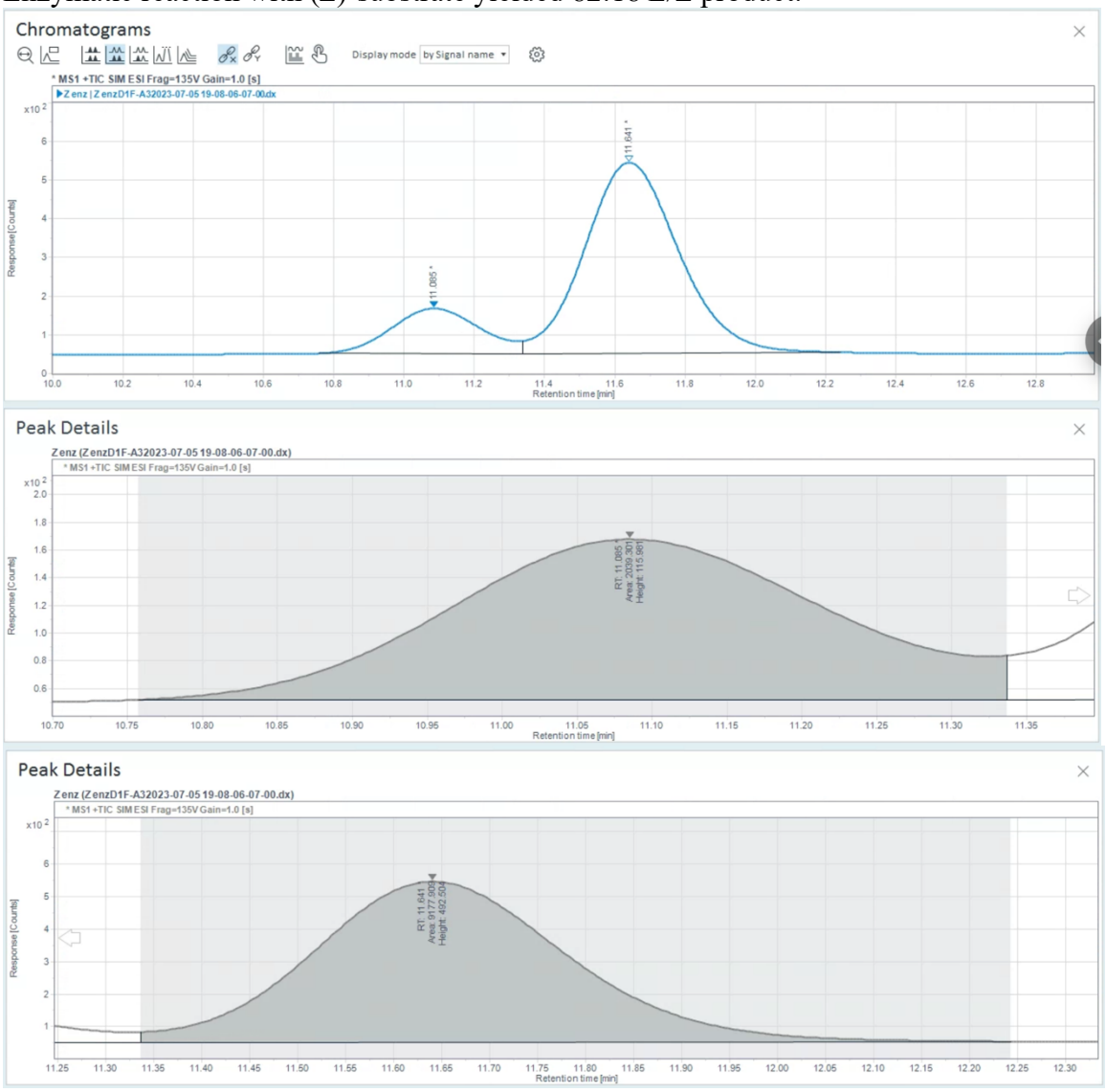


The peak area ratio in the LCMS single ion channel for this mixture is also 8:1. Thus, the peak area ratio can accurately represent the real concentration ratio of the isomers.

Enzymatic reaction with (*E*)-substrate yielded 92:8 *E/Z* product.

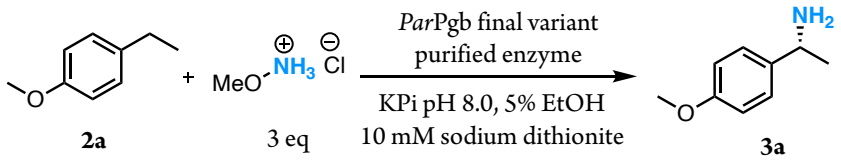


Enzymatic reaction with (*Z*)-substrate yielded 82:18 *E/Z* product.



C.9.3. Reaction with *O*-methylhydroxylamine

The reactions were set up following the procedure in Section 4.1. The reactions were analyzed following the procedure in Section 4.2. Standard curve in Section 4.3. product **3a** was used to calculated [Pdt] (mM)



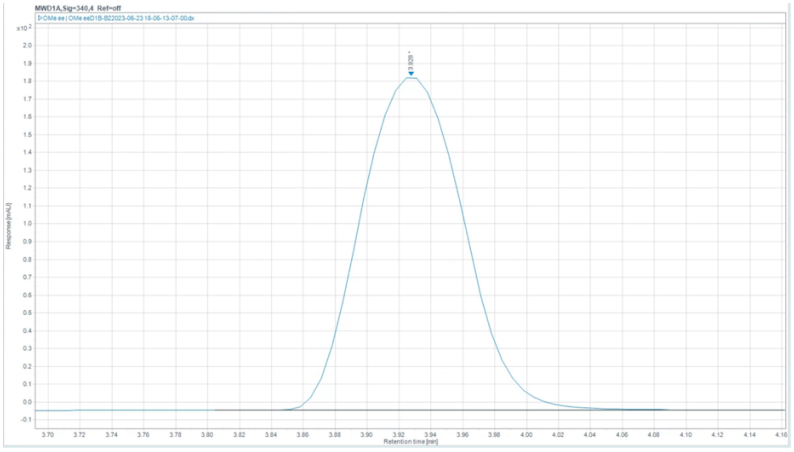
2a 3 eq **3a**

*1-(4-methoxyphenyl)ethan-1-amine (**3a**)*

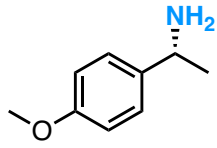
Pdt	IS	Pdt/IS	[Pdt] (mM)	yield %	TTN	Average yield	SD yield	Average TTN	SD TTN
599.09	116.45	5.14	8.25	82.5	165				
598.43	117.45	5.10	8.17	81.7	163				
596.49	118.65	5.03	8.06	80.6	161	81.6	0.77	163	1.54

Notes: Pdt = product area, IS = internal standard area, [Pdt] = product concentration in reaction, [PC] = protein concentration in reaction, Avg yield = average of percentage yield, SD yield = standard deviation of percentage yield, Avg TTN = average of total turnover number, SD TTN = standard deviation of TTN.

The sample for enantioselectivity determination was prepared following the procedure in Section 5.1. **3a-enz** – reverse phase – enzymatic reaction >99% *ee*

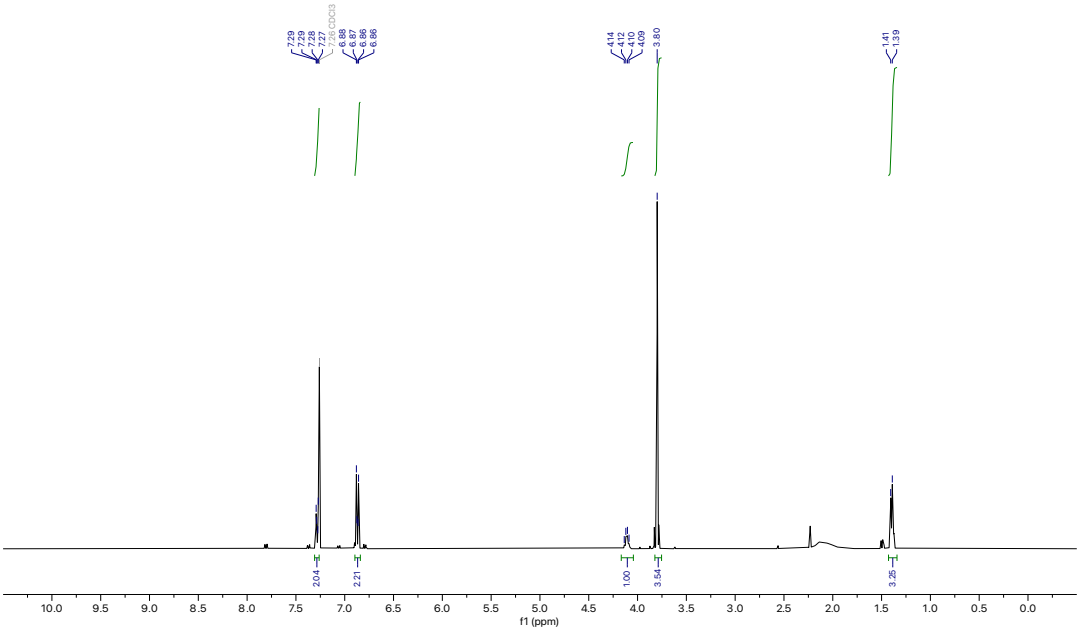


C.10. NMR Spectra of Synthesized Compounds and Enzymatic Product.

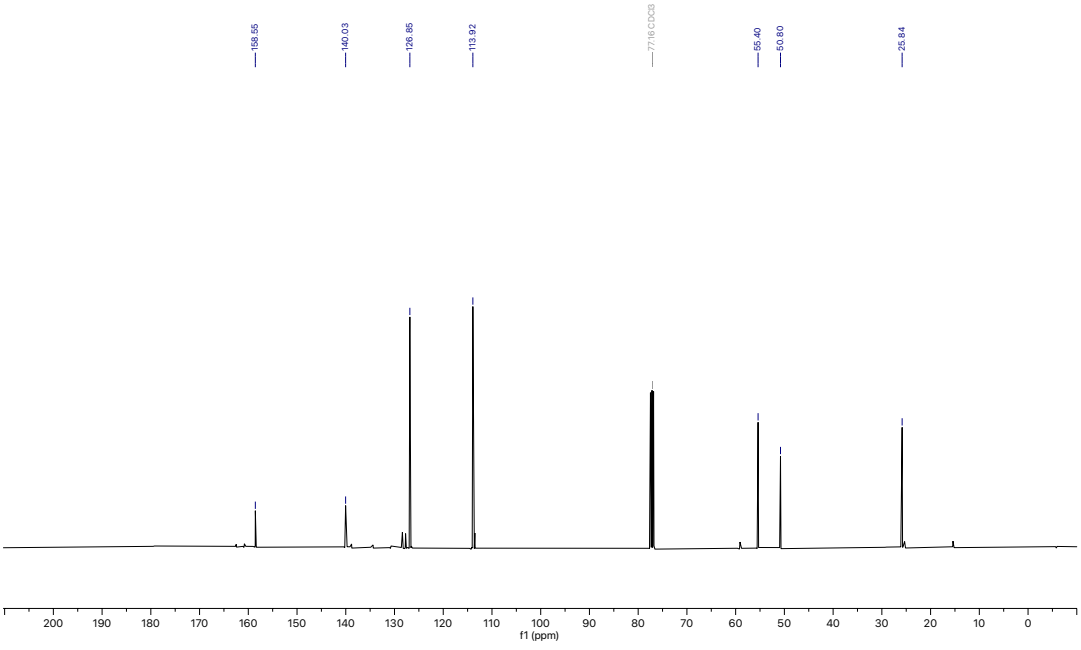


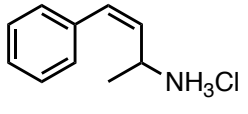
3a – enzymatic product

^{1}H NMR (400MHz, $CDCl_3$)



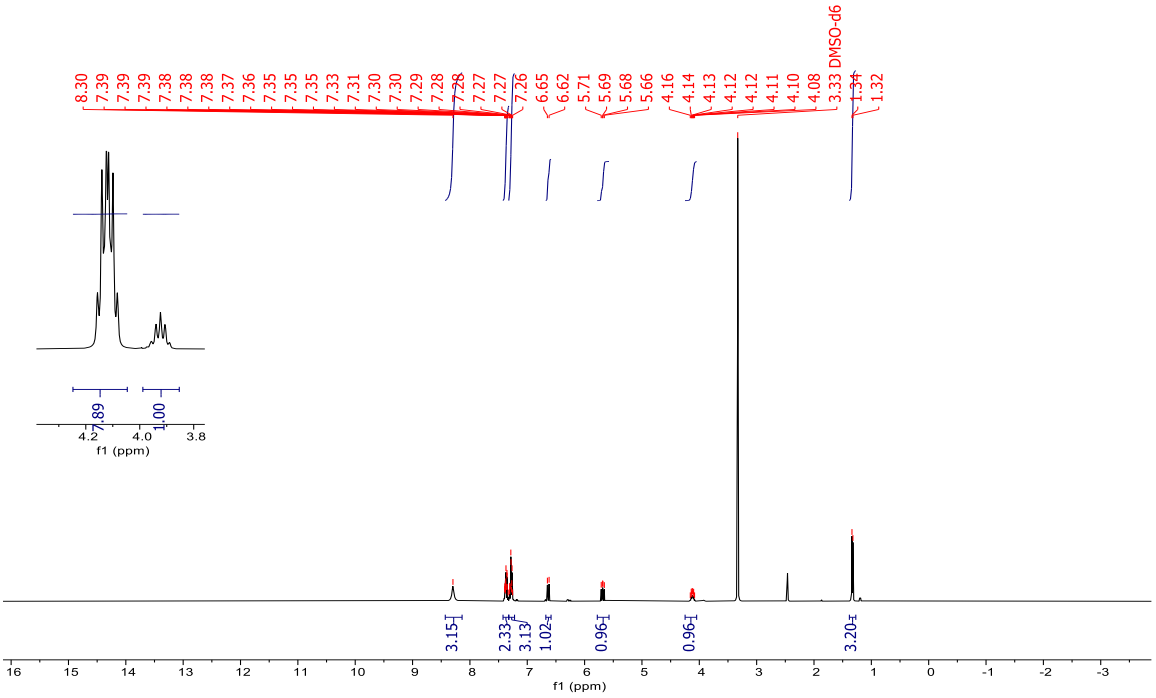
^{13}C NMR (101MHz, $CDCl_3$)





(*Z*)-7

^{1}H NMR (400MHz, DMSO-d$_6$)



C.11. References

1. 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* (2), 83–92.
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* (5), 343–345.
3. Boville, C. E.; Scheele, R. A.; Koch, P.; Brinkmann-Chen, S.; Buller, A. R.; Arnold, F. H. Engineered Biosynthesis of β-Alkyl Tryptophan Analogues. *Angew. Chem. Int. Ed.* **2018**, *57* (45), 14764–14768.
4. Zhao, H.; Giver, L.; Shao, Z.; Affholter, J. A.; Arnold, F. H. Molecular Evolution by Staggered Extension Process (StEP) in Vitro Recombination. *Nat. Biotechnol.* **1998**, *16* (3), 258–261.
5. Barr, I.; Guo, F. Pyridine Hemochromagen Assay for Determining the Concentration of Heme in Purified Protein Solutions. *BIO-PROTOCOL* **2015**, *5* (18).
6. Lewis, J.C.; Bastian, S.; Bennett, C.S.; Fu, Y.; Mitsuda, Y.; Chen, M.M.; Greenberg, W.A.; Wong, C.H. and Arnold, F.H. Lewis, J. C.; Bastian, S.; Bennett, C. S.; Fu, Y.; Mitsuda, Y.; Chen, M. M.; Greenberg, W.A.; Wong, C.H.; Arnold, F. H. Chemoenzymatic elaboration of monosaccharides using engineered cytochrome P450BM3 demethylases. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*(39), 16550-16555.
7. Jia, Z.-J.; Gao, S.; Arnold, F. H. Enzymatic Primary Amination of Benzylic and Allylic C(Sp^3)–H Bonds. *J. Am. Chem. Soc.* **2020**, *142* (23), 10279–10283.
8. Knight, A. M.; Kan, S. J.; Lewis, R. D.; Brandenberg, O. F.; Chen, K.; Arnold, F. H. Diverse engineered heme proteins enable stereodivergent cyclopropanation of unactivated alkenes. *ACS Cent. Sci.* **2018**, *4*(3), 372-377.
9. Cho, I.; Prier, C. K.; Jia, Z. J.; Zhang, R. K.; Görbe, T.; Arnold, F. H. Enantioselective aminohydroxylation of styrenyl olefins catalyzed by an engineered hemoprotein. *Angew. Chem. Int. Ed* **2019**, *58*(10), 3138-3142.
10. Sigmund, M. C.; Poelarends, G. J. Current state and future perspectives of engineered and artificial peroxygenases for the oxyfunctionalization of organic molecules. *Nat. Cat.* **2020**, *3*(9), 690-702.
11. Van Der Waals, D.; Pettman, A.; Williams, J. M. Copper-catalysed reductive amination of nitriles and organic-group reductions using dimethylamine borane. *RSC Adv.* **2014**, 4(94), 51845-51949.
12. Mastandrea, M. M.; Cañellas, S.; Caldentey, X.; Pericàs, M. A. Decarboxylative Hydroalkylation of Alkynes via Dual Copper-Photoredox Catalysis. *ACS Catal.* **2020**, *10* (11), 6402–6408.
13. Hu, Y.; Zou, Y.; Yang, H.; Ji, H.; Jin, Y.; Zhang, Z.; Liu, Y.; Zhang, W. Precise Synthesis of Chiral Z -Allylamides by Cobalt-Catalyzed Asymmetric Sequential Hydrogenations. *Angew. Chemie Int. Ed.* **2023**, *62* (15).