

---

**Production of Lignin-Based Aromatic Amino Acids with  
Computational Designed Phenylalanine Aminomutase  
from *Taxus chinensis***

Zur Erlangung des akademischen Grades eines

DOKTORS DER INGENIEURWISSENSCHAFTEN (DR. -ING.)

von der KIT-Fakultät für Chemieingenieurwesen und Verfahrenstechnik des

Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

M. Sc. Fei Peng

aus China

Tag der mündlichen Prüfung: 07.04.2022

Erstgutachter: Dr. rer. nat. Jens Rudat

Zweitgutachter: Prof. Dr. Matthias Franzreb

# ERKLÄRUNG

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig angefertigt und keine anderen als die darin angegebenen Quellen und Hilfsmittel benutzt, sowie die wörtlich oder inhaltlich übernommenen Stellen als solche kenntlich gemacht und die Satzung des Karlsruher Institut für Technologie (KIT) zur Sicherung guter wissenschaftlicher Praxis in der gültigen Fassung beachtet habe. Zudem versichere ich, dass die elektronische Version dieser Arbeit mit der schriftlichen übereinstimmt.

Ich versichere außerdem, dass ich diese Dissertation nur in diesem und keinem anderen Promotionsverfahren eingereicht habe und dass diesem Promotionsverfahren keine endgültig gescheiterten Promotionsverfahren vorausgegangen sind.

-----

Ort, Datum

-----

Unterschrift

---

## ACKNOWLEDGEMENT

This thesis not only represents a life-changing experience for me but also is filled with warm and lovely memories at Tebi. It would not be possible to finish this work without the help and support from my family, friends, and colleagues.

I sincerely thank Prof. Dr. Christoph Syldatk to allow me to research the interesting thesis in Tebi at Karlsruher Institut für Technologie. He is a very kind professor with domain knowledge and a great role model for me and our colleagues.

I would like to deeply thank my supervisor, Dr. Jens Rudat. He is an enlightened scientist and supported fresh ideas about the thesis. Without his guidance, support, and constant feedback this work would never be achievable. I bet all of the students are enjoying the course which he taught since he is a great teacher with professional knowledge and a sense of humor. To me, he is a “beer-type” supervisor and makes me more creative with his novel perspective.

I would like to give my gratitude to Dr. Ulrike Engel, who gave me much supports in my experiments about molecular biology. On the other side, she has strong organizational skills and attention to detail. I learned not only the tips of experimental methods but also her serious attitude on works from her.

A special thanks to Dr. Habibu Aliyu, who is always an optimistic and rational person. With the huge knowledge of bioinformatics and the motivation of science, he gives me many commends on both my work and life.

I sincerely thank Prof. Dr. Stefan Bräse for his helpful feedback and collaboration in this project. I would like to thank members of the institute for organic Chemie at Karlsruher institute for technology (IOC-KIT), especially my cooperation partner, Dr. Zhen Zhang. They helped me a lot in the field of chemical synthesis.

---

I would like to thank Dr. Kersten Rabe for the collaboration and scientific discussion. I learned a lot from him in the enzyme design and mutagenesis field.

I want to thank Prof. Matthias Franzreb for his friendly discussion and feedback during the TAC meetings.

I gratefully acknowledge the funding received towards my Ph.D. from China Scholarship Council with bioeconomy graduate program BBW ForWerts of the German state of Baden-Württemberg. Thanks to the people in the education department of the embassy of the people's republic of China in Frankfurt. They gave me full support for the application and the working life in Karlsruhe, especially during COVID-19.

I am also grateful to Dr. André Delavault for his useful discussion during the group meetings, publications, and writing this thesis.

My deep appreciation goes out to all family-like staff members at Tebi. Susanne and Beate, are the best two secretaries in the world and always lend a hand to me when I have problems. Pascal, the sweet-toothed HPLC expert, provided me with his assistance throughout my project. Michaela, thanks for the conversation and many laughs.

A tremendous thank is to address to my lovely colleagues: Alba, Carolin, Olga, Rebecca, Teresa, Aline, Christin, Flávius, Marcus, Vanessa, Jens, Oliver, Sascha, Stepfan, Alberto, Nebyat, Magda, especially, the members in "girl's office". Thank you for your friendship, support, and help.

I want to also thank Beatrix, Sabrina, and Robin for their student support of the project.

I thank from the bottom of my heart to my family and friends for their love and encouragement. Their unwavering support and trust give me the courage to make my dream come true. Xinyan Yu and Chao Zhang, thank you for giving me tremendous

---

encouragement when I hesitated to take the position for promotion. Xinshuang Ge, I learned a lot of tips about molecular biologic experiments to improve my project. Jingyi Liu, thanks for all the enthusiastic discussion with me about protein purification and structure studies. Hui Li, my best friend since high school, is always by my side and gives me support. Finally, I would like to thank my parents for their believing in and understanding me and encouraging me to challenge myself.

---

## LIST OF PUBLICATIONS

### Peer-reviewed original publications

#### **Computational-Designed Enzyme for $\beta$ -Tyrosine Production in Lignin Valorization.**

**Peng, Fei.,** Aliyu, Habibu., Delavault, André., Engel, Ulrike., Rudat, Jans.

*Catalysts* 11.11 (2021): 1310.

[DOI: 10.3390/catal11111310](https://doi.org/10.3390/catal11111310)

#### **Origin and Evolution of Enzymes with MIO Prosthetic Group: Microbial Coevolution After the Mass Extinction Event.**

**Peng, Fei.,** Engel, Ulrike., Aliyu, Habibu., Rudat, Jans.

Submitted to: *Frontiers in Genetics*.

#### **Production of (*S*)- and (*R*)- $\beta$ -tyrosine by redesigned phenylalanine aminomutase.**

**Peng, Fei.,** Aliyu, Habibu., Delavault, André., Engel, Ulrike., Rudat, Jans.

Submitted to: *ChemCatChem*.

### Conference Poster

#### **Lignin-based Aromatic Amino Acids (LIGAROM).**

**Peng, Fei.,** Aliyu, Habibu., Engel, Ulrike., Rudat, Jans.

*Biotrans* (2021), Graz, Austria.

#### **Lignin-based Aromatic Amino Acids (LIGAROM).**

**Peng, Fei.,** Aliyu, Habibu., Engel, Ulrike., Rudat, Jans.

*3<sup>rd</sup> International Bioeconomy Congress Baden-Württemberg (2020), Stuttgart, Germany.*

---

## ABSTRACT

Over the past half-century, the growing energy demand and increased consumption of fossil resources have caused various environmental problems, such as smog, acid rain, and depleting ozonosphere. On the other hand, the overconsumption of non-renewable coal, petroleum, and natural gas will lead to natural resource depletion. To solve these global problems, lignocellulosic biomass has been developed as an alternative. In contrast with the carbohydrate fraction of lignocellulose, lignin is underutilized due to its complexity of structure. Although lignin valorization is an attractive topic right now, the construction of an efficient conversion of lignin to fine chemicals remains a huge challenge.

Optically pure  $\beta$ -amino acids have been considered as fundamental building blocks in bioactive products. Considering the antibacterial and antitumor properties, they have gained much attention from researchers in the pharmaceutical and agrochemical field. One clinical and commercially successful example is paclitaxel (Taxol®), which was isolated from *Taxus* species for the first time in the year 1991. The paclitaxel is an antimicrotubule drug and can use for the treatment of various leukemias and solid tumors. With the growing mortality rates of cancers, the associated socioeconomic burden is a huge challenge for several developing countries. Therefore, the establishment of an effective strategy for the generation of bioactive products with the antitumor feature is an attractive topic to solve these problems.

The topic of this thesis is Lignin-based Aromatic Amino Acids (LIGAROM), which is embedded in BBW ForWerts Graduate Program of the State of Baden-Württemberg with support from the China Scholarship Council (CSC). The aim of this work is the construction of an efficient strategy to valorize the lignin monomer: *p*-hydroxycinnamic acid exemplary for underutilized sustainable aromatic resources. To achieve the main goal of this thesis, a chemoenzymatic cascade and an enzymatic conversion with a computationally designed enzyme were performed for the production of  $\beta$ -tyrosine

---

from *p*-hydroxycinnamic acid. Tyrosine ammonia-lyase (TAL) with aminomutase activity was selected as the template enzyme at the beginning. However, the enantioselectivity of bacterial TALs is too low to support the subsequent pharmaceutical application. While the tyrosine aminomutase from *Oryza sativa* (OsTAM), the only example in a plant with great enantioselectivity, is tricky to produce  $\beta$ -tyrosine from *p*-hydroxycinnamic acid. After phylogenetic analyses of enzymes with 3,5-dihydro-5-methylidene-4H-imidazole-4-one (MIO) moiety, the (*R*)- $\beta$ -selective phenylalanine aminomutase from *Taxus chinensis* (TchPAM), as the closest relative family member of OsTAM, was utilized for computational enzyme design by using PyRosetta. After a screening of twenty-eight designed candidates from the mutant library, TchPAM was successfully mutated for acceptance of  $\beta$ -tyrosine. Among the designed mutants, the variants Tyr424Asn and Tyr424Cys are the most efficient for production of (*R*)- $\beta$ -tyrosine with excellent enantiopurity by using *trans-p*-hydroxycinnamic acid as substrate. The variant Ile431Val can also produce  $\beta$ -tyrosine. However, this variation results in an enantioselectivity switch, giving both (*R*)- and (*S*)-configured products. In summary, a lignin-based  $\beta$ -tyrosine synthesis process was successfully established, depending on the rational enzyme design of TchPAM by *in silico* methods.

## Zusammenfassung

Die wachsende Nachfrage von Energie und der zunehmende Verbrauch fossiler Ressourcen haben unterschiedlichen Problemen von Umwelt geführt, z. B. zu Smog, saurem Regen und der Zerstörung der Ozonosphäre. Andererseits wird der übermäßige Verbrauch von nicht-erneuerbarer Rohstoff wie Kohle, Erdöl und Erdgas zu der Erschöpfung der natürlichen Ressourcen führen. Die Lignocellulose wurde als Alternative entwickelt, um diese globalen Probleme zu lösen. Im Gegensatz zur Kohlenhydratfraktion der Lignocellulose wird Lignin nicht ausreichend genutzt, wegen seiner komplexen Struktur. Obwohl die Valorisierung von Lignin derzeit ein attraktives Thema ist, stellt die Konstruktion einer effizienten Umwandlung von Lignin in Feinchemikalien noch eine große Herausforderung dar.

Die optisch reinen  $\beta$ -Aminosäuren werden als grundlegende Bausteine für bioaktive Produkte bezeichnet. Unter Berücksichtigung der Eigenschaften als Antibiotikum und Antitumormittel haben sie bei den Wissenschaftlern im pharmazeutischen und agrochemischen Bereich große Aufmerksamkeit erregt. Ein klinisches und kommerziell erfolgreiches Beispiel ist Paclitaxel (Taxol®), welches im Jahre 1991 zum ersten Mal aus der Taxusart isoliert wurde. Paclitaxel ist ein Wirkstoff, der gegen die Tubuli wirkt und für die Behandlung verschiedener Leukämien und solider Krebsarten eingesetzt werden kann. Die wachsende Sterblichkeitsrate bei den Krebserkrankungen und die daraus folgende sozioökonomische Problematik sind für mehrere Entwicklungsländer eine große Herausforderung. Die Einführung einer wirksamen Strategie für die Erzeugung von bioaktiven Produkten mit Antitumorwirkung ist daher ein interessantes Thema zur Lösung dieser Probleme.

Das Promotionsthema ist Lignin-basierte aromatische Aminosäuren (LIGAROM) und ist mit Unterstützung des China Scholarship Council (CSC) in das Graduiertenprogramm BW ForWerts des Landes Baden-Württemberg eingebettet.

---

Ziel der Arbeit ist die Entwicklung einer effizienten Strategie zur Valorisierung des Lignin-Monomers *p*-Hydroxyczimtsäure, die als Beispiel für eine ungenutzte, nachhaltige aromatische Ressource gilt. Um dieses Ziel zu erreichen, wurden eine chemoenzymatische Kaskade und eine enzymatische Umsetzung mit einem computergestützten Design-Enzym zur Herstellung des  $\beta$ -Tyrosin aus *p*-Hydroxyczimtsäure durchgeführt. Am Anfang wurde die Tyrosin-Ammoniak-Lyase (TAL) mit Aminomutase-Aktivität als Template-Enzym ausgewählt. Die Enantioselektivität der bakteriellen TALs ist jedoch zu gering, um die spätere pharmazeutische Anwendung zu gewährleisten. Die Tyrosin-Aminomutase aus *Oryza sativa* (OsTAM), das einzige Beispiel in der Pflanze mit guter Enantioselektivität, ist dagegen sehr knifflig bei der Herstellung von  $\beta$ -Tyrosin aus *p*-Hydroxyczimtsäure. Nach der phylogenetischen Untersuchung von Enzymen mit 3,5-Dihydro-5-methyliden-4H-imidazol-4-on (MIO) Struktur wurde die (*R*)- $\beta$ -selektive Phenylalanin-Aminomutase aus *Taxus chinensis* (TchPAM), als das am nächsten relative Familienmitglied von OsTAM, für das computergestützte Enzymdesign mit Hilfe von PyRosetta verwendet. Nach einer Überprüfung von achtundzwanzig Kandidaten aus der Mutantenbibliothek wurde TchPAM erfolgreich zur Aufnahme des  $\beta$ -Tyrosin mutiert. Die Varianten Tyr424Asn und Tyr424Cys sind unter den designten Mutanten die effizientesten für die Produktion des (*R*)- $\beta$ -Tyrosins mit hervorragender Enantiopurität bei Verwendung der *trans-p*-Hydroxyczimtsäure als Substrat. Die Variante Ile431Val kann ebenfalls  $\beta$ -Tyrosin produzieren. Diese Variante erzeugt jedoch einen Enantioselektivitätswechsel, sodass es sowohl (*R*)- als auch (*S*)-konfigurierte Produkte gibt. In Kürze wurde ein Lignin-basierter  $\beta$ -Tyrosin-Syntheseprozess erfolgreich aufgebaut, der auf dem rationalen Enzymdesign von TchPAM durch *in silico*-Methoden beruht.

## TABLE OF CONTENTS

ACKNOWLEDGEMENT .....	II
LIST OF PUBLICATIONS.....	III
ABSTRACT.....	IV
ZUSAMMENFASSUNG.....	V
TABLE OF CONTENTS .....	VI
ABBREVIATIONS.....	VII
1 INTRODUCTION .....	1
1.1 Lignin valorization.....	1
1.2 $\beta$ -Amino acids and anticancer drugs.....	4
1.3 MIO-dependent enzymes .....	6
1.3.1 Structure of MIO-dependent enzymes .....	7
1.3.2 MIO moiety.....	11
1.3.3 Mechanism of MIO-dependent enzymes .....	13
1.3.4 The regulatory pathway involving MIO-dependent enzymes .....	15
1.3.5 The application of MIO-dependent enzymes.....	18
1.4 Computational enzyme design.....	20
2 RESEARCH PROPOSAL .....	23
3 MATERIAL AND METHODS.....	25
3.1 Materials .....	25
3.1.1 Chemicals.....	25
3.1.2 Enzymatic Kits, plasmids, and microorganisms .....	27
3.1.3 Media .....	27
3.1.4 Buffer and solutions.....	29
3.1.5 Solutions for HPLC.....	31
3.1.6 Primer.....	31
3.1.7 Software developer .....	33
3.2 Instruments.....	33
3.3 Methods.....	36
3.3.1 Chemical synthesis of 6'-(4-hydroxyphenyl)-5,6-dihydrouracil (OH-PheDU) .....	36
3.3.2 Phylogenetic methods .....	41
3.3.3 Computational enzyme design.....	41
3.3.4 Molecular biology methods .....	49
3.3.5 Activity assay for Ammonia Elimination Reaction .....	56
3.3.6 Kinetics Measurements.....	56
3.3.7 HPLC Analysis for the ammonia addition reaction .....	57
4 Results.....	59
4.1 Chemical synthesis.....	59
4.2 Determination of the suitable enzyme template.....	62

---

4.3	Selection of suitable amino acid residues for enzyme design.....	66
4.3.1	Multiple sequence alignments.....	66
4.3.2	Conservation analysis .....	69
4.3.3	The selection of the designed residues in the active site of TchPAM 71	
4.4	Computational enzyme design by PyRosetta.....	74
4.4.1	Preparation of the input files for ligand and MIO moiety .....	74
4.4.2	Preparation of the input files for enzyme template .....	75
4.4.3	Enzyme design .....	76
4.5	Optimization of enzyme expression.....	77
4.6	Enzyme purification.....	79
4.7	Comparison of relative activity in the ammonia elimination reaction....	81
4.7.1	Ammonia elimination of phenylalanine.....	81
4.7.2	Ammonia elimination of tyrosine .....	84
4.8	Kinetic parameters .....	86
4.9	HPLC analysis for amination reaction .....	90
5	DISCUSSION.....	93
5.1	Chemical synthesis of OH-PheDU .....	93
5.2	The origin of MIO-dependent enzymes.....	94
5.3	The designed TchPAM for enantiopure ( <i>R</i> )- $\beta$ -tyrosine production: Tyr424Asn and Tyr424Cys .....	97
5.3.1	Influence of the mutations on ammonia elimination reaction .....	97
5.3.2	Influence of the mutations on ammonia addition reaction.....	100
5.4	The designed TchPAM for ( <i>S</i> )- and ( <i>R</i> )- $\beta$ -tyrosine production: Ile431	103
5.5	The proposed mechanisms of enantioselectivity in TchPAM.....	105
5.6	The proposed function of the surrounding residues.....	108
5.6.1	The highly conserved residues in the active site of TchPAM .....	108
5.6.2	The residues in the aromatic binding pocket .....	114
6	CONCLUSION AND OUTLOOK.....	119
7	REFERENCES .....	123
8	APPENDIX.....	139
8.1	Figure legend .....	139
8.2	Table legend .....	144
8.3	NMR-spectra.....	146
8.4	Plasmid map.....	147
8.5	The amino acid sequence of wild type TchPAM .....	147
8.6	SDS-PAGE gels by His SpinTrap protein purification .....	148

---

## ABBREVIATIONS

<b>4CL</b>	4-coumaric acid-CoA ligase
<b>Å</b>	Ångström
<b>C3H</b>	Cytosolic coumarate 3-hydroxylase
<b>C4H</b>	Cinnamic acid 4-hydroxylase
<b>COMT</b>	3-O-methyltransferase
<b>DMSO</b>	Dimethylsulfoxid
<i>E. coli</i>	<i>Escherichia coli</i>
<i>ee</i>	Enantiomeric excess
<b>GFP</b>	Green fluorescent protein
<b>HAL</b>	Histidine ammonia-lyase
<b>HGT</b>	Horizontal gene transfer
<b>HPLC</b>	High-performance liquid chromatography
<b>HUT</b>	Histidine pathway
<b>IBLC</b>	<i>N</i> -isobutyryl-L-cysteine
<b>IPTG</b>	Isopropyl- $\beta$ -D-thiogalactopyranosid
<b>KLD</b>	Kinase, Ligase and DpnI
<b>LB</b>	Lysogeny broth
<b>MIO</b>	3,5-dihydro-5-methylidene-4H-imidazole-4-one
<b>mM</b>	Milli molar

---

<b>MS</b>	Mass spectroscopy
<b>NMR</b>	Nuclear magnetic resonance spectroscopy
<b>OD<sub>600</sub></b>	Optical density at 600 nm
<b>OH-PheDU</b>	6'-(4-hydroxyphenyl)-5,6-dihydrouracil
<b>OPA</b>	<i>ortho</i> -Phtaldialdehyde
<b>PAL</b>	Phenylalanine ammonia-lyase
<b>PAM</b>	Phenylalanine aminomutase
<b>PheDU</b>	6'-phenyldihydrouracil
<b>PTAL</b>	Phenylalanine/tyrosine ammonia-lyase
<b>PTM</b>	Posttranslational modification
<b>RFP</b>	Red fluorescent protein
<b>rpm</b>	Revolutions per minute
<b>TAL</b>	Tyrosine ammonia-lyase
<b>TAM</b>	Tyrosine aminomutase
<b>TLC</b>	Thin-layer chromatography
<b>Tris-HCl</b>	Tris(hydroxymethyl)aminomethane-hydrochloride
<b>UV</b>	ultraviolet

# 1 INTRODUCTION

## 1.1 Lignin valorization

Considering the decline of the fossil fuels, such as coal, petroleum, and natural gas, the discovery, and development of alternative renewable resources have received enormous attention during the past decade (Islam and Hossain 2021). Among them, lignocellulosic biomass, which is generated from the forestry and agricultural plant residues, presents excellent potential for the production of renewable feedstock (Yousuf et al. 2020)(Gutiérrez-Antonio et al. 2021)(Ge et al. 2018). Lignocellulose, the most abundant biopolymer on earth, consists of three foundational compounds: cellulose, hemicellulose, and lignin. In the biorefinery, the carbohydrate fraction in the lignocellulose, including cellulose and hemicellulose, is usually extracted for fermentation of glucose or conversion to fine chemicals, such as ethanol (Sorokina et al. 2017)(Takkellapati et al. 2018). Lignin, as the aromatic polymer in lignocellulose, is composed of the aromatic monolignol units with ether bonding, supporting the conduction of the water and nutrients, providing the mechanical support, as well as protecting the plants from pathogen attack (Thomas, Brian, Denis J. Murphy 2016).

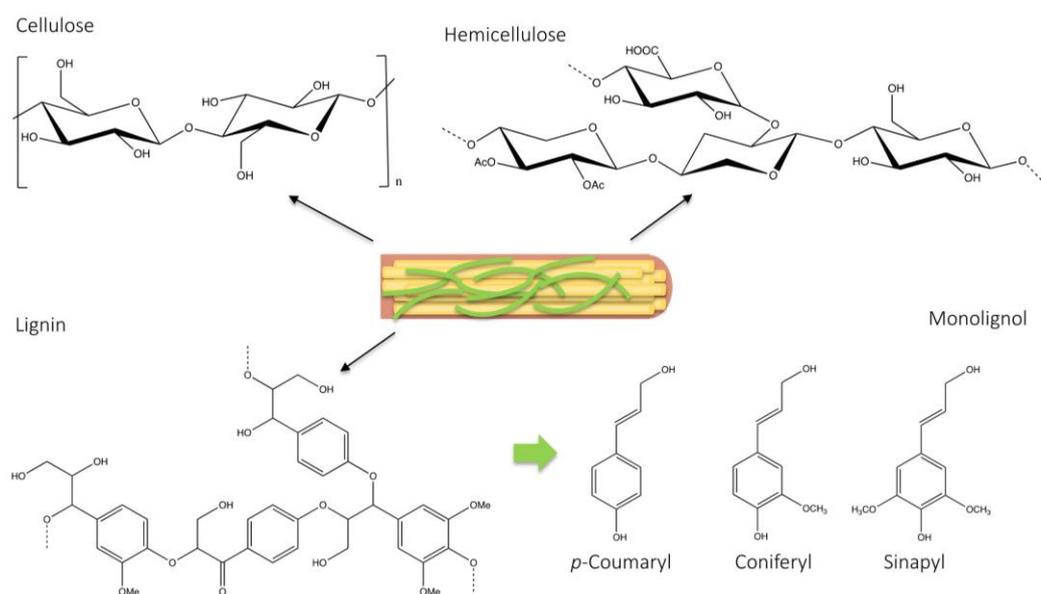
Compared to cellulose and hemicellulose, the valorization of lignin is more difficult due to the resistance of degradation with the complex structure. Therefore, until recently, over 50 million tons of lignin have been released in the pulp, paper, and bioethanol industries after extraction of the cellulose and hemicellulose (Demuner et al. 2019). Herein, the generated lignin is treated as a waster and recovered for heat and energy, even though this aromatic heterogeneous polymer has a great potential for conversion of the high-value fine chemicals. Recently, the development of the depolymerization technique allows researchers to move their eyes from the carbohydrate fraction of the lignocellulose to lignin (Mahmood et al. 2016). However, research is mainly focused on the production of lignin-based macromolecules. Of these applications of the

---

valorization of lignin, the production of carbon fibers currently appears as the most profitable example (Mahmood et al. 2016). Carbon fibers with the high specific stiffness, tensile strength, and high resistance of chemicals and temperature are mainly used for reinforcement of composites in high-end lightweight applications, such as aerospace, civil engineering, and motorsport (Chung and Chung 2012) (Khayyam et al. 2020). The development of lignin-based carbon fiber reduces the production cost, allowing the utilization of this fiber in our daily life. Furthermore, lignin, containing the substituted phenol compounds, is considered as a cost-effective candidate for replacement of the phenol-formaldehyde, which is produced as a traditional polymeric adhesive from petroleum (Ghaffar and Fan 2014) (Dongre et al. 2015). Currently, Lu *et al.* reported that the lignin can serve as a material binder to adhesive the  $\text{LiFeO}_4$  positive and graphite negative electrodes (Lu et al. 2016). Another example is the utilization of lignin as an emulsifier. In the pulping industry, the lignosulfonates are isolated from the lignocellulose after the extraction of the cellulose through sulfite pulping processes. The lignosulfonates with the sulfonate group are dispersible in the water phase, allowing their application as emulsifiers (Ekielski and Mishra 2021) (Mikkonen 2020) (Zhou et al. 2019). The most recent application of lignin in the macromolecule area is the lignin-based nanoparticles, which display antibacterial activity and high water solubility (Chauhan 2020). In contrast to the mature processing for the generation of the macromolecules from the lignin, the conversion of lignin to fine chemicals with low molecular weight remains challenging and requires further work. Until recently, there are only few examples of application to large scale manufacturing, such as vanillin (Plesu Popescu et al. 2021) (Fache et al. 2016). The price of lignin based vanillin is about 100 – 200 \$/kg, whereas the natural vanillin has been maintained in the range of 2000 – 4000 \$/kg (Sadhukhan et al. 2014) (Khwanjaisakun et al. 2020). The global demand for vanillin was over 18,000 tons in 2016, and around 20 % of them was lignin based vanillin (Ludmila, Hodásová 2015) (Martău et al. 2021). Even though the price of synthesized vanillin from petroleum-derived phenol and glyoxylic acid is more

convenient, production of vanillin in lignocellulose biorefineries is a commercial success and has received much attention from companies in these years.

Lignin is composed of three monomers, including *p*-coumaryl (H)-, coniferyl (G)-, and sinapyl (S) alcohol (Figure 1), depending on the degree of methoxylation. These monomers, known as monolignols, are formed in the phenylpropanoid pathway, responding to the drought (M. Li et al. 2018), salt (Sarker and Oba 2018), and heat stress (Wang et al. 2019). Apart from the influence factors from the environment, the lignin content differs in the species of the plants, even in the tissues and the growth stage of the individual plant. The lignin of softwood is dominantly formed by coniferyl alcohol, while both coniferyl and sinapyl alcohol are the fundamental units of the hardwood in an equal ratio (Table 1). The *p*-coumaryl alcohol is the typical monolignol with a varying amount in grasses, promoting the formation of the *p*-hydroxycinnamic acids (Tian et al. 2016). It is noteworthy that the content of the *p*-coumaryl alcohol is around 30 % in curaua leaf fiber (Timokhin et al. 2020).



**Figure 1:** lignocellulosic biomass is composed of carbohydrates, including cellulose and hemicellulose, and aromatic polymer: lignin. Lignin contains mainly three monolignol subunits: *p*-coumaryl (H), coniferyl (G), and sinapyl (S).

**Table 1:** Distribution of the monolignol content in different types of lignin (Azadi et al. 2013).

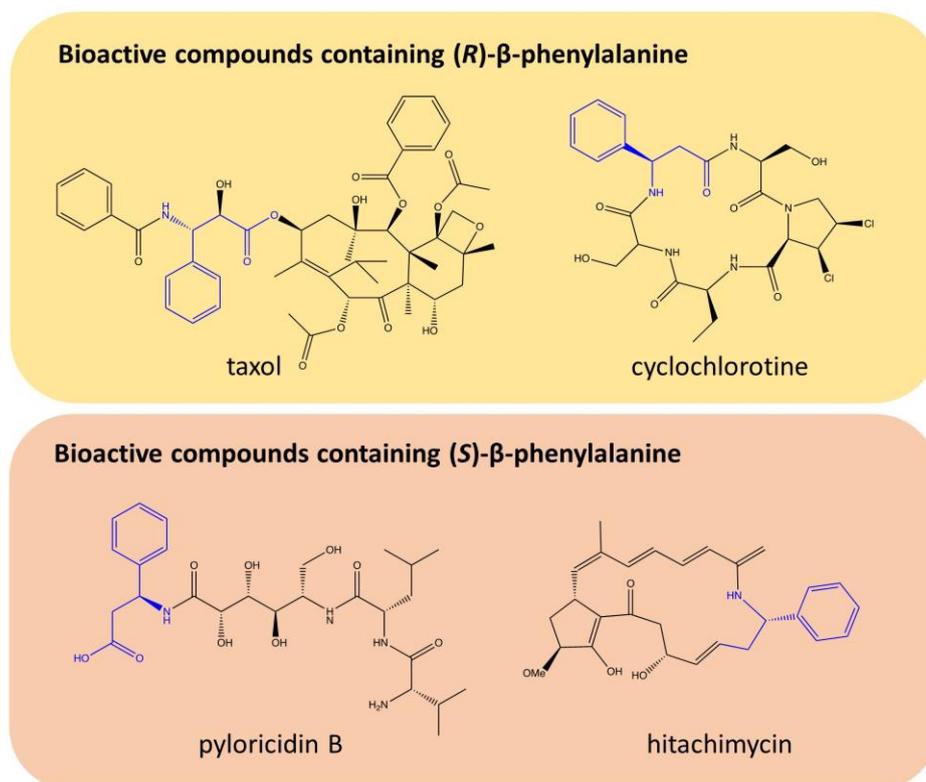
Monomers	Softwood	Hardwood	Grasses
<i>p</i> -Coumaryl (H)	-	-	5-30 %
Coniferyl (G)	90-95 %	50 %	75 %
Sinapyl (S)	5-10 %	50 %	25 %

## 1.2 $\beta$ -Amino acids and anticancer drugs

Cancer is the second major cause of mortality. Despite the reduction of the death rate in past years, it still takes approximately 10 million people's lives only in 2020, which is reported by WHO (WHO 2020). Considering population aging, the disease burden of cancer, spending on anticancer drugs, has been enhancing in many countries (Hofmarcher et al. 2020). Meanwhile, the cancer incidence and death rate are much higher for patients with lower socioeconomic status due to the unaffordable drugs (Tabuchi 2020) (Prasad et al. 2017). Therefore, the establishment of an effective pathway for the production of anticancer drugs is important for not only the pharmaceutical industries but also the national health care system and the patients.

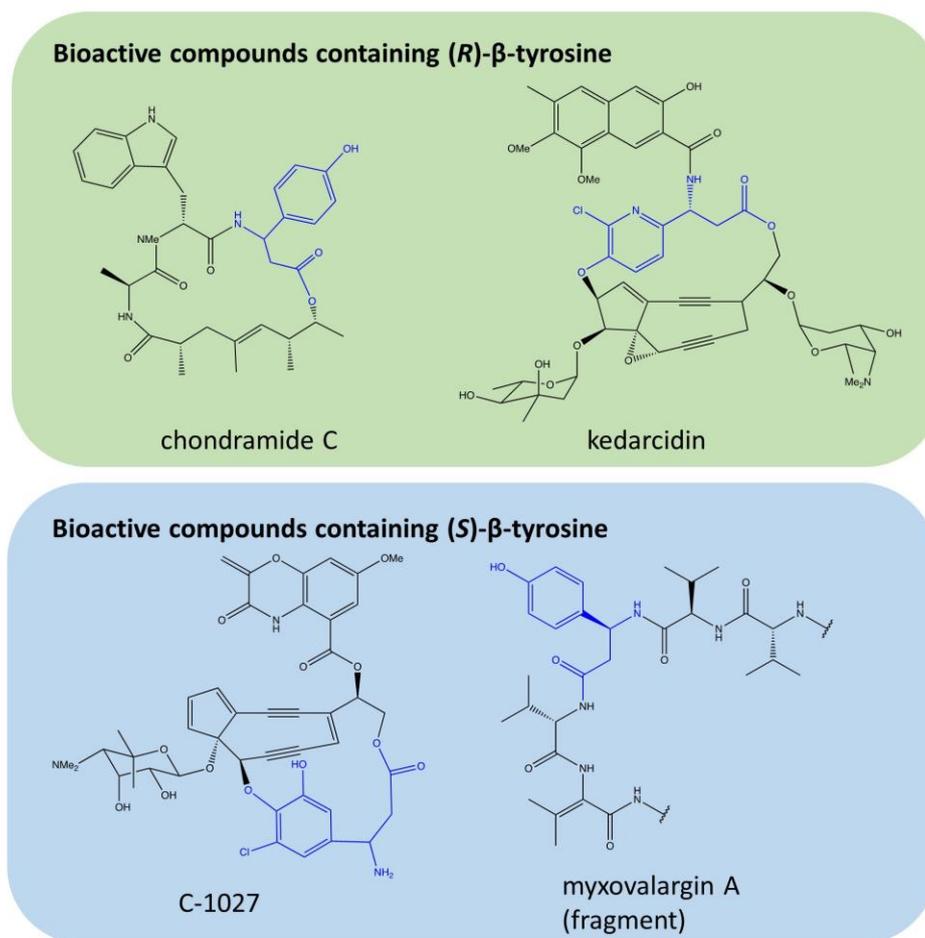
In contrast to the proteinogenic  $\alpha$ -amino acids, which are encoded by sixty-one triplet codons, the noncanonical-amino acids are not incorporated in the natural polypeptide chains (Rich et al. 2021). Among them,  $\beta$ -amino acids have attracted much attention in the pharmaceutical application during the past decade, since they are recognized as the chiral building blocks of bioactive compounds due to their anticancer and antibiotic effect. Taxanes, including paclitaxel, docetaxel, and cabazitaxel, present the well-known natural antimicrotubule agents by targeting the cytoskeleton (McGrogan et al. 2008). With a taxadiene core, paclitaxel (Taxol®) is the most investigated member of the taxane family of the marketed antitumor drug for the treatment of breast, lung, and ovarian cancer (Figure 2). After discovery from the Pacific yew tree, *Taxus brevifolia*, paclitaxel is also found in the endophytic, epiphytic, pathogenetic, and saprophytic

fungal, especially, the *Penicillium* species. (Malik et al. 2011) (Bi et al. 2011) (Soliman and Raizada 2013). Paclitaxel is assembled by a functionally necessary C13-side chain with  $\beta$ -phenylalanine-derivate and promotes cell death by stabilization of the microtubule against depolymerization (Malik et al. 2011) (Gond et al. 2014).



**Figure 2:** Example of bioactive compounds containing  $\beta$ -phenylalanine. The unit of  $\beta$ -phenylalanine is colored in blue (Parmeggiani, et al. 2018).

As the outstanding candidate of the anticancer drug, C-1027, containing the unique  $\beta$ -tyrosine derivate, (*S*)-3-chloro-4,5-dihydroxy- $\beta$ -phenylalanine, has developed in phase II of a clinical trial (Figure 3) (Wang et al. 2012). This chromoprotein is formed by the Gram-positive bacterium, *Streptomyces globisporus*, and has the antibiotic and antitumor activity with the enediyne structure by breaking the DNA double-strand (Christianson et al. 2007) (Rachid et al. 2007). With a similar structural feature and mechanism, the maduropeptin, which is extracted from *Actinomadura madurae*, also belongs to the enediyne family with antitumor and antibiotic activity (Van Lanen et al. 2007).



**Figure 3:** Example of bioactive compounds containing  $\beta$ -tyrosine. The unit of  $\beta$ -tyrosine is colored in blue (Parmeggiani, et al. 2018).

### 1.3 MIO-dependent enzymes

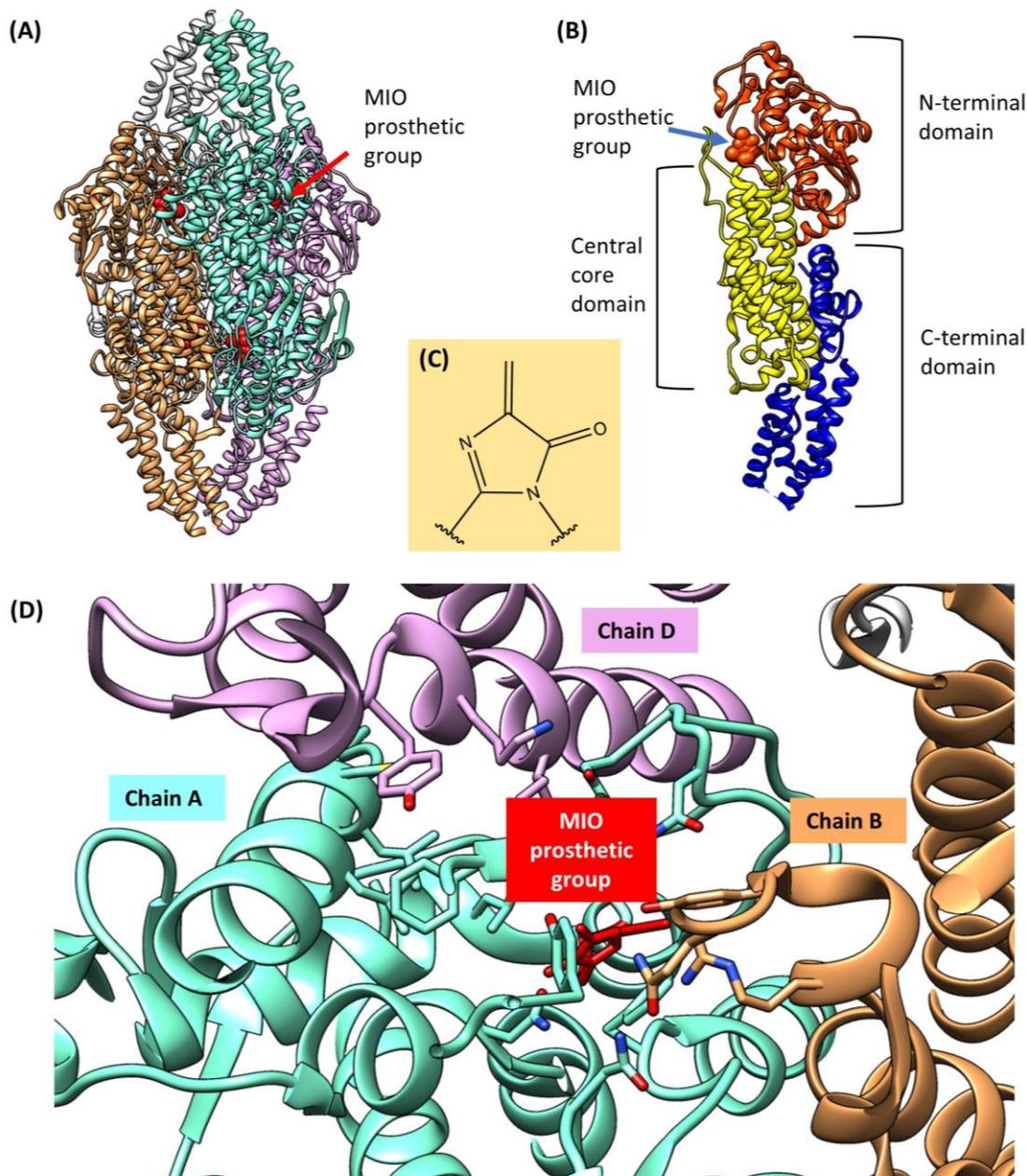
The MIO prosthetic group is observed in the aromatic ammonia-lyase because of the poor acidity of the  $\beta$ -proton in the aromatic amino acids (Punekar 2018) This highly electrophilic moiety is unique in the aromatic amino acids ammonia-lyases, including phenylalanine ammonia-lyase (PAL; EC 4.3.1.24), tyrosine ammonia-lyase (TAL; EC 4.3.1.23), and histidine ammonia-lyase (HAL, EC 4.3.1.3), which are considered as the MIO-dependent enzymes. The ammonia-lyase can catalyze the elimination of the ammonia from the aromatic amino acids and form the corresponding unsaturated carboxylic acids as products. Currently, the PALs with bifunctional activity, which confer tyrosine acceptance, are sorted in the phenylalanine/tyrosine ammonia-lyase

---

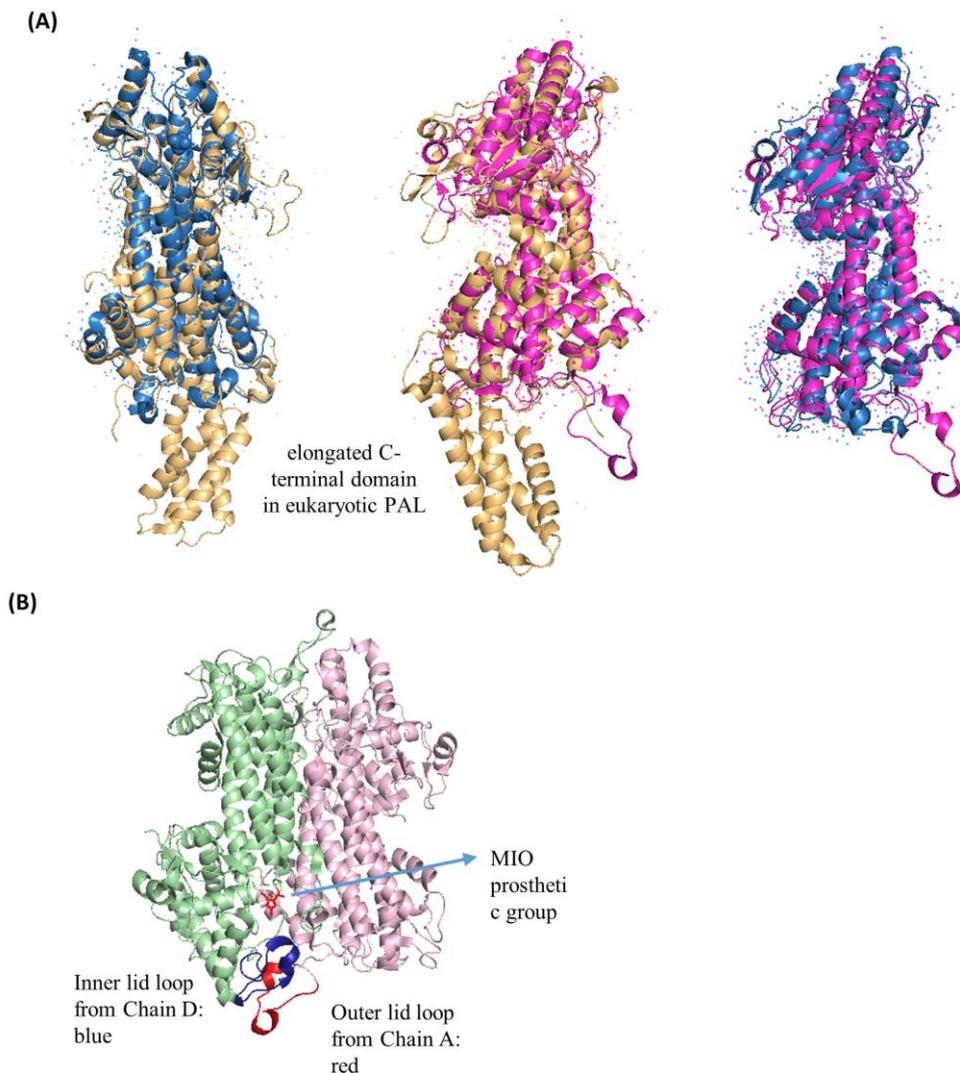
(PTAL; EC 4.3.1.25). While, the aromatic amino acid aminomutases, which catalyze the isomerization of proteinogenic L- $\alpha$ -amino acids to their  $\beta$ -regioisomers, share similar structures and mechanisms. Therefore, the phenylalanine aminomutase (PAM; EC 5.4.3.10) and tyrosine aminomutase (TAM; EC 5.4.3.6) also belong to the MIO-dependent enzyme family. In addition, the enzyme, which can catalyze the 2,3-shift of the  $\alpha$ -amino group of histidine to produce the  $\beta$ -histidine, as a putative histidine aminomutase is not discovered until now.

### **1.3.1 Structure of MIO-dependent enzymes**

Upon the previous X-ray crystallographic studies, the members of MIO-dependent enzymes are homotetramers (Moffitt et al. 2007). Each monomeric subunit contains individual active sites and connects with the other two monomers in a head-to-tail way (Matthew M Heberling et al. 2015). In common, the rigid of the monomer is formed by parallel  $\alpha$ -helices. In the TchPAM, the core domain is built by the bundles of four  $\alpha$ -helices, which are surrounded by the six  $\alpha$ -helices from the elongated C-terminal domain (MacDonald and D’Cunha 2007) (Jun et al. 2018). Apart from the C-terminus, the central core domain is connected by the globular N-terminal domain, in which the MIO prosthetic group is observed, indicating that each active site is placed between the core- and N-terminal domains and consists of the residues from three different monomers (Figure 4)(Feng et al. 2011).



**Figure 4:** (A) The X-ray crystal structure of TchPAM (PDB ID: 4C5R) with four monomeric chains, which are colored in white, light blue, orange, and purple. The MIO prosthetic group is marked in a red sphere (Wybenga et al. 2014). (B) The monomer contains three main domains: N-terminal domain (orange-red), central core domain (yellow), and C-terminal domain (dark blue), The MIO prosthetic group is placed in the N-terminal domain (red) (Feng et al. 2011). (C) Chemical Structure of MIO prosthetic group. (D) Each active site is composed of the residues from three monomeric chains. The MIO prosthetic group, which is marked in red, is the moiety in chain A, is surrounded by the residues from chain A (light blue), B (orange), and D (purple).



**Figure 5:** (A) Comparison of the X-ray structure of PAL, TAL, and HAL. The color of the RtPAL from *Rhodotorula toruloides* (PDB entry 1T6J) is yellow, while the PpHAL from *Pseudomonas putida* (PDB entry 1GKM) and RsTAL from *Rhodobacter sphaeroides* (PDB entry 2O6Y) are colored in blue and pink, respectively. (B) The inner loop (blue) and out loop (red) are shown in RsTAL with the MIO prosthetic group.

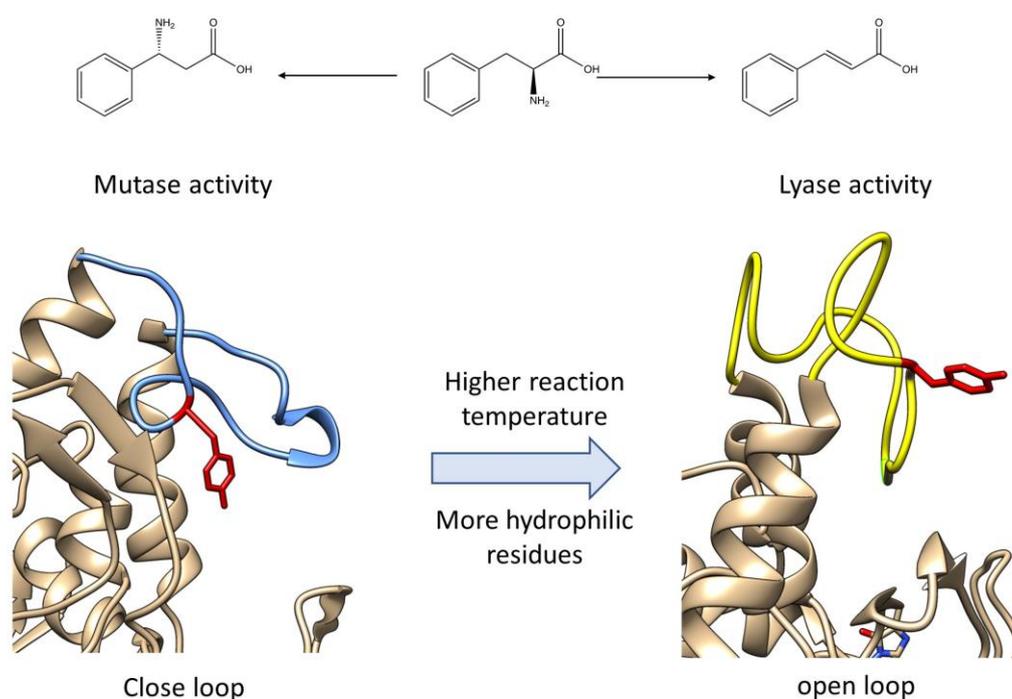
From the structural point of view, the sequence length of the prokaryotic HALs is around 500 amino acids, which is similar to the bacterial TALs/TAMs. While the eukaryotic PALs show sequences longer than 700 amino acids, which are bigger than these members from the MIO-dependent enzyme family since eukaryotic PALs contain an extended C-terminal multi helix domain (Figure 5) (Allwood et al. 1999) (Pilbák et

---

al. 2006) (Ritter and Schulz 2004). The additional domain in eukaryotic PALs is functionally essential for higher plants that can deactivate the enzyme to regulate the biosynthesis of the phenylpropanoids, responding the unfavorable environmental stress. In *Phaseolus vulgaris*, the phosphorylation site of the PAL is identified as Thr545. While the PAL from *Rhodospiridium toruloides* can be cleaved by trypsin and chymotrypsin at the residues Arg123 and Tyr110, respectively. The residue Thr545 in PvPAL and the residues Arg123 and Tyr110 in RtPAL are located in the above-mentioned extended C-terminal multi helix domain, indicating that this domain plays a key role in the regulation of the phenylpropanoid pathway by reduction of the enzyme lifetime (Wang et al. 2005).

A narrow tunnel composition is formed by an inner and an outer loop in the N-terminal region, carrying the functional essential MIO prosthetic group. The HALs show a closed tunnel structure by two covered loops, avoiding the contact between their active site and the solvents. In contrast with HALs, the PALs display a more mobile structure, which affects not only the enzyme activity but also the ratio of mutase/lyase activity. Apart from the MIO prosthetic group, the catalytic tyrosine residue is placed on the inner loop, which is highly conserved in the MIO-dependent enzymes (Tyr53 in PpHAL, Tyr110 in PcPAL, Tyr60 in RsTAL, Tyr80 in TchPAM). This tyrosine residue is strictly required for the enzyme activation since mutants harboring variation of this residue to alanine and phenylalanine are completely abolished (Röther et al. 2002) (Wybenga et al. 2014). The inner loop, which is formed by more hydrophilic residues (Ala77Thr, Ile79Ser, Cys89Thr, and Leu97Gly) has a flexible conformation. Similarly, at a higher temperature, the inner loop is more conformationally flexible, leading to a shift in the aminomutase activity to lyase activity. Additional evidence comes from the research about the distance between a catalytic tyrosine residue and the MIO prosthetic group at the different reaction temperatures, indicating that the phenolic O-atom of the residue is in close contact with the exocyclic methylene C-atom of the MIO moiety at a lower temperature, affording aminomutase to lyase features (Attanayake et al. 2018)

(Matthew M. Heberling et al. 2015) (Pilbák et al. 2006). Furthermore, the PaPAM from *Pantoea agglomerans* and PAL from *Streptomyces maritimus* show predominantly aminomutase activity and less lyase activity at a lower temperature, but increase the ability for ammonia addition reaction by enhancing the reaction temperature (Chesters et al. 2012). In conclusion, the hydrophilic feature of the inner loop and the reaction temperature have a great influence on the flexibility of the inner loop, as well as the distance between the catalytic tyrosine residue and MIO moiety, varying the ratio of the aminomutase and lyase activity (Figure 6).

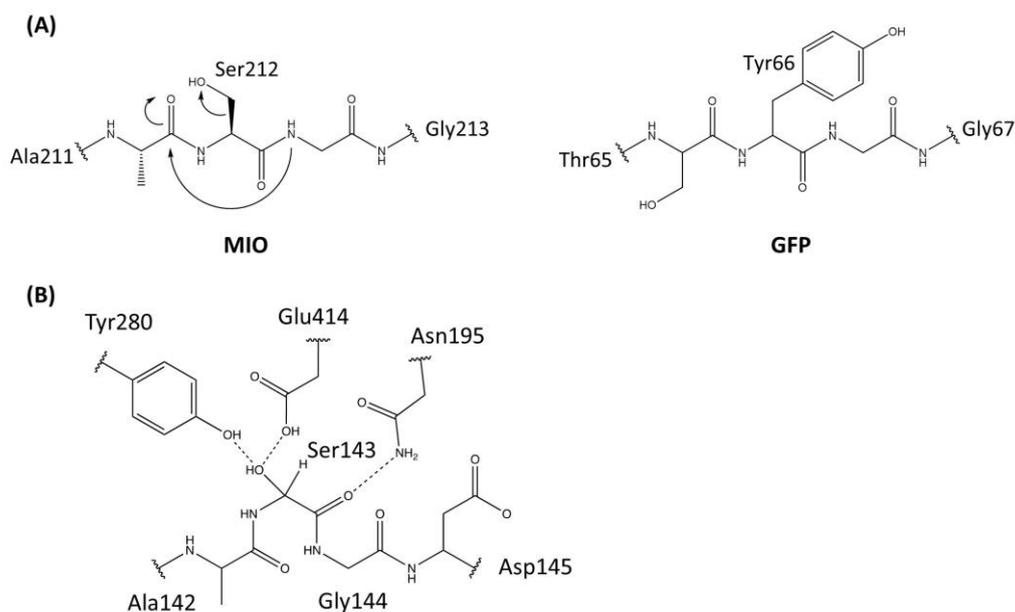


**Figure 6:** High reaction temperature and more hydrophilic residues led to a flexible lid-like inner loop and a shift in the aminomutase activity to lyase activity.

### 1.3.2 MIO moiety

The side chains of proteogenic amino acids provide the nucleophilic group in the active site for enzymatic catalysis in common. While, the metal ions and organic molecules from the environment serve as the electrophilic group, promoting the corresponding catalysis (Punekar 2018) Expecting from these cofactors, some enzymes can form the electrophilic group through posttranslational modifications (PTMs) of the amino acid

side chains by themselves (Müller 2018). Of these PTMs, the MIO prosthetic group assists the elimination and addition of ammonia from aromatic amino acids. This strongly electrophilic MIO moiety can be found in the active site of the aromatic amino acid ammonia-lyase and aminomutase, which are sorted in the MIO-dependent enzyme family.



**Figure 7:** (A) Comparison of the tripeptide in PpHAL from *Pseudomonas putida* and AvGFP from *Aequorea victoria*. (B) The tripeptide for the formation of the MIO prosthetic group with its functional essential neighboring residues in PpHAL.

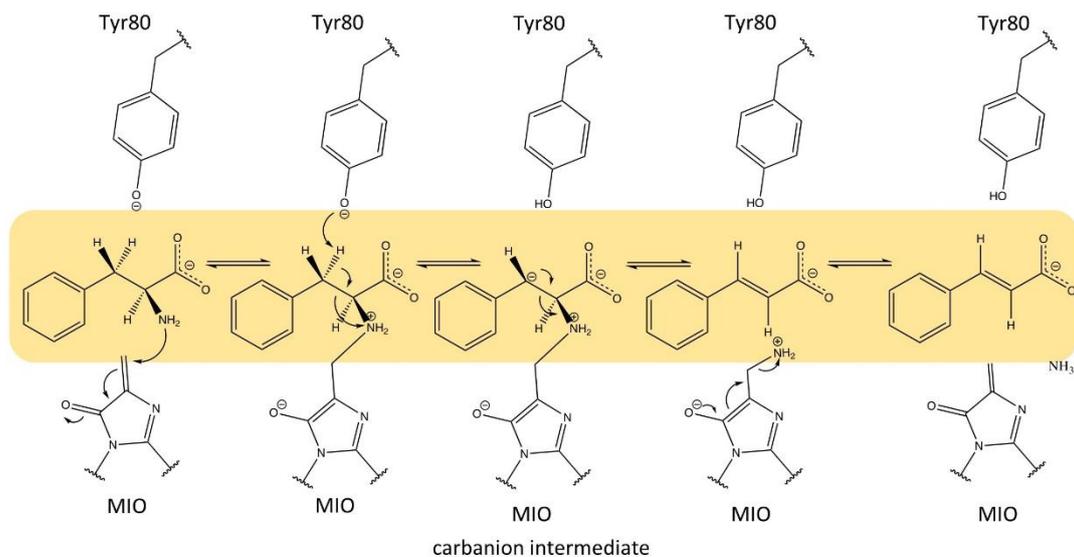
The MIO prosthetic group, which is spontaneously formed by a tripeptide Ala-Ser-Gly through cyclization and two-step dehydration during the late stage of protein folds, has been identified by the crystal structure study on PpHAL (Figure 7 (A)) (Schwede et al. 1999) (Baedeker and Schulz 2002). In contrast to the highly conserved residues serine and glycine, the first residue in this inner amino acid triad can be replaced by serine, cysteine, and threonine in a few members of the MIO-dependent enzymes. According to the previous mechanism researches on green fluorescent protein from *Aequorea victoria* (AvGFP), which contains a similar tripeptide Thr-Tyr-Gly, the formation mechanisms of the MIO moiety have been proposed by Torsten F. *et al* (Reid and Flynn 1997). The first step of the cyclization is the nucleophilic attack by the amide of the

---

glycine (Gly67 in AvGFP, Gly144 in PpHAL, and Gly177 in TchPAM) to the carbonyl of alanine (Ser65 in AvGFP, Ala142 in PpHAL, and Ala175 in TchPAM). Meanwhile, the neighboring residues (Asp145 in PpHAL and Asp178 in TchPAM) provide the mechanical compression to press the tripeptide loop. This mechanical compression is strictly required for MIO moiety formation since the nucleophilic attack is electronically unfavorable. Additional supporting evidence comes from the single-point mutagenesis on PpHAL. The MIO prosthetic group has been not observed in the active site of the mutant Asp145Ala in PpHAL (Baedeker and Schulz 2002). The current molecular simulation study on PpHAL has reported that the residue Asn195, Glu414, and Tyr280, which are located in close contact with the MIO moiety. Among them, the residue Asn195 forms the hydrogen bond with the reacting amide, promoting the cyclization of the MIO prosthetic group. For exocyclic dehydration and removal of the hydroxyl group from the residue Ser143, the residues Glu414 and Tyr280 (Gln459 and Tyr322 in TchPAM) can build a hydrogen-bonding network with the residue Ser143 (Figure 7 (B)) (Sánchez-Murcia et al. 2016). Furthermore, the results from calculations have pointed out the importance of the water molecules on the formation of the MIO prosthetic group (Sánchez-Murcia et al. 2016).

### 1.3.3 Mechanism of MIO-dependent enzymes

As the highly electrophilic moiety, the exocyclic alkene of the MIO prosthetic group forms an amino-MIO intermediate with the  $\alpha$ -ammonium group of the substrate. In the E1cB mechanism, which is most accepted, the amino-MIO intermediate with a secondary ammonium ion facilitates the abstraction of the  $\beta$ -proton by the catalytic tyrosine residue Tyr80 in TchPAM (Figure 8) (Wybenga et al. 2014). Following the deprotonation and formation of the carbanion intermediate, the amino-MIO moiety is released, which subsequently removes the ammonia and generates the electrophilic MIO moiety. Finally, the corresponding unsaturated carboxylic acid is produced (Hanson and Havir 1970).



**Figure 8:** The E1cB catalytic mechanism of the MIO-dependent enzymes with phenylalanine and cinnamic acid as accepted substrates

According to the position of ammonia readdition, the aromatic amino acid aminomutases have two mechanisms of isomerization. By using (*S*)- $\alpha$ -phenylalanine as the starting substrate, the PaPAM from *Pantoea agglomerans* produces (*S*)- $\beta$ -phenylalanine by reanimation on the same stereoheterotopic face, which is often called (*S*)-selective PAM (Ratnayake et al. 2011). While, the  $\beta$ -proton is abstracted on the opposite face in the (*R*)-selective TchPAM, promoting the formation of the (*R*)- $\beta$ -phenylalanine as a product (Wu, et al. 2012). When the (*S*)- $\alpha$ -phenylalanine enters the active site, the carboxylate group of the (*S*)- $\alpha$ -phenylalanine is anchored by the residue Arg325 with a bidentate salt bridge. Meanwhile, the carboxylate group is also stabilized by the residue Gln319. In the PaPAM, the (*S*)- $\beta$ -phenylalanine is formed, following the readdition of the ammonia on the  $\beta$ -position of cinnamic acid intermediate (Ratnayake et al. 2011). Whereas, the bond rotation of this intermediate is required in the active site of the TchPAM for reanimation on the opposite face. The mechanism of the conformational change has been debated for many decades, with the different hypotheses, containing the rotation around  $C_1-C_\alpha$  bond or  $C_\beta-C_{ipso}$  bond or both bonds together (Ratnayake et al. 2011) (Wu, et al. 2012) (Feng et al. 2011) (Wang et al. 2013). After the rotation of the carbon bonds, the cinnamic acid intermediate is reoriented. At

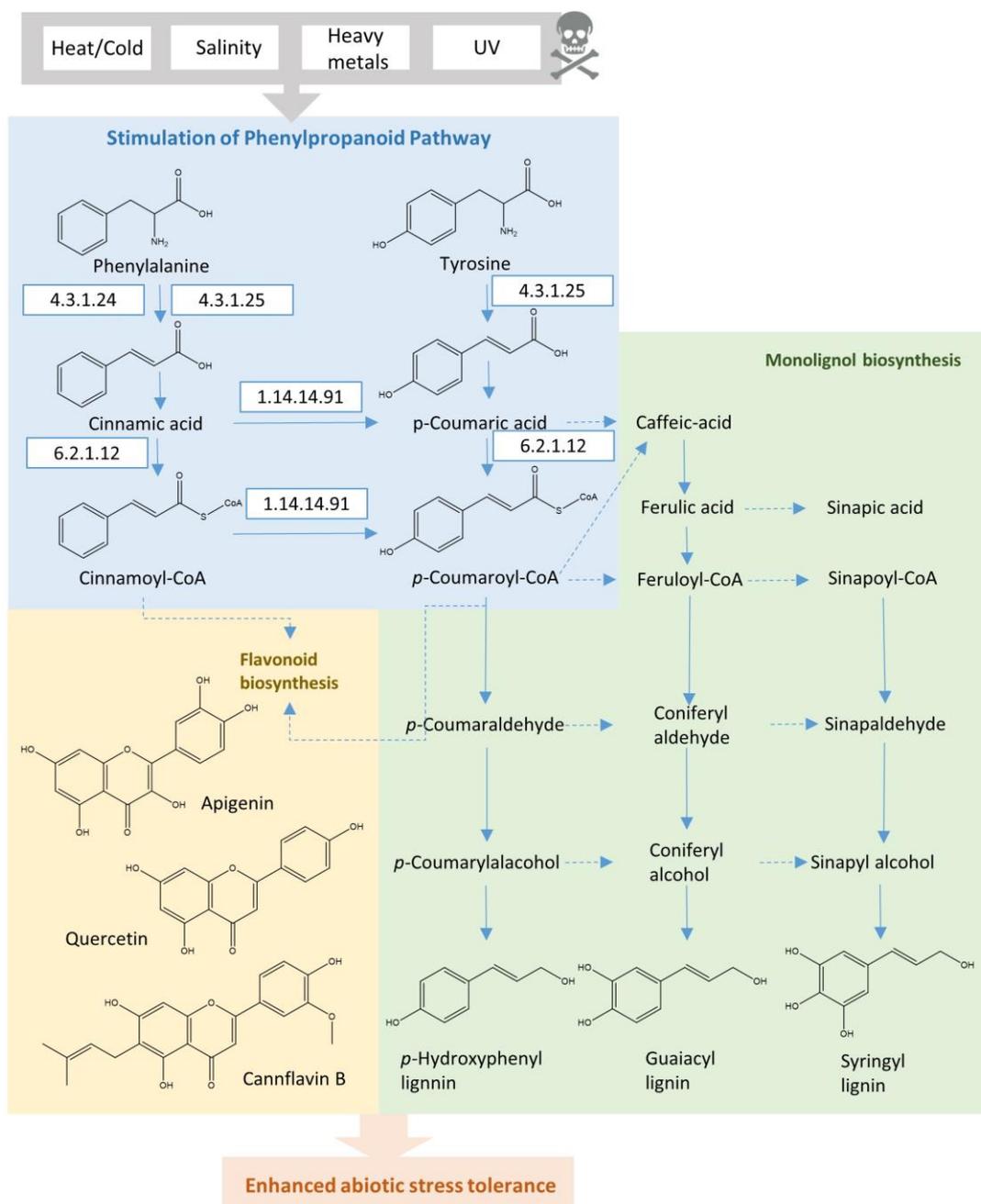
---

this moment, the carboxylate group of the intermediate is bound by the residues Asn231, Arg325, and Asn355 (Wu, et al. 2012). These two proposed binding modes have implied the influence of the residues Asn231, Gln319, Arg325, and Asn355, which are sorted in the carboxyl binding pocket, on the substrate selectivity of the TchPAM. Apart from these polar residues, the residues such as the Tyr322 and Phe371 have been also summarized in the carboxyl binding pocket in the previous research on TchPAM by Wu *et al.* In contrast to these residues, they have also pointed out an aromatic binding pocket, in which the residues stabilize the aromatic ring of the substrate, including Phe86, Leu104, Cys107, Leu108, Leu179, and Gln459 (Wu, et al. 2012).

### **1.3.4 The regulatory pathway involving MIO-dependent enzymes**

#### **Phenylpropanoid pathway**

The phenolic metabolites are accumulated in plants through the phenylpropanoid pathway under unfavorable environmental conditions. Of these secondary metabolites, flavonoids are the major compounds, which can protect the plants from abiotic stress, such as drought, heavy metal, salinity, and UV radiation (Sharma et al. 2019). Among them, the avoiding of the plants from the damage, which is caused by exposure to UV radiation, is the most fundamental function of the accumulation of the flavonoids. Especially, the ancient plants, for example, yew species, which are faced many challenges on the photoprotection. These accumulated flavonoids under the epidermal layer act as a shield for adsorbing visible radiations. Meanwhile, they can prevent dimerization of thymine, reducing the damage of the protein and DNA, as well as enhancing the tolerance of the reactive oxygen species under the harmful UV radiation (Kootstra 1994) (Treutter 2006). The other secondary metabolite, monolignols, which are the building blocks of the lignin, provide mechanical support to plant cell walls under cold stress (Figure 9).



**Figure 9:** Responding to the unfavorable stress parameters such as drought, heavy metal, salinity, and UV radiation, several phenolic metabolites are accumulated by regulation of the corresponding enzymes in the phenylpropanoid pathway of higher plants.

The accumulation of the phenolic metabolite in plants is caused by the regulation of a series of biosynthetic enzymes to respond to unfavorable conditions (Sharma et al. 2019). As the primary regulatory enzyme in the phenylpropanoid pathway, PALs catalyze the conversion of the L-phenylalanine and generate the *trans*-cinnamic acid, connecting the shikimate pathway by controlling the carbon flux. Subsequently, this

---

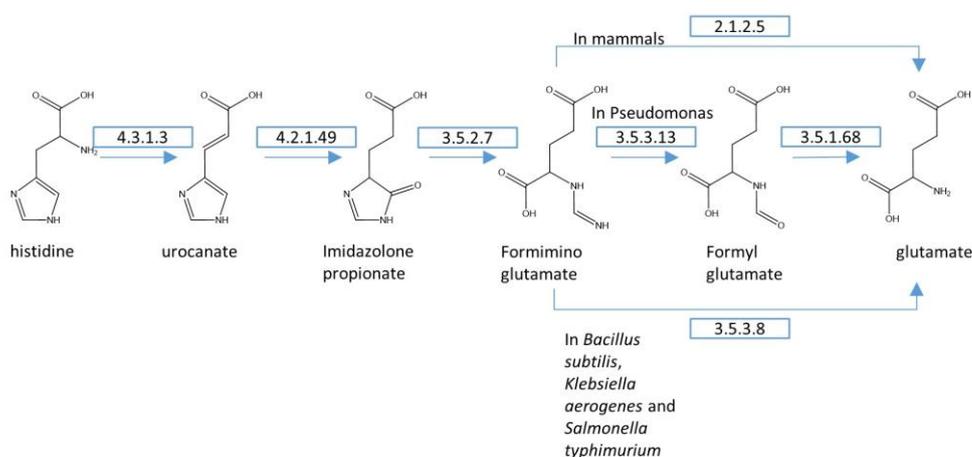
formed *trans*-cinnamic acid is transformed to *trans-p*-hydroxycinnamic acid by cinnamic acid 4-hydroxylase (C4H; EC 1.14.13.11). Then, the ATP-dependent formation of the *p*-coumaroyl CoA employs the 4-coumaric acid-CoA ligase (4CL; EC 6.2.1.12). This formation step is the branch point of the biosynthesis of the phenylpropanoids, leading to different secondary metabolites. In conclusion, PAL, C4H, and 4CL are the fundamental enzymes and represent the major conversion step in the phenylpropanoid pathway in higher plants.

Apart from the L-phenylalanine, which comes from the shikimate pathway, L-tyrosine is also used as the starting material for the phenylpropanoid pathway in certain plants, especially the monocots species. In contrast to the dicotyledonous plant, these plant species employ the PTAL with both phenylalanine- and tyrosine-ammonia-lyase activity, which can directly generate the *trans-p*-hydroxycinnamic acid from L-tyrosine. The cell walls of these monocots species consist of the higher contents of the syringyl (S)-rich lignins, flavonoid tricetin, and more esterified hydroxycinnamic acid (Barros et al. 2016). Avoiding the hydroxylation by C4H may indicate the alternative routes by employing the cytosolic coumarate 3-hydroxylase (C3H), promoting an efficient conversion of the *trans-p*-hydroxycinnamic acid to caffeic acid. In addition, the proportions of the (S)-monolignol is higher in the cell walls of monocots species due to the bypassing the biosynthesis of the (H)- and (G)-monolignol by cytochrome P450. In these species, the alternative route is efficient by employing the bifunctional PTAL and C3H (Barros et al. 2019).

### **Histidine pathway**

HAL is the initial enzyme in the histidine catabolism that transforms histidine into *trans*-urocanic acid. The subsequent enzyme is urocanate hydratase (EC 4.2.1.49), which is used to produce 4-imidazole-5-propionate by using the *trans*-urocanic acid. Then, the formimino glutamate is formed by imidazolonepropionase (EC 3.5.2.7). For the generation of the final product L-glutamate, the histidine pathway branches into various routes with different enzymes, depending on the type of the organism. In

mammals, the tetrahydrofolate (THF)-dependent glutamate formiminotransferase (EC 2.1.2.5) is employed for the production of L-glutamate. Whereas, the N-formimino-L-glutamate is hydrolyzed by formiminoglutamase (EC 3.5.3.8) in *Bacillus subtilis*, *Klebsiella aerogenes*, and *Salmonella typhimurium*. The bacterial genera *Pseudomonas* generate the L-glutamate by two-step conversion with formimidoylglutamate deiminase (EC 3.5.3.13) and N-formylglutamate deformylase (EC 3.5.1.68), which are coregulated by the *hutF* and *hutG* genes (Figure 10) (Bender 2012) (Kohlmeier 2015).



**Figure 10** Histidine pathway in mammals, *Pseudomonas*, and the other bacterial genera with various enzymes.

HAL is mainly found in the skin and liver and is important for the growth of children. In the stratum corneum of the skin, the *trans*-urocanic acid is accumulated due to the lack of urocanase, which can eliminate the ammonia from the *trans*-urocanic acid. Meanwhile, under UV exposure, the *trans*-urocanic acid is isomerized into *cis*-isomer, which may initiate the immunoregulation (Norval 2001).

### 1.3.5 The application of MIO-dependent enzymes

The MIO-dependent enzymes occupy a key position in pharmaceutical and medical applications. Considering the antibiotic and antitumor features of the  $\beta$ -aromatic amino acids, which are produced by the MIO-dependent enzymes, these enzymes have been

---

of great interest to researchers in the investigation of the treatment of cancer during the last decade. On the other hand, MIO-dependent enzymes have been also directly used in the therapeutic application. There are nine essential amino acids in the human diet, which are not synthesized by humans, including phenylalanine and histidine. Compared to the normal cells, the tumor cells are more sensitive under the limited condition with the lower concentration of the essential amino acids in the blood plasma (Tabe et al. 2019) (Parmeggiani, et al. 2018). For instance, melanoma has a higher demand for tyrosine and phenylalanine. Therefore, the PAL or TAL, which can be used as the targeted therapeutic agents through depletion of the corresponding amino acids, present great potential for inhibition of the cell growth and restriction of the uncontrolled cell division of the melanoma (Fu et al. 1999) (Dhankhar et al. 2020).

Another remarkable application of MIO-dependent enzymes is the treatment of certain hereditary diseases. For example, phenylketonuria is caused by the mutation of the gene, encoding the phenylalanine hydroxylase with abnormal activity, leading to an accumulation of the phenylalanine until the potentially toxic concentration in the body (Blau et al. 2010) (Fusetti et al. 1998). The symptoms of phenylketonuria include intellectual disability, mental disorders, seizures, and congenital heart defects (Brown and Lichter-Konecki 2016) (Rouse et al. 2000). The tyrosinemia type I is the other inborn error of catabolism, involving the accumulation of the tyrosine in the body through the defective activity of the fumarylacetoacetate hydrolase, which leads to progressive liver disease, renal tubular, neurological disorders, and defective blood clotting (Kvittingen 1986) (Holme and Lindstedt 1995). While the similar disease, which is caused by the lack of the HAL with normal activity, is histidinemia (Auerbach et al. 1962). Instead of the traditional treatment with the strict dietary restrictions of the phenylalanine, the PALs have been used as medical agents by injection or oral administration to reduce the concentration of the phenylalanine in the patient's body. As a promising strategy for the treatment of phenylketonuria, the recombinant PAL conjugated with polyethylene glycol, which is named pegvaliase from the BioMarin

---

Pharmaceutical Inc., has been tested in the phase III of the clinical trials (Gupta et al. 2018).

#### **1.4 Computational enzyme design**

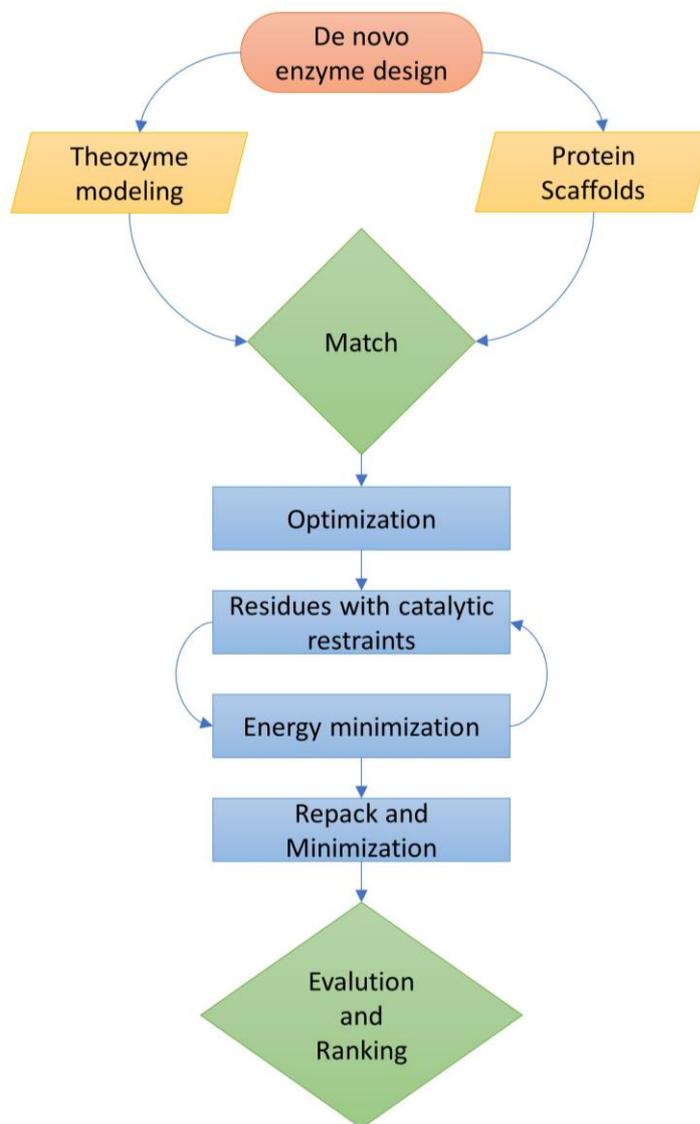
With the increasing impact on the manufacturing of biocatalysis, the traditional chemical industry has continued to face challenges during the past decade, considering the utilization of harmful organic solvents and the production of toxic waste. In contrast with chemical synthesis, enantioselectivity is one of the most attractive advantages of enzymatic conversion, especially in the production of chiral compounds (Kataoka et al. 2016). However, enzymes are unavailable for many of the desired synthesis processes because of their narrow substrate scope and poor stability against extreme temperature, organic solvent, and conditions with extreme pH values. To overcome the limitation of a natural enzyme, protein engineering became a trending topic in the area of chiral synthesis.

Currently, unlike the directed evolution with randomly screening in a large mutants library, enzyme engineering based on rational approaches has been developed, depending on the prior information on structures and function of proteins. For decades, the determination of protein structure by X-ray crystallography has been limited by the crystallization step, since not all of the proteins can be crystallized. With the recent technological advances in instrumentation, like cryo-electron microscopy (Cryo-EM), which can determine the structure without crystallization, more protein structures are available (Bai et al. 2015). On the other hand, the three-dimensional structures of protein can be also predicted (McGuffin et al. 2000). A more recent protein structural prediction program is AlphaFold, which is a trending artificial intelligence program with a deep learning algorithm (AlQuraishi 2019). Compare to the other prediction method, AlphaFold, has achieved a demonstrating accuracy (Jumper et al. 2021a). Considering its ability of machine learning to involve physical and biological knowledge, AlphaFold will help researchers to reduce the time-consuming works on the fundamental elimination step of false structures and the costly experimental

---

structure studies (Jumper et al. 2021b). AlphaFold is not a well-functioning protein engineering program right now. For example, it has not been trained for the prediction of ligands, cofactors, post-translational modifications (Perrakis and Sixma 2021). However, as the top-ranked prediction method with a strong machine learning algorithm, AlphaFold still present great potential for protein structural studies.

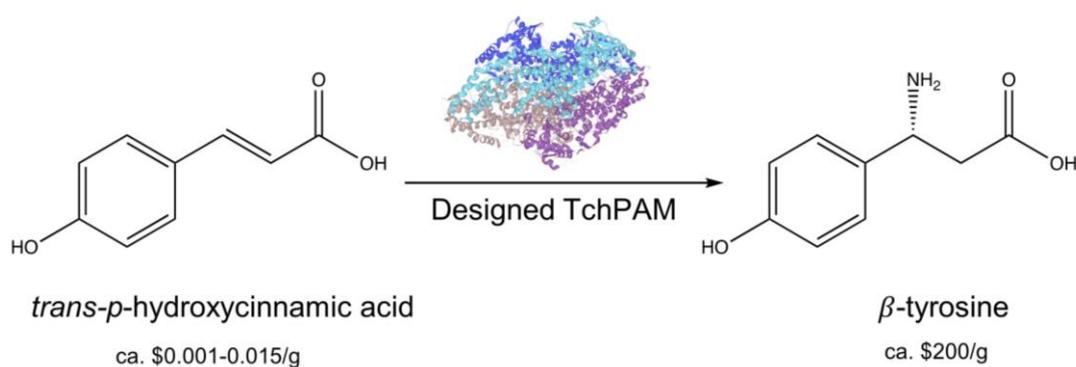
PyRosetta is one of the strong computational protein design tools for macromolecular modeling with a broad range of functions, including structure prediction, docking, analysis, and design (Chaudhury et al. 2010). Among them, the *de novo* enzyme design has gained much attention and presents a huge potential in the field of chemical and biotechnological production. Usually, the engineered enzyme is designed by the *de novo* enzyme design approach in four stages (Figure 11). Firstly, depending on the catalytic mechanism of the enzyme, a reaction transition-state model is built with a series of catalytically essential residues, which is defined as theozyme. Then, the hypothetical theozyme is placed into a protein scaffold by using the RosettaMatch module. After the matching stage, the surrounding residues with the transition-state model are optimized. The last stage is evaluation and ranking the results (Richter et al. 2011). The other application of PyRosetta is ligand docking, which is mainly used for virtual screening in enzyme engineering and is the key step of drug design development.



**Figure 11:** Flowchart of the *de novo* enzyme design (Richter et al. 2011).

## 2 RESEARCH PROPOSAL

Unlike the carbohydrate fraction in lignocellulosic biomass, lignin has mostly been treated as technical waste during the commercial production of cellulosic ethanol over the last few years. Even though it is the second most abundant organic substance in the biosphere and contains numerous aromatic compounds. Until today lignin has been mainly utilized for the production of valuable macromolecules, such as carbon fibers. Only a few examples of lignin-based fine chemicals have been reported, such as vanillin. Therefore, the construction of the efficient valorization strategy to generate the low molecular weight and high-value chemicals from lignin remains a major challenge. Considering the growing mortality rates of cancer, the synthesis of chiral  $\beta$ -aromatic amino acids, which are usually considered as a precursor for corresponding drugs, has been of particular interest to researchers in the field of pharmaceutical production.



**Figure 12:** The main goal of this doctoral research is the computational design of TchPAM for the production of value-added lignin-based aromatic amino acids, such as  $\beta$ -tyrosine, from monolignol: *trans-p*-hydroxycinnamic acid.

The aim of this thesis is the enzymatic conversion of the monolignol *trans-p*-hydroxycinnamic acid into the value-added optically pure  $\beta$ -tyrosine (Figure 12). At first, based on the modified hydantoinase process, a strategy with chemical synthesis and two-step enzymatic conversion has been investigated to reach this goal. However, the production yield of OH-PheDU in the first chemical synthesis step was too low, and

---

no enzyme has been found to convert *N*-carbamoyl- $\beta$ -phenylalanine to  $\beta$ -phenylalanine. As an alternative strategy, the production of  $\beta$ -tyrosine from monolignol by computational designed phenylalanine aminomutase has been pointed out. With the development in algorithms of computational protein design, a wide range of engineered ammonia lyases has been reported to expand their substrate scope or enhance enzyme activity in the past decade. Among them, the most studied member of this enzyme family is aspartase, which can catalyze the ammonia addition reaction without MIO moiety (Asano et al. 2005). Based on the computational approach of de novo protein design, Wu and co-workers have redesigned a series of aspartase variants from *Bacillus sp.*. These variants are able to produce the enantiopure aliphatic, polar, or aromatic  $\beta$ -amino acids, such as  $\beta$ -phenylalanine (R. Li et al. 2018) (Cui et al. 2021). However, it is more difficult to design MIO-dependent enzymes through a similar computational approach method, since the MIO moiety, as a noncanonical amino acid part in the active center of the enzyme, is not identified by the computational program. The researchers usually focused on the space of the aromatic binding site of MIO-dependent enzymes for expansion of substrate scope, such as the introduction of methoxy group on the different substituted position of the aromatic ring (Filip et al. 2018) (Zhu et al. 2018) (Nagy et al. 2019). One of the challenging points of this project is how to define the MIO moiety in the active center of TchPAM during the enzyme design *in silico*. To pursue the strategy of  $\beta$ -tyrosine production with designed TchPAM, the following research packages were established:

- Selection of the enzyme from the suitable organism as an enzyme template.
- Determination of the residues in the active center of the enzyme template
- Obtainment of a mutant library based on results from computational enzyme design
- Expression and purification of the corresponding designed enzymes
- Determination of the enzyme activity for mutants by using the natural substrate  $\alpha$ -phenylalanine as well as the target substrate  $\beta$ -tyrosine
- Chiral analysis of the produced aromatic amino acids from designed mutants

### 3 MATERIAL AND METHODS

#### 3.1 Materials

##### 3.1.1 Chemicals

**Table 2:** List of all chemicals used in this work

Chemical	Supplier
( <i>R</i> )- $\beta$ -phenylalanine	PepTech Corporation
( <i>R</i> )- $\beta$ -tyrosine	PepTech Corporation
( <i>S</i> )- $\beta$ -phenylalanine	PepTech Corporation
( <i>S</i> )- $\beta$ -tyrosine	PepTech Corporation
	Aber GmbH
1,4-Dithiothreit (DTT)	Carl Roth GmbH & Co. KG
3-amino-3-(4-hydroxyphenyl)-propionic acid	Aber GmbH
4-(Dimethylamino-)benzaldehyd	Merck
Acetic acid	Carl Roth GmbH & Co. KG
Acetone	
Acrylamid	Carl Roth GmbH & Co. KG
Agar	Sigma Aldrich
Agarose	Carl Roth GmbH & Co. KG
Ammoniumperoxodisulfat	Carl Roth GmbH & Co. KG
Ammonia	
Ammonium sulfate	Carl Roth GmbH & Co. KG
Ampicillin sodium salt	Carl Roth GmbH & Co. KG
BlueStar Plus Prestained Protein Marker	Nippon Genetics Europe GmbH
Boric acid	Carl Roth GmbH & Co. KG
Bromphenol blue	Sigma Aldrich
BugBuster 10X Protein Extraction Reagent	Merck
Chloroform	Carl Roth GmbH & Co. KG
Chloroform	Carl Roth GmbH & Co. KG
Coomassie Brilliant Blue R250	Carl Roth GmbH & Co. KG
DL- $\beta$ -Phenylalanine	Sigma Aldrich
Dimethyl sulfoxide (DMSO)	Carl Roth GmbH & Co. KG
Dipotassium phosphate	Sigma Aldrich
Disodium phosphate	Carl Roth GmbH & Co. KG
DNA Ladder (1 kb, Quick-Load®)	New England Biolabs
Ethanol	Carl Roth GmbH & Co. KG
Ethyl acetate	Carl Roth GmbH & Co. KG

Ethylenediaminetetraacetic acid (EDTA)	
Gel Loading Dye, Purple (6X)	New England Biolabs
Glucose	Carl Roth GmbH & Co. KG
Glycerine	Carl Roth GmbH & Co. KG
Glycin	Roth
Hydrochloric acid	Carl Roth GmbH & Co. KG
Imidazole	Roth
Isopropanol	Carl Roth GmbH & Co. KG
Isopropyl- $\beta$ -D-thiogalactopyranosid	Carl Roth GmbH & Co. KG
Kanamycin sulfate	Carl Roth GmbH & Co. KG
L-phenylalanine	Sigma Aldrich
L-tyrosine	Sigma Aldrich
Magnesium chloride	Carl Roth GmbH & Co. KG
Magnesium sulfat Heptahydrat	Carl Roth GmbH & Co. KG
Methanol	Carl Roth GmbH & Co. KG
Monosodium phosphate	Carl Roth GmbH & Co. KG
N,N-Dimethylformamide	Carl Roth GmbH & Co. KG
n-Butanol	
n-Hexane	Carl Roth GmbH & Co. KG
Nickel (II) chloride Hexahydrate	Sigma Aldrich
<i>N</i> -isobutyryl-L-Cystein (IBLC)	Sigma Aldrich
Nuclease free water	Thermo Fisher Scientific Inc.
ortho-Phthaldialdehyde (OPA)	Sigma Aldrich
Phosphoric acid	Carl Roth GmbH & Co. KG
Potassium chloride	Carl Roth GmbH & Co. KG
Potassium cyanate	Fluka
Roti®-Gel Stain	Carl Roth GmbH & Co. KG
Sodium chloride	Carl Roth GmbH & Co. KG
Sodium dodecyl sulfate (SDS)	Carl Roth GmbH & Co. KG
Sodium hydroxide	Carl Roth GmbH & Co. KG
Sulfuric acid	
Tetramethylethylendiamin (TEMED)	Carl Roth GmbH & Co. KG
Trans-cinnamic acid	Sigma Aldrich
Trans-p-coumaric acid	Sigma Aldrich
Tris-hydrochlorid	Carl Roth GmbH & Co. KG
Tryptone	Carl Roth GmbH & Co. KG
Urea	Carl Roth GmbH & Co. KG
Yeast extract	Carl Roth GmbH & Co. KG

### 3.1.2 Enzymatic Kits, plasmids, and microorganisms

**Table 3:** Commercial enzymes, plasmids, and microorganisms

<b>Name</b>	<b>Supplier</b>
Benzonase	Merck
Lysozyme	Sigma Aldrich
<i>E. coli</i> C2566	New England BioLabs Inc.
Monarch® Plasmid Miniprep Kit	New England BioLab Inc.
Primer	Thermo Fisher Scientific Inc.
Q5 Hot Start High Fidelity 2x Master Mix	New England BioLab Inc.
Q5 site-directed Mutagenesis Kit	New England BioLab Inc.
pam gene in pET-28a plasmid	ProteoGenix
Anza Restriction enzyme	Thermo Fisher Scientific Inc.
DNA Extraction Kit	New England BioLab Inc.
Monarch® DNA Gel Extraction Kit	New England BioLab Inc.
DpnI	Thermo Fisher Scientific Inc.

### 3.1.3 Media

**Table 4:** The list of used commercial media

<b>Medium</b>	<b>Supplier</b>
SOC Medium	New England BioLab Inc.
AIM	Formedium, UK

**Table 5:** Supplementary materials in media

<b>Name</b>	<b>Composition</b>	<b>Final concentration</b>
Ampicillin	Ampicillin stock solution, 100 mg/ml	100 µg/ml
Kanamycin	Kanamycin stock solution, 50 mg/ml	50 µg/ml
IPTG	1 M IPTG stock solution	0.1 mM – 1 mM

**Table 6:** The list of the medium constitution

<b>Medium</b>	<b>Constitution</b>	<b>Concentration</b>
LB-Medium (lysogeny broth)	Tryptone	10 g/l
	Yeast extract	5 g/l
	NaCl	10 g/l
LB-Agar	pH 7.2, dissolved in water, autoclaving	
	Agar in LB-Medium	15 g/l
	Amp or Kan was added in autoclaved LB-Agar	
TB-Medium (terrific broth)	Tryptone	12 g/l
	Yeast extract	24 g/l
	Glycerol	4 g/l
	KH <sub>2</sub> PO <sub>4</sub>	0.17 M
	K <sub>2</sub> HPO <sub>4</sub>	0.72 M
SOC-Medium (super optimal broth)	pH 7.2, dissolved in water, autoclaving	
	Tryptone	20 g/l
	Yeast extract	5 g/l
	NaCl	0.5 g/l
	1M KCl (pH 7)	2.5 g/l
	MgCl <sub>2</sub>	2.5 g/l
	MgSO <sub>4</sub>	20 g/l
	Glucose	1 M
	Sterile filtration	
AIM (Auto Induction Medium)	Tryptone	10 g/l
	Yeast extract	5 g/l
	(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	3.3 g/l
	KH <sub>2</sub> PO <sub>4</sub>	6.8 g/l
	Na <sub>2</sub> HPO <sub>4</sub>	7.1 g/l
	MgSO <sub>4</sub>	0.24 g/l
	Glycerol	0.2 % (v/v)
	Glucose	0.5 g/l
	Lactose	2 g/l
	dissolved in water, autoclaving	

### 3.1.4 Buffer and solutions

**Table 7:** Solution for protein purification

Name	Composition	Concentration
Tris-Buffer for cells (pH 8.0)	Tris	20 mM
	NaCl	50 mM
Binding Buffer (pH 8.0)	Tris	20 mM
	NaCl	500 mM
	Imidazole	10 mM
Washing Buffer (pH 8.0)	Tris	20 mM
	NaCl	500 mM
	Imidazole	20 mM
Elution Buffer (pH 8.0)	Tris	20 mM
	NaCl	500 mM
	Imidazole	300 mM
Ending Buffer (pH 8.0)	Tris	20 mM
	NaCl	500 mM
	Imidazole	500 mM
Tris-Buffer for activity assay (pH 8.8)	Tris	100 mM
Ammonium sulfate Buffer (pH 10)	Ammonia solution	100 ml
	Concentrated sulfuric acid	Titrated until pH 10

**Table 8:** Solution for thin-layer-chromatography

Name	Composition	Concentration
UV-TLC	n-hexane	8 ml
	Ethyl acetate	2 ml
	Acetic acid	100 µl
Ninhydrin-TLC	Ninhydrin	1.5 g
	n-butanol	100 ml
	Acetic acid	3 ml

**Table 9:** Buffers and solution for agarose- and SDS-PAGE gels

<b>Name</b>	<b>Composition</b>	<b>Concentration</b>
SDS sample buffer	Tris-HCl (pH 6.8)	225 mM
	Glycerine	50 % (v/v)
	DTT	250 mM
	SDS	5 % (w/v)
	Bromphenol blue	0.05 % (w/v)
SDS-PAGE gel		
Separation gel (12.5 %)	1.88 M Tris-HCl (pH 8.8)	2.4 ml
	10 % SDS	120 $\mu$ l
	Water	4.48 ml
	10 % APS	60 $\mu$ l
	TEMED	10 $\mu$ l
	30 % Acrylamide	5 ml
Stacking gel (5 %)	0.625 M Tris-HCl (pH 6.8)	800 $\mu$ l
	10 % SDS	40 $\mu$ L
	Water	2.5 ml
	10 % APS	20 $\mu$ l
	TEMED	4 $\mu$ l
	30 % Acrylamide	660 $\mu$ l
Coomassie blue staining solution for SDS gels	Coomassie blue G250	1.5 g
	Ethanol	455 ml
	Acetic acid	75 ml
	Water	Filled up to 1 l
TGS SDS-gel buffer (10x)	Tris-HCl	50 mM
	Glycine	1.92 M
	SDS	1 %
TAE agarose gel buffer (10x) (pH 8.0)	Tris-HCl (pH 8.0)	0.4 M
	Acetic acid	0.2 M
	EDTA	10 mM
Agarose gel (0.8 %)	Agarose	0.4 g
	TAE buffer (1x)	50 ml
	Roti® Gel Stain	2.5 $\mu$ l

### 3.1.5 Solutions for HPLC

**Table 10:** Solution for HPLC

Name	Composition	Concentration
NaB Buffer (pH 10.4)	Boric acid	133 mM
OPA/IBLC	OPA in MeOH	100 mM
	IBLC in NaB Buffer	133 mM
	Fresh prepared, 1:1 mixture of OPA and IBLC solution	
Acetic acid solution	Acetic acid in water	3 %
NaPP Buffer (pH 6.5)	Na <sub>2</sub> HPO <sub>4</sub> /NaH <sub>2</sub> PO <sub>4</sub>	40 mM
HPLC for amino acids		
Mobile Phase A	NaPP Buffer	
Mobile Phase B	MeOH	
HPLC for OH-PheDU		
Mobile Phase A	Phosphoric acid	0.1 %
Mobile Phase B	MeOH	

### 3.1.6 Primer

**Table 11:** Primer list of variants Tyr424 by using KLD treatment

Mutant	Primers		
Tyr424Cys	69	fw	GAG CGT TGA <b>TTG</b> <b>CGG</b> CCT GAA AGG
	66	rw	AGA TCC GGA CCC AGA CTC
Tyr424Asp	67	fw	GAG CGT TGA <b>TGA</b> <b>TGG</b> CCT GAA AG
	66	rw	AGA TCC GGA CCC AGA CTC
Tyr424Met*	74	fw	GAG CGT TGA <b>TAT</b> <b>G</b> GG CCT GAA AGG TCT GGA TAT TG
	66	rw	AGA TCC GGA CCC AGA CTC
Tyr424Asn	67	fw	GAG CGT TGA <b>TAA</b> <b>TGG</b> CCT GAA AG
	66	rw	AGA TCC GGA CCC AGA CTC
Tyr424Gln	69	fw	GAG CGT TGA <b>TCA</b> <b>AGG</b> CCT GAA AGG
	66	rw	AGA TCC GGA CCC AGA CTC

\* The mutant Tyr424Met was not successfully collected from the single colony in the LB-agar plate.

**Table 12:** Primer list of Touch-Down PCR products

Mutant	Primers	
Phe86Tyr	fw	TGT GAC CAC CGG <b>TTA</b> TGG TG
	rw	ACC GGT GGT CAC ACC ATA G
Leu104Ser	fw	CTG CAG GAA AGT <b>AGC</b> ATT CGC TGT CT
	rw	ACT TTC CTG CAG TTC GCT CAG G
Leu104Thr	fw	CTG CAG GAA AGT <b>ACC</b> ATT CGC TGT CT

	rw	ACT TTC CTG CAG TTC GCT CAG G
Leu108Gln	fw	CTG ATT CGC TGT CAA CTG GCA G
	rw	ACA GCG AAT CAG ACT TTC CTG CA
Leu108Ser	fw	CTG ATT CGC TGT AGC CTG GCA G
	rw	ACA GCG AAT CAG ACT TTC CTG CA
Leu179Ser	fw	GCA AGC GGC GAT AGC ATT CCG
	rw	ATC GCC GCT TGC GCT AAC A
Leu179Thr	fw	GCA AGC GGC GAT ACC ATT CCG
	rw	ATC GCC GCT TGC GCT AAC A
Leu427Ala	fw	GAT TAT GGC CTG GCT GGT CTG GAT ATT
	rw	CAG GCC ATA ATC AAC GCT CAG A
Lys427Cys	fw	GAT TAT GGC CTG TGC GGT CTG GAT ATT
	rw	CAG GCC ATA ATC AAC GCT CAG A
Lys427Glu	fw	GAT TAT GGC CTG GAA GGT CTG GA
	rw	CAG GCC ATA ATC AAC GCT CAG A
Lys427Met	fw	GAT TAT GGC CTG ATG GGT CTG GA
	rw	CAG GCC ATA ATC AAC GCT CAG A
Lys427Asn	fw	ATC TGA GCG TTG ATA ATG GCC TG
	rw	CAG GCC ATA ATC AAC GCT CAG A
Lys427Gln	fw	GAT TAT GGC CTG CAG GGT CTG GA
	rw	CAG GCC ATA ATC AAC GCT CAG A
Ile431Glu	fw	TGA AAG GTC TGG ATG AAG CAA TGG C
	rw	ATC CAG ACC TTT CAG GCC ATA ATC
Ile431Asn	fw	TGA AAG GTC TGG ATA ATG CAA TGG C
	rw	ATC CAG ACC TTT CAG GCC ATA ATC
Glu455Asp	fw	TGCA TAG TGC AGA TCA GCA TAA TCA GGA
	rw	GCA CTA TGC ACA TGG GTG GTC
Glu455His	fw	TGC ATA GTG CAC ATC AGC ATA ATC AGG A
	rw	GCA CTA TGC ACA TGG GTG GTC
Glu455Asn	fw	TGT GCA TAG TGC AAA TCA GCA TAA TCA GGA
	rw	GCA CTA TGC ACA TGG GTG GTC
Asn458Ala	fw	GCA GAA CAG CAT GCA CAG GAT ATT AAT A
	rw	ATG CTG TTC TGC ACT ATG CAC A
Asn458Cys	fw	GCA GAA CAG CAT TGC CAG GAT ATT
	rw	ATG CTG TTC TGC ACT ATG CAC A
Gln459Cys	fw	AGA ACA GCA TAA TTG CGA TAT TAA TAG TCT G
	rw	ATT ATG CTG TTC TGC ACT ATG CAC
Gln459Leu	fw	AGA ACA GCA TAA TCT GGA TAT TAA TAG T
	rw	ATT ATG CTG TTC TGC ACT ATG CAC
Gln459His	fw	AGA ACA GCA TAA TCA CGA TAT TAA TAG T
	rw	ATT ATG CTG TTC TGC ACT ATG CAC

### 3.1.7 Software developer

**Table 13:** List of used software and web services

**BLAST**

**Genome Compiler**

**PDB**

**MestReNova**

**PyMOL**

**SnapeGene® Viewer**

**UniProt**

**PyRosetta**

**Chimera**

**OpenBabel 2.4**

**Gaussian09**

**Biology and Chemistry Library Project (BCL)**

**MAFFT version 7.3**

**Gblocks**

**IQ-TREE version 1.4**

**PhyML SMS servers**

**Minimal ancestor deviation approach (MAD)**

**Interactive Tree of Life (iTol)**

### 3.2 Instruments

**Table 14:** Devices

<b>Device</b>	<b>Type</b>	<b>Supplier</b>
Autoclave	V-150	Systec, Wetztenberg, Germany
	HX-430	Thermo Fisher Scientific, US
Balance	BP3100S	Sartorius AG, Germany
	BP61S	Sartorius AG, Germany
	Entris®	Sartorius AG, Germany
	PR2003 DeltaRange	Mettler-Toledo
Centrifuge	Megafuge 40R	Thermo Fisher Scientific Inc.
	5415D Rotor: F45-24-11	Eppendorf AG, Germany
	Heraeus™ Pico™ 17	Thermo Fisher Scientific Inc.
	Heraeus Multifuge X3 FR	Thermo Fisher Scientific Inc.
Clean bench	Maxisafe 2020	Thermo Fisher Scientific Inc.
Electrophoresis	Biometra P35T 1000/500	Analytik Jena AG
	Minigel-Twin Power P25	Analytik Jena AG
Fast protein liquid chromatography (FPLC)	Äkta™ Start	GE Healthcare Life Sciences

FPLC column	HisTrap™ High Performance (1 ml)	GE Healthcare Life Sciences
Flat shaker	Polymax 1040	Hedolph Instruments GmbH & CO. KG
Gel documentation system	GP-05LED	NIPPON Genetics Europe
High-pressure liquid chromatography (HPLC)	1100 Series	Agilent Technologies, Germany
	1200 Series	
	1290 Infinity II LC System	
HPLC columns	HyperClone™ 5 µm ODS (C18) 120 Å, LC Column 150 x 4.6 mm	Phenomenex
HPLC precolumn	SecurityGuard™ cartridges for C18 HPLC columns with 3.2 to 8.0 mm internal diameters, 10/Pk	Phenomenex
Incubator	BBD 6220	Heraeus Deutschland GmbH & Co. KG
Incubator with shaking	Multitron	Infors HT, Switzerland
Liquid handling system	Liquid Handling Station	BRANDTech Scientific, Inc, Germany
Lyophilization systems		Martin Christ Gefriertrocknungsanlagen, Germany
Magnetic stirrer	RCT basic	IKA®-Werke GmbH & Go. KG
Melting point apparatus	Fine control	Electrothermal Engineering Ltd
Microwave		Severin
		Siemens
Microwave Reactor	Monowave 400	Anton Paar GmbH
Microwave pressure tube	G10	Anton Paar GmbH
Micro-Volume-Plate	Take 3	BioTek Instruments, Inc.
MilliQ water Ultrapure Lab Water System	Purelab flex	ELGA LabWater
pH-meter	inoLab	Xylem Inc.
	pH Level 1	Xylem Inc.

Pipettes	Research plus	Eppendorf AG
Photometer for cuvettes	Novaspec II	Amersham Biosciences Europe
Photometer for microtiter plates	Infinite® 200 PRO	Tecan Trading AG
Pressure tube	15 ml/300 ml	Ace Glass, Inc.
Purification System	Reveleris® PREP	Büchi Labortechnik AG, Switzerland
Pure Chromatography-System Column	4 g Silica Flash Cartridge	Büchi Labortechnik AG, Switzerland
Rotator	Intelli mixer	neoLab Migge GmbH
	SB3	Stuart Equipment, UK
Rotary Evaporator		Heidolph Instruments
SDS gel supply	Mini-PROTEAN Tetra	Bio-Rad Laboratories, Inc. Cell
Spectrophotometer	Epoch	BioTek Instruments, Inc.
Thermocycler for PCR	MJ Research PTC-200 Gradient Thermal Cycler	Marshall Scientific LLC
Thermo shaker	Thermomixer compact	Eppendorf AG, Germany
	Thermomixer comfort	Eppendorf AG, Germany
	BioShake iQ	Analytik Jena, Germany
Ultrasonic bath	Sonorex Super RK 100	Bandelin electronic GmbH & Co. KG
Ultrasonic homogenisator	Sonopuls UW3100	Bandelin electronic GmbH & Co. KG
Vacuum pump		Vacuubrand GmbH & Co. KG
Vortex	Vortex Genie 2	VWR
	Vortex Mixer	NeoLab Migge GmbH

**Table 15:** List of used supplies

Supply	Type	Supplier
His Spintrap columns	TALON®	GE Healthcare Life Sciences
Silica TLC-plates	Alugram® SIL G	Macherey-Nagel
Syringe filter	Rotilabo®, 0.2 µm	Carl Roth GmbH & Co. KG
Silica gel on TLC Al foils	Silica gel matrix, with fluorescent indicator 254 nm	Sigma Aldrich
Protein Concentrator	Pierce™ PES, 10K MWCO, 0.5 µL	Thermo Fisher Scientific Inc.

---

### 3.3 Methods

#### 3.3.1 Chemical synthesis of 6'-(4-hydroxyphenyl)-5,6-dihydrouracil (OH-PheDU)

##### 3.3.1.1 Preparation of the standard of OH-PheDU with $\beta$ -tyrosine and urea

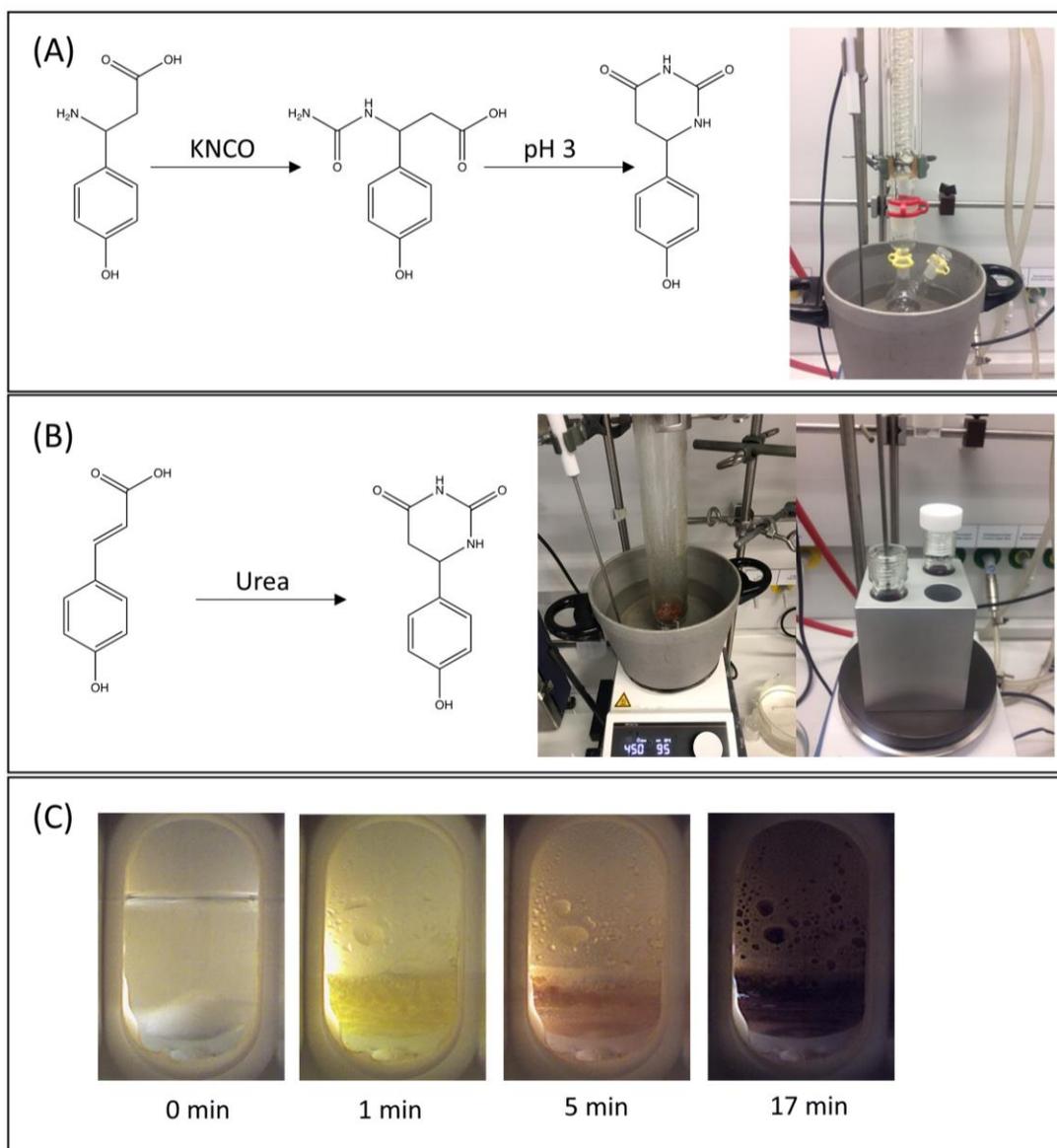
Based on the protocol by Azaryan et.al,  $\beta$ -tyrosine and urea were mixed in a molecular ratio of 1:3. The reaction was conducted with a reflux condenser in an oil bath at 145 °C for 30 min. Until cooling to room temperature, the formed solid mixture was taken up with 10 % aqueous NaOH and decolorized by shaking with active charcoal. After elimination of the activated charcoal with filters, HCl was added to pH 3. The synthesized crystals were filtered from the acid solution, washed with cold water, and resolved in methanol (Azaryan et al. 2010).

The other similar protocol came from the patent WO 2004/11448. Instead of the solid-solid reaction, trimethylamine was used as the solvent during the reaction of the  $\beta$ -tyrosine and urea. This reaction was conducted for 4 h and the product was resolved in boiling water.

##### 3.3.1.2 Preparation of the standard of OH-PheDU with $\beta$ -tyrosine and KNCO

To synthesize the OH-PheDU, the *N*-carbamoyl- $\beta$ -tyrosine was prepared as starting material, which was formed by heating the mixture solution in a two-necked flask with a reflux condenser at 90 °C (Figure 13 (A)). For the preparation of the mixture solution, 200 mg  $\beta$ -tyrosine and 113,6 mg KNCO were solved in 3 ml and 6 ml distilled hot water, respectively. After 1 h, the *N*-carbamoyl- $\beta$ -tyrosine was produced, which was crystallized by cooling at 4 °C. For the subsequent synthesis of the OH-PheDU, the *N*-carbamoyl- $\beta$ -tyrosine solution was acidified by titration of the concentrated HCl until the pH value decreased to 3. Finally, the white precipitate was slowly formed by cooling

at 4 °C, which was collected by a filter, washed with distilled water, as well as dried and stored in a desiccator.



**Figure 13:** Chemical synthesis reactions and methods. (A) OH-PheDU was synthesized in a two-necked flask from  $\beta$ -tyrosine. (B) OH-PheDU was synthesized in pressured glass tubes of different sizes from *p*-hydroxycinnamic acid. (C) The monitor of chemical synthesis in a microwave-assisted reactor at different times.

### 3.3.1.3 Synthesis of the OH-PheDU from *trans-p*-hydroxycinnamic acid

#### Synthesis of OH-PheDU in pressure tubes

According to the protocols for the synthesis of the PheDU, the monolignol *trans-p*-

---

hydroxycinnamic acid was used as the starting material, instead of the cinnamic acid. Considering the high melting point of the *trans-p*-hydroxycinnamic acid (214 °C), the melting point of the eutectic mixture with urea was measured for reduction of the reaction temperature. The mixture of *trans-p*-hydroxycinnamic acid and urea in various proportions was grinded and mixed before heating in a glass sealed tube (Ace Glass Inc, USA). The 300 ml pressure tube was heated in an oil bath, or the 15 ml pressure tubes were heated in a metal block (Figure 13 (B)). To achieve, the best production yield, the optimization of the process was performed with different reaction parameters, such as temperature, time, and extracted solvent.

### **Synthesis of OH-PheDU through the microwave-assisted method**

Subsequently, the microwave-assisted method was used to reduce the reaction temperature and time. The grinded mixture of *trans-p*-hydroxycinnamic acid and urea was reacted in the G10 glass tube, which fits in the Monowave 400 (Anto Paar, Ostfildern, Germany). After heating to the desired temperature in 2 min, both compounds were reacted under microwave irradiation with controlling the reaction state and pressure (Figure 13(C)). Finally, the products were extracted by hot water or methanol, which were dried by lyophilizer or rotary evaporator.

#### **3.3.1.4 Purification of the extracted OH-PheDU**

To obtain the OH-PheDU with high purity, the undesired products were removed by using liquid chromatography with a UV-light detector (BÜCHI, Germany). The products were redissolved in an organic solvent, containing the n-hexane and ethyl acetate with a proportion in 8:2. Then around 2.5 ml, liquid products were injected through a silica cartridge (4 g silica bed weight, Reveleris™, Germany) into the chromatography system. The mobile phase was n-hexane (solvent A) and ethyl acetate (solvent B). The flow rate was set at 15 ml/min and the isolated products were detected at 210 and 254 nm. After purification, the OH-PheDU in organic solvent was removed

---

by a rotary evaporator.

### **3.3.1.5 Characterization of OH-PheDU and *N*-carbamoyl- $\beta$ -tyrosine**

#### **Dissolving analysis**

To improve the production yield, the solvent screening of OH-PheDU and *N*-carbamoyl- $\beta$ -tyrosine was conducted to select the suitable solvent for the extraction of the products. The candidate solvents were chloroform, methanol, cooled water, and ethyl acetate. In the beginning, 1 mg OH-PheDU or *N*-carbamoyl- $\beta$ -tyrosine were weighted in a 15 ml falcon tube. Then, 1 ml solvent was pipetted into the tube and mixed by vortex mixer. If the OH-PheDU or *N*-carbamoyl- $\beta$ -tyrosine was insoluble in the corresponding solvent with the concentration of 1 mg/ml, the other 4 ml or 9 ml solvents were added into the tube to attempt to dissolve the products at 0.2 or 0.1 mg/ml.

#### **Melting point measurement**

To measure the melting point, the OH-PheDU and *N*-carbamoyl- $\beta$ -tyrosine were put into a glass capillary with a sealed bottom. The height of the samples was 2-3 mm in the glass capillary. Then, the capillary was inserted into the melting point apparatus (Electrothermal Engineering Ltd, England). By monitoring the sample state through the viewfinder and the thermometer, the samples were heated at slowly increased temperatures. Once the appearance of the visible drop in the samples, the temperature was recorded and carefully adjusted until the samples were completely melted.

#### **Detection by Thin-Layer-Chromatography (TLC)**

To prepare the samples for TLC analysis, the methanol was pipetted in the tube with 10 mg OH-PheDU, *N*-carbamoyl- $\beta$ -tyrosine, or  $\beta$ -tyrosine until the chemicals were completely dissolved. 2  $\mu$ l samples were pipetted on the silica plates, which contained the fluorescence indicator. Subsequently, the TLC plate was eluted with the mixture of

---

hexane, ethyl acetate, and acetate with a volume ratio of 5:5:0.1. Finally, the plate was dried at room temperature and observed under a UV lamp with a 254 nm wavelength. Moreover, the eluent with ninhydrin was also prepared to identify the compounds. Herein, 2  $\mu$ l of samples, which were dissolved in methanol, were pipetted on the silica plate (DC ready-made films, Alugram® SIL G, Macherey-Nagel, Germany). After elution, the plate was stained with ninhydrin staining solution (Table 8) and air-dried.

### **High-Performance Liquid Chromatography (HPLC) for chemical synthesis**

The production yield of the OH-PheDU and *N*-carbamoyl- $\beta$ -tyrosine was determined through detection at 210 nm by HPLC (Agilent 1100 or 1200 series, Germany). Before injection in the C18 HPLC column (5  $\mu$ m, 4.6 x 150 mm, Hyperclone, Phenomenex, USA), the samples were dissolved in methanol, filtered, and centrifuged. The injection volume of samples was 5  $\mu$ l. For analysis of the compounds, an isocratic elution with 30 % methanol and 70 % H<sub>3</sub>PO<sub>4</sub> solution (0.1 % v/v, pH 3) and the flow rate of 1 ml/min were operated.

### **Nuclear Magnetic Resonance and Mass Spectroscopy**

The nuclear magnetic resonance (<sup>1</sup>H-NMR) spectroscopy spectra were used for the determination of the synthesized products. 10 mg purified OH-PheDU and *N*-carbamoyl- $\beta$ -tyrosine were dissolved in 0.5 ml deuterated DMSO, which were measured by the 300 NMR instrument. Furthermore, mass spectrometry (MS) was conducted with electrospray ionization (ESI) to verify the products. The MS was performed on a quadrupole Q Exactive Plus with positive mode. Both NMR and MS were measured at the institute of organic chemistry (KIT-IOC). The results from NMR and MS were analyzed by MestReNova (Willcott 2009).

---

### 3.3.2 Phylogenetic methods

#### 3.3.2.1 Multiple sequence alignment

According to the classification of the related enzymes, the amino acid sequences were identified from the SwissProt Knowledgebase (Bairoch et al. 2004). The multiple sequence alignments were carried out by using MAFFT version 7.3. and online tool Clustal Omega (Kato and Standley 2013) (Sievers and Higgins 2014). The resulting alignments were visualized and analyzed by the Jalview (Waterhouse et al. 2009).

#### 3.3.2.2 Phylogenetic analysis

The gaps were deleted from the resulting alignments by Gblocks. Meanwhile, the amino acid sequences, which were 100 % identified to the other related sequences were manually removed. The PhyML SMS servers were used to determine the best-fit substitution model for the construction of the phylogenetic tree (Lefort et al. 2017). Based on the results from the PhyML SMS, the LG+G model with 1000 bootstrap replicates number was operated. Subsequently, the maximum likelihood trees were built by using IQ-TREE version 1.4. (Minh et al. 2020). Then, the phylogenetic tree was rooted in a minimal ancestor deviation (MAD) approach (Tria et al. 2017). Finally, the online tool “interactive Tree of Life” (iTol) was used to visualize and analyze the resulting tree (Letunic and Bork 2019).

### 3.3.3 Computational enzyme design

The catalytically essential residues in the active center of TchPAM was defined with the following commons:

For definition of the residue Tyr80, which acted as the key position for activation:

CST::BEGIN

TEMPLATE:: ATOM\_MAP: 1 atom\_name: C3 C1 N1

TEMPLATE:: ATOM\_MAP: 1 residue3: BTR

TEMPLATE:: ATOM\_MAP: 2 atom\_type: OH

---

```
TEMPLATE:: ATOM_MAP: 2 residue1: Y
CONSTRAINT:: distanceAB: 3.20 0.3 150.0 0 0
CONSTRAINT:: angle_A: 115.8 5.0 100.0 360.0 1
CONSTRAINT:: angle_B: 110.3 5.0 100.0 360.0 1
CONSTRAINT:: torsion_A: -118.9 5.0 80.0 360.0 6
CONSTRAINT:: torsion_B: -67.8 2.0 80.0 360.0 6
CONSTRAINT:: torsion_AB: 162.2 25.0 80.0 360.0 3
```

```
CST::BEGIN
```

For definition of the residue Tyr322, which might affect the MIO moiety formation:

```
CST::BEGIN
```

```
TEMPLATE:: ATOM_MAP: 1 atom_name: N1 C1 C3
TEMPLATE:: ATOM_MAP: 1 residue3: BTR
TEMPLATE:: ATOM_MAP: 2 atom_type: OH
TEMPLATE:: ATOM_MAP: 2 residue1: Y
CONSTRAINT:: distanceAB: 2.90 0.2 150.0 0 0
CONSTRAINT:: angle_A: 82.5 2.0 100.0 360.0 1
CONSTRAINT:: angle_B: 160.3 2.0 100.0 360.0 1
CONSTRAINT:: torsion_A: 136.10 5.0 80.0 360.0 6
CONSTRAINT:: torsion_B: 133.4 5.0 80.0 360.0 6
CONSTRAINT:: torsion_AB: 80.4 5.0 80.0 360.0 3
```

```
CST::END
```

For definition of the residues Asn231, Gln319, Arg325, Asn355, and Asn458, which sorted in carboxyl binding pocket:

```
CST::BEGIN
```

```
TEMPLATE:: ATOM_MAP: 1 atom_name: O3 C8 C3
TEMPLATE:: ATOM_MAP: 1 residue3: BTR
TEMPLATE:: ATOM_MAP: 2 atom_type: NH2O
TEMPLATE:: ATOM_MAP: 2 residue1: N
```

---

CONSTRAINT:: distanceAB: 5.60 0.4 20.0 0 0  
CONSTRAINT:: angle\_A: 68.30 25.0 20.00 360.00 1  
CONSTRAINT:: angle\_B: 120.20 15.0 20.00 360.00 1  
CST::END

CST::BEGIN

TEMPLATE:: ATOM\_MAP: 1 atom\_name: O1 C8 C3  
TEMPLATE:: ATOM\_MAP: 1 residue3: BTR  
TEMPLATE:: ATOM\_MAP: 2 atom\_type: NH2O  
TEMPLATE:: ATOM\_MAP: 2 residue1: Q  
CONSTRAINT:: distanceAB: 2.4 0.5 50.0 0 0  
CONSTRAINT:: angle\_A: 141.5 2.0 20.00 360.0 1  
CONSTRAINT:: angle\_B: 131.1 2.0 20.00 360.0 1  
CST::END

CST::BEGIN

TEMPLATE:: ATOM\_MAP: 1 atom\_name: O1 C8 C3  
TEMPLATE:: ATOM\_MAP: 1 residue3: BTR  
TEMPLATE:: ATOM\_MAP: 2 atom\_name: NH2 CZ NE  
TEMPLATE:: ATOM\_MAP: 2 residue1: R  
CONSTRAINT:: distanceAB: 2.40 0.20 50.00 0 0  
CONSTRAINT:: angle\_A: 106.10 15.00 20.00 360.00 1  
CONSTRAINT:: angle\_B: 85.60 15.00 20.00 360.00 1  
CST::END

CST::BEGIN

TEMPLATE:: ATOM\_MAP: 1 atom\_name: O1 C8 C3  
TEMPLATE:: ATOM\_MAP: 1 residue3: BTR  
TEMPLATE:: ATOM\_MAP: 2 atom\_name: NH2 CZ NE

---

```
TEMPLATE:: ATOM_MAP: 2 residue1: R
CONSTRAINT:: distanceAB: 2.30 0.20 50.00 0 0
CONSTRAINT:: angle_A: 115.20 15.00 20.00 360.00 1
CONSTRAINT:: angle_B: 90.70 15.00 20.00 360.00 1
CST::END
```

```
CST::BEGIN
```

```
TEMPLATE:: ATOM_MAP: 1 atom_name: O1 C8 C3
TEMPLATE:: ATOM_MAP: 1 residue3: BTR
TEMPLATE:: ATOM_MAP: 2 atom_type: NH2O
TEMPLATE:: ATOM_MAP: 2 residue1: N
CONSTRAINT:: distanceAB: 3.90 0.5 20.0 0 0
CONSTRAINT:: angle_A: 116.1 2.0 20.00 360.00 1
CONSTRAINT:: angle_B: 149.9 2.0 20.00 360.00 1
CST::END
```

```
CST::BEGIN
```

```
TEMPLATE:: ATOM_MAP: 1 atom_name: O1 C8 C3
TEMPLATE:: ATOM_MAP: 1 residue3: BTR
TEMPLATE:: ATOM_MAP: 2 atom_type: NH2O
TEMPLATE:: ATOM_MAP: 2 residue1: N
CONSTRAINT:: distanceAB: 4.90 0.5 20.0 0 0
CONSTRAINT:: angle_A: 60.5 2.0 20.00 360.00 1
CONSTRAINT:: angle_B: 129.1 2.0 20.00 360.00 1
CST::END
```

To avoid the huge change of the conformation, the enzyme structure was relaxed with constraints in the following commons:

---

```
-relax:constrain_relax_to_start_coords
-relax:coord_constrain_sidechains
-ex1
-ex2
-use_input_sc
-flip_HNQ
-no_optH false
```

A file for rotation limitation of the MIO moiety was defined with the following commons:

```
<LIGAND_AREAS>
  <LigandArea name="docking_sidechain_Y" chain="Y" cutoff="6.0" add_nbr_radius="true"
all_atom_mode="true" minimize_ligand="0.1" />
  <LigandArea name="final_sidechain_Y" chain="Y" cutoff="6.0" add_nbr_radius="true"
all_atom_mode="true" />
  <LigandArea name="final_backbone_Y" chain="Y" cutoff="7.0" add_nbr_radius="true"
all_atom_mode="true" Calpha_restraints="0.01" />
</LIGAND_AREAS>
```

The residue specification file during the first round of enzyme design was shown in the following:

```
ALLAA
AUTO
start
* A NATAA
* B NATAA
* D NATAA
80 B EMPTY
```

---

322 B EMPTY  
175 B EMPTY  
86 B PIKAA FY  
104 B PIKAA KRLHDSTQYNE  
107 B PIKAA KRHDSTCQYNE  
108 B PIKAA KRLHDSTQYNE  
179 B PIKAA KRLHDSTQYNE  
424 D PIKAA KRHDSTQYNE  
427 D PIKAA KRHNSTQYED  
431 D PIKAA KRHISTNQYED  
455 B PIKAA KRHNQYEDST  
459 B PIKAA KRHNQYEDST

For design the surrounding residues, a program to limit the movement and rotation of ligand in the active center was wrote by using the scripting language RosettaScripts, which is shown in the following:

```
<ROSETTASCRIPTS>  
<SCOREFXNS>  
<ScoreFunction name="ligand_soft_rep" weights="ligand_soft_rep" />  
<ScoreFunction name="hard_rep" weights="ligandprime" />  
</SCOREFXNS>  
<TASKOPERATIONS>  
<DetectProteinLigandInterface name="design_interface" cut1="6.0" cut2="8.0" cut3="10.0"  
cut4="12.0" design="1" resfile="matching.resfile" />  
</TASKOPERATIONS>  
<LIGAND_AREAS>  
<LigandArea name="docking_sidechain_X" chain="X" cutoff="6.0" add_nbr_radius="true"  
all_atom_mode="true" minimize_ligand="10" />
```

---

```
<LigandArea name="final_sidechain_X" chain="X" cutoff="6.0" add_nbr_radius="true"
all_atom_mode="true" />
<LigandArea name="final_backbone_X" chain="X" cutoff="7.0" add_nbr_radius="false"
all_atom_mode="true" Calpha_restraints="0.3" />
</LIGAND_AREAS>
<INTERFACE_BUILDERS>
<InterfaceBuilder name="side_chain_for_docking" ligand_areas="docking_sidechain_X" />
<InterfaceBuilder name="side_chain_for_final" ligand_areas="final_sidechain_X" />
<InterfaceBuilder          name="backbone"          ligand_areas="final_backbone_X"
extension_window="3" />
</INTERFACE_BUILDERS>
<MOVEMAP_BUILDERS>
<MoveMapBuilder          name="docking"          sc_interface="side_chain_for_docking"
minimize_water="true" />
<MoveMapBuilder          name="final"          sc_interface="side_chain_for_final"
bb_interface="backbone" minimize_water="true" />
</MOVEMAP_BUILDERS>
<MOVERS>
single movers_X
<FavorNativeResidue name="favor_native" bonus="0.5" />
<StartFrom name="start_from_X" chain="X">
<Coordinates x="-31.9" y="7.5" z="-5.5" />
</StartFrom>
<CompoundTranslate          name="compound_translate"          randomize_order="false"
allow_overlap="false">
<Translate chain="X" distribution="gaussian" angstroms="0.5" cycles="50" />
</CompoundTranslate>
<Rotate name="rotate_X" chain="X" distribution="uniform" degrees="360" cycles="500" />
```

---

```
<SlideTogether name="slide_together" chains="X" />

<HighResDocker      name="high_res_docker"      cycles="6"      repack_every_Nth="3"
scorefxn="ligand_soft_rep" movemap_builder="docking" />

<PackRotamersMover      name="design_interface"      scorefxn="hard_rep"
task_operations="design_interface" />

<FinalMinimizer name="final" scorefxn="hard_rep" movemap_builder="final" />

<InterfaceScoreCalculator name="add_scores" chains="X" scorefxn="hard_rep" />

compound movers

<ParsedProtocol name="low_res_dock">

<Add mover_name="start_from_X" />

<Add mover_name="compound_translate" />

<Add mover_name="rotate_X" />

<Add mover_name="slide_together" />

</ParsedProtocol>

<ParsedProtocol name="high_res_dock">

<Add mover_name="high_res_docker" />

<Add mover_name="final" />

</ParsedProtocol>

</MOVERS>

<PROTOCOLS>

<Add mover_name="favor_native" />

<Add mover_name="low_res_dock" />

<Add mover_name="design_interface" />

<Add mover_name="high_res_dock" />

<Add mover_name="add_scores" />

</PROTOCOLS>

</ROSETTASCRIPTS>
```

### 3.3.4 Molecular biology methods

#### 3.3.4.1 Site-directed Mutagenesis

Based on the sequence of the protein from *Taxus chinensis*, the pET28a plasmid with the *pam* gene was synthesized by ProteoGenix (France) after codon optimization. The plasmid was transformed into the T7 express competent *E. coli* C2566 (NEB, USA) and prepared as the cryo-culture. After Sanger sequencing by GATC (Germany), the extracted plasmid was diluted in 10-20 ng/ $\mu$ l and used as the template for subsequent single point mutagenesis. The reagents were pipetted in a PCR tube with the following list in

Table 16. The thin-walled PCR tubes were gently centrifugated for 3 s to collect the solution on the tube wall.

Table 16: PCR batch with Q5® Hot Start High-Fidelity 2X Master Mix

Reagents	PCR samples	Negative control
Q5 Hot Start High-Fidelity 2X Master Mix	12.5 $\mu$ l	12.5 $\mu$ l
10 $\mu$ M Forward Primer	1.25 $\mu$ l	1.25 $\mu$ l
10 $\mu$ M Reverse Primer	1.25 $\mu$ l	1.25 $\mu$ l
Template DNA	1 $\mu$ l	0 $\mu$ l
Nuclease-free water	9 $\mu$ l	10 $\mu$ l
Final volume	25 $\mu$ l	25 $\mu$ l

For amplification of the DNA, touch down PCR was run with the protocol shown in Table 17, except for the mutants Tyr424Asn, Tyr424Asp, Tyr424Cys, and Tyr424His, which were prepared based on the protocol of Q5® Hot Start High-Fidelity 2X Master Mix (Table 18, NEB, England).

**Table 17:** Touch-Down PCR protocol

Step	Temperature [°C]	Duration [s]
Pre-heating	95	
Initial denaturation	98	30
10 Cycles	98	10
	72*	30
	72	228
15 Cycles	98	10
	62	30
	72	228
Final extension	72	120
Hold	4	

\* A decreasing gradient of annealing temperature by 1 °C per cycle was performed until 62 °C was reached in the 10<sup>th</sup> cycle.

**Table 18:** PCR protocol with Q5® Hot Start High-Fidelity 2X Master Mix

Step	Temperature [°C]	Duration [s]
Pre-heating	95	
Initial denaturation	98	30
25 Cycles	98	10
	Depending on primers	30
	72	240
Final extension	72	120
Hold	4	

Agarose gel electrophoresis was applied to identify the size of amplified DNA. 0.4 g agarose was dissolved in 50 mL 1X TBE-buffer by heating in a microwave oven. After cooling to around 60 °C, 2.5 µl of Roti® Gel Stain was added in a 0.8 % (w/v) agarose

---

solution. 5  $\mu$ l PCR product were mixed with 2.5  $\mu$ l of Gel Loading Dye (Purple, NEB, England) and pipetted in a pocket of the hardening gel. Meanwhile, a 7.5  $\mu$ l of DNA Ladder (1 kb, NEB, England) was used as a marker. The electrophoresis was performed at 80 V by using the 1 X TBE-buffer as a running buffer. Then, the size of the DNA in the agarose gel was determined by detection with a gel documentation system (NIPPON Genetics Europe GmbH, Germany). The desired DNA band on the gel would be cut out if necessary and extracted with Monarch® DNA Gel Extraction Kit (NEB, England).

#### **3.3.4.2 Transformation**

For the touch-down PCR product, the DpnI enzyme was used to eliminate the template DNA. While the PCR products from Tyr424 variants were treated by kinase, ligase, and DpnI (NEB, England, Table 19) and were incubated at room temperature for 5 min. The designed plasmids were transformed into T7 Express Competent *E. coli* C2566 (NEB, England). After thawing on the ice for 10 min, 1  $\mu$ l (around 20-50 ng) plasmid DNA was pipetted into the 50  $\mu$ l *E. coli*. The cell-plasmid mixture was gently pipetted 4-5 times and incubated on ice for the other 30 min. Subsequently, the mixture was heated at 42 °C without shaking for exactly 90 sec and immediately incubated on ice for 5 min. Then, 950  $\mu$ l pre-warmed SOC medium was added and mixed with the cell-plasmid mixture, which was incubated at 37 °C and 600 rpm for 90 min. After the incubation, 100 – 300  $\mu$ l mixture with cells was placed on agar plates with LB-medium containing the ampicillin. The plates were incubated at 37 °C overnight.

**Table 19:** Reagents for Kinase, Ligase, and DpnI (KLD) Treatment

Reagents	Volume
PCR product	1 $\mu$ l
2X KLD reaction buffer	5 $\mu$ l
10X KLD enzyme mix	1 $\mu$ l
Nuclease-free water	3 $\mu$ l

### 3.3.4.3 Plasmid sequencing and preparation of glycerol culture

The well-isolated colony was picked out from the agar plate and added in 10 ml LB-medium with 50  $\mu$ g/ml kanamycin. Then, the cells were incubated at 37 °C with 120 rpm for 12-14 h. After the incubation, the plasmid DNA was extracted by using the Monarch® Plasmid DNA Miniprep Kit (NEB, England). According to the protocol from the company, 3 ml cultures were harvested through centrifugation at 13,000 rpm (Heraeus™ Pico™ 17) for 1 min. According to the manufacturers' instructions, 30  $\mu$ l of DNA was eluted in nuclease-free water. The concentration of extracted DNA was determined through the UV spectrophotometer with Take3 micro-volume plate (Epoch, BioTek Instruments, Inc., USA) by using the pre-programmed dsDNA protocol. Finally, the mutated enzymes were controlled by DNA sequencing (GATC, Germany) and analyzed through Genome Compiler and SnapGene Viewer (Amirav-Drory et al. 2015). Moreover, the glycerol stocks were prepared by a mixture of 850  $\mu$ l preculture and 150  $\mu$ l sterile glycerol in a sterile cryovial. Before being frozen at -80 °C, the cryovial was gently inverted 5-6 times.

### 3.3.4.4 Expression of enzymes

In 10 ml LB-medium with 50  $\mu$ g/ml kanamycin, 10  $\mu$ l glycerol culture of wild type or designed mutants were incubated at 37 °C for 12-14 h by shaking at 120 rpm. Main cultures were incubated with starting OD<sub>600</sub> of 0.05 in 1 l shaking flasks with 100 ml

---

LB-medium (50  $\mu\text{g/ml}$  kanamycin) at 37 °C and 120 rpm. Once  $\text{OD}_{600}$  of the main cultures reached 0.3-0.4, IPTG was added to a final concentration with 0.1 mM. Then, the shaking flasks were cooled on the ice for 10 min. The expression temperature and duration were set at 25 °C and 22 h.

By using TB-medium for the main culture, the parameters were identical with the above-mentioned LB-medium incubation. Without stopping the incubation by adding IPTG, auto inducing medium (AIM) was also used for enzyme expression. Starting with the 0.05  $\text{OD}_{600}$ , the main cultures were incubated in 100 ml AIM with 50  $\mu\text{g/ml}$  kanamycin at 37 °C. After shaking at 120 rpm for 3 h, the temperature was decreased to 25 °C. The main cultures were then incubated for another 21 h. The other simpler expression protocol with AIM was also carried out. In this case, the main cultures were incubated at 30 °C for 24 h. All pre-cultures and main cultures were conducted in triplicate. After protein expression, the cells were collected by concentration at 4700 rpm, 4 °C for 10 min. The harvested cells were washed twice with 50 mM Tris-HCl buffer.

### **3.3.4.5 Cells disruption**

#### **Cells disruption by sonication**

The cells, which were harvested from 100 ml main culture, were completely resuspended in 5 ml of the binding buffer by pipetting and sonicated on ice with 40% amplitude (the 20s/30s pulsation for 5 min) (Bandelin, Germany). Then the debris and unbroken cells were removed by centrifugation at 15,200 g for 20 min at 4 °C. Finally, the supernatants were collected in a new tube as crude extracts for subsequent purification and characterization experiments.

#### **Cells disruption by BugBuster® Protein extraction reagent**

Per gram pellet was resuspended in 5 ml of BugBuster mix reagent, which contains 500  $\mu\text{L}$  10X BugBuster® Protein extraction reagent, 5  $\mu\text{l}$  of Benzonase® nuclease, 125  $\mu\text{l}$

---

of lysozyme solution, and 4.5 ml of Tris-HCl buffer (pH 8.0). Then, the resuspended pellet was incubated at room temperature by rotation for 20 min. Finally, the supernatants were collected as a crude extract by centrifugation at 15,200 g and 4 °C for 20 min.

### **3.3.4.6 Enzyme purification**

#### **Enzyme purification by heat treatment**

Considering the high thermal stability of TchPAM, the crude extracts were heated at 55 °C, 60 °C, 65 °C, or 70 °C for 30 min to remove the undesired proteins. After the heat treatment, the enzyme solution was clarified by centrifugation for 20 min at 15,200 g and 4 °C. Subsequently, the supernatants were collected for determination of the protein concentration, purification, and retain activity.

#### **Enzyme purification by Fast Liquid Protein Chromatography (Äkta-Start system)**

After cell disruption by sonication, around 5 ml of crude extract in binding buffer was loaded in equilibrated HisTrap™ High-Performance column (1 ml, GE Healthcare, USA) by Äkta-Start chromatography system. At a constant flow rate of 1 ml/min, the unbound proteins were flushed from the column with washing buffer until no signal was detected at 280 nm. Then the desired enzyme was eluted from the column with elution buffer. Fractions containing the TchPAM were automatically collected in 15 ml reaction tubes, which was confirmed by the peak of the UV absorption. The protein concentration, purification, and activity of the collected TchPAM were determined in subsequent measurements. In the final step, the Tris-HCL buffer with 500 mM imidazole was applied to remove the rest of the proteins from the column.

#### **Enzyme purification by rapid small-scale purification (His SpinTrap™ columns)**

In the first step, the storage solution was removed from the prepacked His SpinTrap™ columns (GE Healthcare, USA) by centrifugation at 100 g for 30 s. Then, the columns were equilibrated with 600 µl of binding buffer, which contained the same

---

concentration of imidazole as the crude extract in the binding buffer solution. Subsequently, 600  $\mu$ l of crude extract solution, which was prepared by sonication as described in 3.3.4.5, was pipetted in each column and incubated at 4 °C by rotation for 5 min. The unbound proteins were removed from columns with an equal volume of washing buffer. In order to improve the yield of the purified enzyme, the other 600 crude extract solution was added in the columns, as well as the incubation and washing steps were repeated as previously described. Finally, the purified enzymes were eluted by a 200  $\mu$ l elution buffer. To collect all of the desired enzymes, the elution step was also repeated one more time. The purified enzymes were desalted and concentrated by a protein concentrator (10 kDa cut-off value) and diluted with Tris-HCl buffer (pH 8.0) to a certain concentration (0.5 mg/ml). After short storage on ice, the purified directly determined the concentration and activity.

#### **3.3.4.7 Determination of the enzyme concentration**

To determine the concentration of enzyme, 2  $\mu$ l of purified enzyme solutions were pipetted in microspot locations of the Take3 Micro-Volume plate (BioTek, USA). The UV absorption of the purified enzyme was detected at 280 nm by spectrophotometer (Epoch, BioTek, USA). By using the Gen5-Data-Analysis Software with protocol about BSA, the protein concentration was calculated with the formula of absorption.

#### **3.3.4.8 SDS-PAGE**

To control the expression and purification of enzyme, 20  $\mu$ l of lysate samples (0.5 mg/ml) or resuspended pellet debris were mixed with 5  $\mu$ l of SDS sample buffer and heated at 95 °C for 10 min. The SDS-PAGE gel, containing a 12.5 % (w/v) separation gel and a 5 % (w/v) stacking gel, was placed in the electrophoresis cell (Bio-RAD, USA) with 1 x TGS buffer. After loading the prepared samples and prestained protein marker (NEB, England) in the wells, the gel was run at 200 V around 1 h. Then, the gel was carefully washed with dH<sub>2</sub>O before dyeing the gel with Coomassie blue staining solution under gentle shaking over 30 min. In the final destaining step, the gel was rinsed in

---

water with shanking until blue bands were visible.

#### **3.3.4.9 Lyophilization of enzymes**

At first, the purified enzyme was frozen in a 1.5 ml reaction vessel with liquid nitrogen. After cutting off the lid, the reaction vessel was covered by parafilm. Subsequently, the reaction vessel with purified enzyme was freeze-dried at -80 °C for 72 h and was stored at the same temperature until use. To resolve the lyophilized enzyme, the Tris-HCl buffer (pH 8.0) was used in a certain volume.

#### **3.3.5 Activity assay for Ammonia Elimination Reaction**

Activity assay for ammonia elimination reaction was carried out in triplicate in 96-well UV-Microplates (Greiner Bio-One, Austria) for 30 min at 40 °C and 600 rpm. 1 mM stock solution of aromatic amino acid, such as (*S*)- $\alpha$ -phenylalanine, (*R*)- $\beta$ -phenylalanine, (*S*)- $\alpha$ -tyrosine or (*R*)- $\beta$ -tyrosine, was prepared as substrate solution in Tris-HCl buffer (50 mM, pH 8.8). Meanwhile, this Tris-HCl buffer was used as blank control. 100  $\mu$ l of substrate solution was pipetted in each well and preheated at 40 °C before 10  $\mu$ l of designed enzyme or wild type enzyme with a protein concentration of 0.5 mg/ml was added. Using the Epoch<sup>TM</sup> Spectrophotometer system (BioTek, USA), the ammonia elimination reaction was monitored by measurement of the product absorbance: *trans*-cinnamic acid at 300 nm and *trans*-p-hydroxycinnamic acid at 350 nm. The product absorbance was measured every 3 min. The first measured time point was 0 min.

#### **3.3.6 Kinetics Measurements**

To measure the kinetic parameters, the substrate solution with a varying concentration from 0.01 to 10 mM were pipetted in 96-well UV-Microplates (Greiner Bio-One, Austria) by using a liquid handling station (BRAND, Germany). 2 mM (*S*)- $\alpha$ -tyrosine solution was used as the maximal concentration because of the poor solubility of (*S*)- $\alpha$ -tyrosine in Tris-HCl buffer at 40 °C. The kinetics measurements were performed using protocol from activity assay for ammonia elimination reaction. Kinetic parameters ( $K_M$

and  $k_{\text{cat}}$  values) were determined with Michaelis-Menten curves using the following equation, which were formed by nonlinear fitting on OriginLab (Edwards 2002).

$$v = V_{\text{max}} \frac{[S]}{K_M + [S]}$$

### 3.3.7 HPLC Analysis for the ammonia addition reaction

To 4 M ammonium sulfate solution (pH 10), 3 mM *trans-p*-hydroxycinnamic acid or *trans*-cinnamic acid were added as substrate. Then, 500  $\mu\text{l}$  substrate solution was pipetted in three parallel 150  $\mu\text{l}$  micro reaction tubes, which were preheated at 40  $^{\circ}\text{C}$ . The 50  $\mu\text{l}$  enzyme samples were pipetted and shortly mixed by pipette. After different reaction times (0, 1, 2, 3, 5, 7, and 24 h, some samples were also detected after 72 h), 50  $\mu\text{l}$  samples were taken out from the reaction mixture, directly pipetted into a new micro reaction tubes, which was preheated at 99  $^{\circ}\text{C}$  and heated for 5 min. The samples were directly detected through HPLC (Agilent 1100 Series, USA) with the derivatization of the possible formed aromatic amino acids. According to the protocol by Brucher *et al.*, the sodium borate buffer, OPA/IBLC solution, and the acetic acid solution were always freshly prepared. Initial, 2  $\mu\text{l}$  sodium borate buffer and 3  $\mu\text{l}$  OPA/IBLC solution were mixed in the automated precolumn for 2 times with maximum speed. Then, 0.5  $\mu\text{l}$  samples reacted with the mixture solution for 1 min. The reaction was terminated by adding 3  $\mu\text{l}$  acetic acid solution. Finally, these 7.5  $\mu\text{l}$  mixture solutions were injected into the C18 HPLC column (Hyperclone, 5  $\mu\text{m}$ , 150 x 4.6 mm, Phenomenex Inc, Germany) and detected at 338 nm. For detection of the phenylalanine, the compounds were eluted with 32 % sodium phosphate buffer (40 mM, pH 6.5) and 68 % methanol. While the analysis of the samples with tyrosine was operated with a ratio of the sodium phosphate buffer and methanol in 21:79. The flow rate was set at 1 min/ml. The retention times of the formed aromatic amino acids were (*S*)- $\alpha$ -phenylalanine 36.0 min, (*R*)- $\alpha$ -phenylalanine 32 min, (*S*)- $\beta$ -phenylalanine 57.8 min,

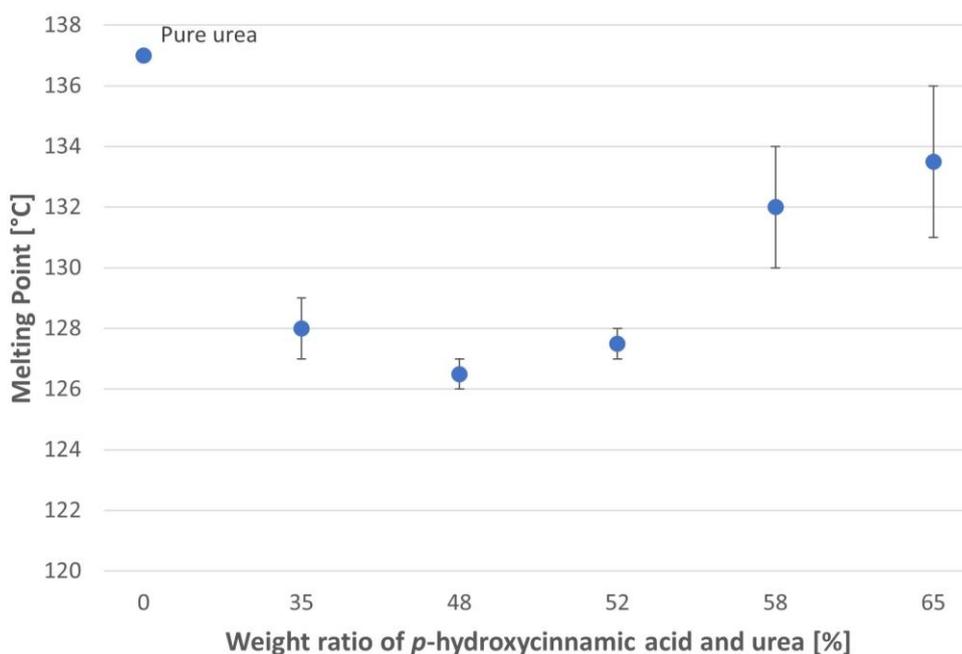
---

(*R*)- $\beta$ -phenylalanine 29.4 min, (*S*)- $\alpha$ -tyrosine 71.0 min, (*R*)- $\alpha$ -tyrosine 45.7 min, (*S*)- $\beta$ -tyrosine 57.6 min, (*R*)- $\beta$ -tyrosine 42.7 min.

## 4 Results

### 4.1 Chemical synthesis

The melting point of the *trans-p*-hydroxycinnamic acid was determined as 214 °C, while the cinnamic acid already melted at 133 °C. Considering the reaction temperature (195 °C) from the literature, which was lower than the melting point of *trans-p*-hydroxycinnamic acid, the eutectic reaction between *trans-p*-hydroxycinnamic acid and urea was carried out to confirm whether the mixture melted at the reacted temperature. Different proportions of *trans-p*-hydroxycinnamic acid and urea were selected based on some literature, including 1:1.5, 1:2, 1:2.5 (Chang and Jeung 1991), 1:3, and 1:5 (Vuano et al. 2000). The results revealed that the melting points of the mixture were lower than 138 °C, while the most suitable *trans-p*-hydroxycinnamic acid and urea proportion was 1:3 (Figure 14).

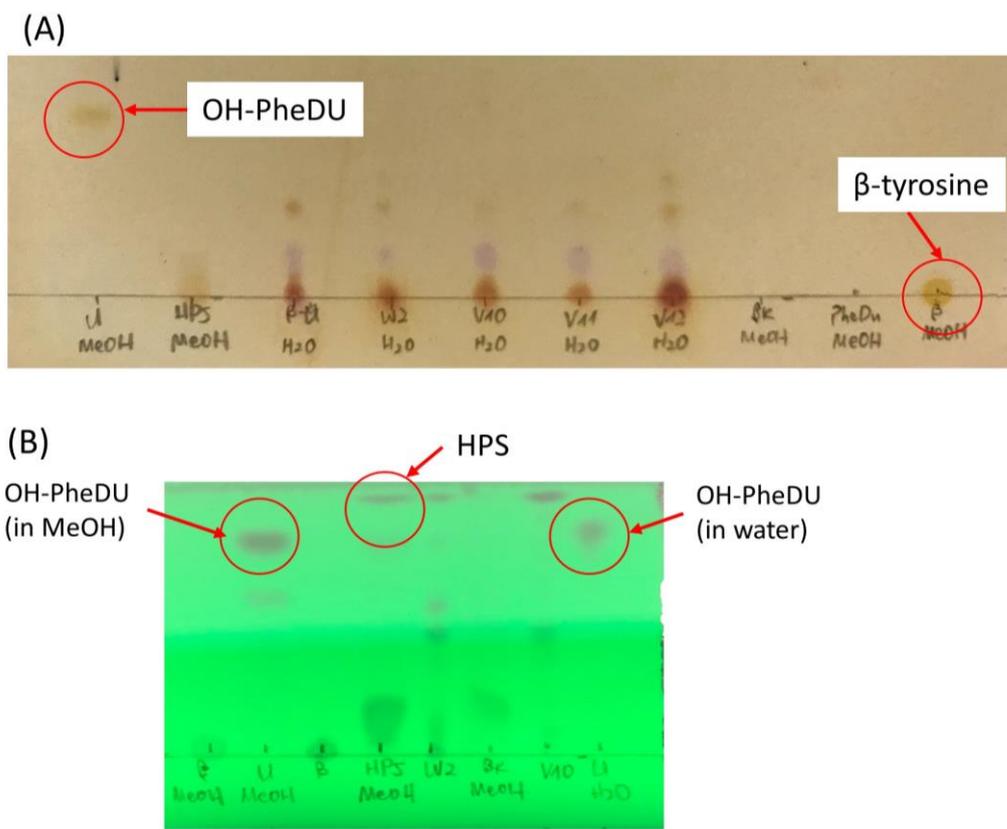


**Figure 14:** Eutectic point test for the mixture of *p*-hydroxycinnamic acid and urea in different proportions.

To prepare the substrate for the modified hydantoinase process, the OH-PheDU and *N*-

---

carbamoyl- $\beta$ -tyrosine were produced through different synthesis methods. The standard of OH-PheDU was prepared by using the *N*-carbamoyl- $\beta$ -tyrosine as starting material, according to the method, described in section 3.3.1.2. By elution with the solvent mixture of methanol and phosphate solution in 30:70, the retention times of OH-PheDU, *N*-carbamoyl- $\beta$ -tyrosine, and  $\beta$ -tyrosine were 2.55 min, 2.18 min, and 2.09 min, respectively. The *N*-carbamoyl- $\beta$ -tyrosine (2.48 min) and  $\beta$ -tyrosine (2.1 min) were detected by HPLC with a ratio of methanol/phosphate solution in 20:80. The identification of the OH-PheDU was proved by  $^1\text{H-NMR}$  spectrum (300 MHz) in  $\text{DMSO-d}_6$ :  $\delta$  2.56 (1H, CH<sub>2</sub>), 2.74 (1H, CH<sub>2</sub>), 4.54 (1H, CHPh), 6.76-7.10 (4H, Ph), 7.86 (1H, CHNHCO), 9.43 (1H, OH), 10.10 (1H, CONHCO) and the observed  $m/z$  value of OH-PheDU by mass spectrometry was 206.2 with a relative abundance of 100 %. Furthermore, the OH-PheDU, *N*-carbamoyl- $\beta$ -tyrosine, and  $\beta$ -tyrosine were also analyzed by different TLC methods (Figure 15).



**Figure 15:** TLC measurement of OH-PheDU, *N*-carbamoyl- $\beta$ -tyrosine (HPS), and  $\beta$ -tyrosine. These products were solved in water or methanol. (A) TLC with ninhydrin reagent. (B) TLC under UV light at 254 nm.

Initially, the sealed pressure tube was used to produce the OH-PheDU from *trans-p*-hydroxycinnamic acid. According to the detected results by HPLC, the production yield of OH-PheDU was quite low by using the *trans-p*-hydroxycinnamic acid as starting material. In this case, the NMR results indicated that the formed OH-PheDU might be disrupted at high temperatures. After isolation of the synthesized products by flash chromatography, one of the primary compounds was styrene, which might indicate the cleavage of the bond of *trans-p*-hydroxycinnamic acid or OH-PheDU during the heating step. The synthesis was carried out at a lower temperature. Moreover, to improve the extraction yield, the dissolving analysis of both OH-PheDU and *N*-carbamoyl- $\beta$ -tyrosine was conducted. As shown in Table 20, methanol was the best solvent to dissolve the educt *N*-carbamoyl- $\beta$ -tyrosine and product OH-PheDU. Therefore, instead

of hot water, methanol was used for the extraction process. Moreover, considering the advantages of microwave synthesis during the solid-solid-reaction: reduction the reaction temperature and time, improvement of the production yield, the synthesis of OH-PheDU was also performed in the microwave reactor. After the optimization of the reaction parameters, microwave synthesis was the best method, giving the highest product yield in 3 %. Considering the notable low conversion of the OH-PheDU, which was used as the substrate for hydantoinase, the chemoenzymatic synthesis of  $\beta$ -tyrosine from *trans-p*-hydroxycinnamic acid by modified hydantoinase process was terminated.

**Table 20:** Solubility of OH-PheDU and *N*-carbamoyl- $\beta$ -tyrosine in different solvents

Solvents	OH-PheDU	<i>N</i> -carbamoyl- $\beta$ -tyrosine
Cooled water	No	Yes
Chloroform	No	No
Ethyl Acetate	Yes	No
Methanol	Yes	Yes

**Table 21:** The production yield of the OH-PheDU under different conditions.

Method	Temperature [°C]	time	$r_m$ (CA : urea)	Yield [%]
Pressure tube	195	5 h	1:2.5	0.5
	195	2 h	1:2.5	0.5
	170	5 h	1:2.5	n.d.
Microwave	250	5 min	1:2.5	0.3
	195	20 min	1:2.5	0.8
	195	50 min	1:2.5	3

CA: cinnamic acid.

## 4.2 Determination of the suitable enzyme template

In order to determine the suitable PAL, 66 available enzymes in the class EC.4.3.1.24 were collected their sequence information from SwissProt (Bairoch et al. 2004).

---

Meanwhile, the homologues, which belong to the MIO-dependent enzyme family with an identity over 30 %, were also chosen to reconstruct a phylogenetic tree for exploration of the evolutionary history of the MIO-dependent enzymes. Apart from PAL, the genetic data of the other 203 protein sequences were available on the SwissProt, including the TAL, HAL, PTAL, PAM, TAM (Bairoch et al. 2004).

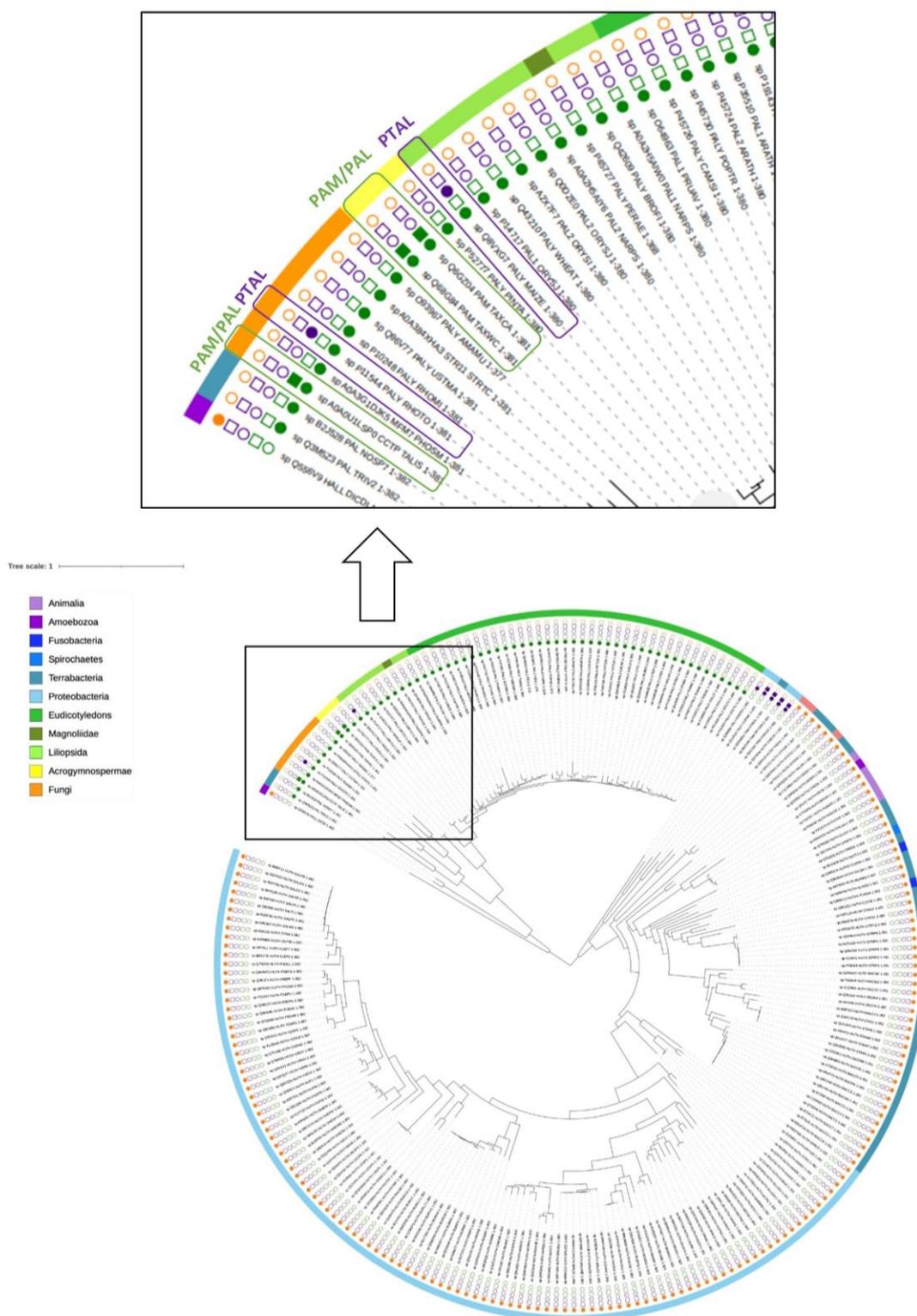
As shown in Figure 16, the corresponding enzymes would be sorted in two well-supported clusters with bootstrap support of 96 %: the one included the eukaryotic HAL from *Dictyostelium discoideum* and the PALs, while, the other cluster contained the rest of the HALs and all TAL/TAM sequences. In conclusion, the phylogenetic tree suggested that the MIO-dependent enzymes shared a common ancestry.

In the first cluster, the eukaryotic PALs consisted of the well-supported fungal and plant branches (100 %). The bacterial PALs from *Anabaena variabilis* and *Nostoc punctiforme* were the closest relatives of these two eukaryotic branches, indicating that the fungi and plants might acquire the genes from a bacterial ancestor. The *Anabaena variabilis* and *Nostoc punctiforme* represent the bacteria with diazotrophic symbiotic associations with plants and fungi. It could be recognized as the additional evidence of the genetic transfer of PALs at an early age by symbiosis between bacterial ancestors, fungi, and plants (Meeks et al. 2001). As shown in the phylogenetic tree, the plant PALs consisted of three major sub-branches: the *Acrogymnospermae*, the *Liliopsida*, and the *Eudicotyledons*. Among them, the PALs from the eudicotyledons represented the largest source of plant PALs. In contrast to the eudicotyledons PALs with high similarity, the clade of the PALs from *Acrogymnospermae* was located between the fungi and angiosperms and diversified into two well-supported branches (100 %). One branch included the PAL from *Pinus taeda*, while the other branch contained the PALs from *Taxus canadensis* and *Taxus wallichiana var. chinensis*. Except for the ability to convert the phenylalanine to the cinnamic acid, the PALs from the latter branch also exhibited the PAM activity, which was employed for isomerization reaction. In addition, the eukaryotic HAL in the first cluster formed the deepest but well-supported branching

---

lineage (100%). This outsider eukaryotic HAL supported the hypothesis of the single event of the horizontal gene transfer (HGT) between prokaryotic ancestors during the early evolution.

The second cluster split into two main well-supported branches (96 %): the one contained all bacterial TALs/TAMs while the other one included the HALs except the HAL from *D. discoideum*. The deep branching lineage in this cluster, which was formed by the bacterial TALs/TAMs, indicates that they were closer to the common ancestor compared to the bacterial HALs. At the branching point from the genus *Thermoplasma*, including *Thermoplasma acidophilum* and *Thermoplasma volcanium*, the clade of the bacterial HALs was developed in a distinct branch with the genus *Streptomyces* (*Streptomyces avermitilis*, *Streptomyces coelicolor*, and *Streptomyces griseus*), albeit with poor statistical support (44 %). This result supports the occurrence of the HGT between the ancient bacteria and archaea. Furthermore, the bacterial HALs separated into two main branches: the one included the genus *Proteobacteria*, while the other one contained the Terrabacteria, Fusobacteria, and Spirochaetes. In contrast with the bacterial HALs, the eukaryotic HALs arose at the position close to the genus *Deinococcus* with poor statistical support.



**Figure 16:** The phylogenetic tree of the MIO-dependent enzymes. The protein sequences were collected from the SwissProt database (Bairoch et al. 2004). The characterized aromatic amino acid ammonia-lyases were displayed in the square, and the characterized aromatic amino acid aminomutases were marked as the circle. The different substrate acceptance was sorted in various colors: histidine in yellow, phenylalanine in green, and tyrosine in purple. The color of the outer ring indicated the source organisms.

---

Among these MIO-dependent enzymes, the PALs with PAM activity as the side reaction present great potential for the generation of the  $\beta$ -tyrosine from the *trans-p*-hydroxycinnamic acid. Three PALs exhibit the ability for isomerization of the  $\alpha$ -phenylalanine. Of these, the genome sequence data of the fungal PAL from *Talaromyces islandicus* was released recently, indicating that the corresponding genes not only exist in plants (Schafhauser et al. 2016). Whereas, the PALs from the genus *Taxus* were the most investigated enzyme in the superfamily, especially, the TchPAM, which could produce enantiopure (*R*)- $\beta$ -phenylalanine and served as the precursor of the famous antitumor drug paclitaxel (Walker et al. 2004). Considering the excellent enantioselectivity of TchPAM, which was reported in earlier studies, TchPAM was chosen as the template for subsequent computational design of the novel enzyme for production of the enantiopure (*R*)- $\beta$ -tyrosine by using the monolignols (Wu et al. 2009) (Zhu et al. 2018).

### **4.3 Selection of suitable amino acid residues for enzyme design**

#### **4.3.1 Multiple sequence alignments**

Apart from the PAL/PAMs, partial PALs in fungi and monocots conferred phenylalanine and tyrosine acceptance, which were named as PTALs. Whereas, the monofunctional TALs were found in bacteria in common. Comparison of the sequences of the PTALs and bacterial TALs might reveal the origin of the PTALs in fungi and plants, as well as the relationship between PAL and TAL, which was important for the selection strategy of the residues in the active site of TchPAM for enzyme design. The multiple sequence alignment was performed with the PTALs in fungi and plants, the PALs in monocots, as well as all enzymes with the TAL activity in bacteria (RsTAL from *Rhodobacter sphaeroides*, SgTAM from *Streptomyces globisporus*, CmTAL from *Cupriavidus metallidurans*, CcTAM from *Chondromyces crocatus*, MfTAM from *Myxococcus fulvus*, BfPAL from *Bromheadia finlaysoniana*, NpPAL from *Narcissus*

---

*pseudonarcissus*, OsPAL from *Oryza sativa subsp. Japonica*, and TaPAL from *Triticum aestivum*) (Figure 17).

The residues, which were located surrounding the active site of the MIO-dependent enzymes, were compared in the multiple sequence alignment. The fungal RtPTAL from *Rhodosporidium toruloides* and plant ZmPTAL from *Zea mays* had the higher sequence similarity to the PALs from the monocots, compared to the bacterial TALs. For instance, the residue Tyr440 in RtPTAL, which was identical with the residue in PALs/PTALs from monocots at the same position, was mutated in cysteine and serine in the bacterial TALs/TAMs. The other example was the residue Lys443 in RtPTAL, which seemed like a highly conserved position in PALs/PTAL from monocots. Unlike the PALs in monocots, this residue was replaced by alanine, glutamic acid, and methionine in bacterial TALs/TAMs. The multiple sequence alignment suggests that the bifunctional PTALs might originate in the monocots through the mutations of the residues in the active site. Subsequently, the corresponding genes were transferred to fungi by HGT.

In ZmPTAL, apart from the residues Tyr440, Lys443, Ile447, and Asp471 (Tyr424, Lys427, Ile431, and Glu455 in TchPAM), it was worth mentioning that the residues His123 and Leu124 (Cys107 and Leu108 in TchPAM) had a strong influence on substrate selectivity. In the TALs/TAMs from bacteria and PTALs, a hydrophilic residue, such as serine and histidine, was preferred at position 123. Whereas, the PALs from monocots harbor phenylalanine at the same position. One special example is the OsPAL, in which both PAL and TAM activity were observed in current research (Walter et al. 2016). Moreover, the mutation of the residue His89 (His123 in ZmPTAL) in phenylalanine led to substrate selectivity switch of the RsTAL in the previous mutagenesis study (Watts et al. 2006).

53 IREARRVGLTGGGLLANRLISGENVRTLQANVHVAASGV.....GVLWDWTTARAMVLARLVSIAQGSAGSEGTIARLIDLLNSELAPAVPSRGTGAS 160  
 56 AEQNIPIGVSTGEMIMYQVDSKSEVELQTNVRSASAGV.....GFLFAEDEARAIVAAERLNLAKGHSVVRPIILERLAQYLNKGIITPAIPEIGVSLGAS 163  
 50 AAANVPVIGVSTGELVHWVDIEGRALQENLRSHCAGV.....GFLSRDEVRAMVAVANALAKGYSVVRPAVIEQLLKYLEAGITPAVPOVGSLOGS 147  
 44 QEAQHPVIGVNTGELVPMVPROHKRELOENIRSHCAGG.....GEPFADVVRAIMLARLNCMLKMGVSGASVETVKLAEFIRRGHPVPIPOOGLGAS 141  
 34 SNEKTPMVGVTGEGELINVLIPPGKSDLQKALRSHCAGG.....GEPFDEVVRAIMTVRINCLMKMGVSGISPEALQQLATMLNRRGHPVPIPMOOSLOGS 131  
 105 L...SMSVVGVTGEGGADTIRI...EDALISLQKALELQGLVLSSEDSFRLGRLENLPLRVSALTRGHSAVLLVLEALTINRHGJLPIVPLRGTIIRAS 212  
 89 IAHGGDIIGVTTGGTSHRRR...KDGPAQVELRUNAGI.....FGTGS...DGHLPSEVTRAAMLVRIINLLQGYSGIRFEILEAITKLLNTGVSPPCLPRTITAS 190  
 91 VDKGDSVGVTTGGTSHRRR...KQGGALQEKIKFVNAGI.....FGSG...NSNTLPSAATRAAMLVRIINLLQGYSGIRFEILEIKAIATLKNKIPCPLRGTITAS 191  
 95 MAKGDSVGVTTGGTSHRRR...KQGGALQEKIKFVNAGI.....FGSGKESGNTLPASTRAAMLVRIINLLQGYSGIRFEILEAITALLNANVTPCLPRTITAS 197  
 88 IAHGGDIIGVTTGGTSHRRR...KDGPAQVELRUNAGI.....FGTGS...DGHLPSEVTRAAMLVRIINLLQGYSGIRFEILEAITKLLNTGVSPPCLPRTITAS 189  
 86 MNNGTDSVGVTTGGTSHRRR...KEGGALQEKIKFVNAGI.....FGTGT...DGHVLPAAATRAAMLVRVNTLLQGYSGIRFEILEAITKLLNANVTPCLPRTITAS 187  
 151 RDTPLAHMVLCLQGGDF...LDRDGT...LDGAEGLRRGLQP...LDLSHRDALALVIGTSAMTGIALVNAHACRHLGNWAVALTALLAECLRGRTEAWAALS...DLRPH 254  
 154 RDLAPLSHVASTLIGESY...VLRDGRP...VTAQVLAERGIEP...LELRFKEGLALVIGTSMGTGLGSLVVGRALEGAQAAEIVTALLIEAVRGTSPFLAEGHDIARPH 257  
 148 RDLAPLSHVASTLIGESY...LTDGGT...APTAEVLRERGITP...LALAYKEGLALVIGTSAMTGVSCLLLETLRAOQQAAEIAALALEGLSADAFMAHGHDIARPH 252  
 142 RDLAPLSHVASTLIGESY...VSKKQGV...RKTGDVLRREGLP...LELFGKGGTLVIGTSAMTGAACVALGRAYHLFRLLALLATADFOVCCLOGSTGPFEEGRH...LPKNH 244  
 132 RDLAPLSHVASTLIGESY...VRKNGVT...RPTMEVFQEEGLTP...LKLFGKEGLALVIGTSAMTGAASALYARHLLRSLLSASADIQVAMHASTPFFSHIGN...ALKKH 234  
 213 RDLAPLSYIAAALISGHPKSVHVVHHEGKIKLYAREAMALNLEP...VVLGPKKEGLALVIGTAVSASMAILALHDAMHLSLSLSLSTAMTYEAMVHAGSHEFLHDVIREP 322  
 191 RDLAPLSYIAGLITGRPNAQA...VTVDGRK...VDAAEAFKIAGIEGGFKLPNKEGLALVIGTSVGSALAAATVMYDANVILAVSELSAVFCEVMNGKPEYTDLTH...KLKHH 298  
 192 RDLAPLSYIAGLITGRPNSKA...RTPNGST...VDATTAFRLAGISGFFLQPKKEGLALVIGTAVSGVASIVLFTNILLVMAELLSALFCEVMNGKPEYTDLTH...KLKHH 299  
 198 RDLAPLSYIAGMLTGRPNSKA...VGPDDGT...IDASEAFKLAGIPNGFFELQPKKEGLALVIGTAVSGSLASTVLFDFANILAVLSEISAVFCEVMNGKPEYTDLTH...KLKHH 305  
 191 RDLAPLSYIAGLITGRPNAQA...ISPDGRK...VDAAEAFKLAGIEGGFKLPNKEGLALVIGTAVSGSALAAATVMYDFANILAVLSELSAVFCEVMNGKPEYTDLTH...KLKHH 298  
 188 RDLAPLSYIAGLITGRPNSMA...ISPDGRK...VNAAEAFKLAGIEGGFKLPNKEGLALVIGTAVSGSALASMLVDFEANVILAVLSELSAVFCEVMNGKPEYTDLTH...KLKHH 295  
 255 PQQDAARLRARVDGSRVVRHVIARRLDA...GDIQTEPEAGDAYSLECAPVLGAGFTLAWHDRVLTIEL...NAVTD...PVPFPGDGVSPALHGGNFMGQHVALTSD 390  
 258 EQQIDTAANMRMRGSGLTVEHADLRRELQKKEAGKDVORSEIYLKAVSLRIPQVVGVADRTLYHARHKRIEL...NSAND...PVLFFE...GKEIFHGANFHGQPIAFAMD 390  
 253 PQQIRSAANMRALLADSRLSGHGELSAEMKTRAGEAKNTG...TGVI...PKAYTLR...CIPQVLGAVRDTLDHCATVVEREL...NSSND...PVLFFE...DGEFLHGGNFMGQVAFAMD 380  
 246 SQQVIVAREIRKLLAGSGLTSDHQDLKEMVARSQGVNDVVUVDYVYLDATLAVPQILGVLDTLDFARKLIEEEL...NSTND...PVLFDV...PEQTFHGANFHGQVYVAMACD 364  
 235 PQQVIVARIMRDLTQGLMRDHDQIMRAISEITSHSNDVEEIEIYLNANVSL...CMPOVLGVVLETLQMCQRFIEEEA...NSVMD...PVLIDT...PAEITHGANFHGQVYVAMACD 344  
 223 PQLVAVAGNIRKLLGSRFVHHEEVKVKDQ...EGLLR...RDRVPL...ITSPQWLGPLVSDLIHAHAVI...ILEAGSITD...PVLIDVE...NKISHHGGNEFEQAQAVANTME 423  
 300 PQQIEAAAIMEHILEGSSYM...QMAKKLHELDP...LQPKDRYAL...ITSPQWLGQPIEVIRAATKSIEREV...NSVMD...PVLIDVH...RKKALHGGNFMGQPIGVSM 397  
 308 PQQIEAAAIMEHILEGSSYM...QMAKKLHELDP...LQPKDRYAL...ITSPQWLGQPIEVIRAATKSIEREV...NSVMD...PVLIDVH...RKKALHGGNFMGQPIGVSM 398  
 299 PQQIEAAAIMEHILEGSSYM...SHAKKVNEMDP...LQPKDRYAL...ITSPQWLGQPIEVIRAATKSIEREV...NSVMD...PVLIDVH...RKKALHGGNFMGQPIGVSM 397  
 296 PQQIEAAAIMEHILEGSSYM...MLAKKLGELDP...LQPKDRYAL...ITSPQWLGQPIEVIRAATKSIEREV...NSVMD...PVLIDVH...RKKALHGGNFMGQPIGVSM 394  
 361 ALATAVTVLGLAERQIARLTDERLNRGLPFLHRG...PAGLNS...GFMGAGV...PATALLAEMRAT...GPASIHISISTNAAGV...DVVSLGTIARLCREKIDRWAEILAILALCLAGA 470  
 367 FVTIALTQLGVLAERQINRVLNRHLSYGLPEFLVSG...DPGLH...GFMGAGV...PATALLAEMRAT...GPASIHISISTNAAGV...DVVSLGTIARLCREKIDRWAEILAILALCLAGA 470  
 361 FLAIAATQLGVVSERRLNRLLSPHLNNLPAFLAAA...NEGLH...GFMGAGV...PATALLAEMRAT...GPASIHISISTNAAGV...DVVSLGTIARLCREKIDRWAEILAILALCLAGA 471  
 355 YLNIAVTEIGVLAERQLNRLLVDPNPKLPPFLASA...HSGLL...GFMGAGV...PATALLAEMRAT...GPASIHISISTNAAGV...DVVSLGTIARLCREKIDRWAEILAILALCLAGA 465  
 345 YLSIAVEMGVLAERQLNRLLDPHINKPLPFLAHA...KTGLF...GFMGAGV...PATALLAEMRAT...GPASIHISISTNAAGV...DVVSLGTIARLCREKIDRWAEILAILALCLAGA 465  
 424 KTRFLAQIGKLNFTQITEMLAGMNRGLPSCLAEE...DPSLS...YHCKGLD...IAAAYITSELGHANPVTTHVQPA...EMANAVNSLALARRTIESNDVLSLLATHLYCVLQA 534  
 398 NARLAIANIGKLMFAQFSELVNEFYNNGLSNLAGSRNPSLDYGFKGTIEAMASYSELOALANPVTNHVQSA...DHNC...DVNSLGLVSRKTAEAIDLKLMSTYIVALCOA 500  
 399 NTRLAIANIGKLMFAQFSELVNDFYNNGLSPNLSGGRNPSLDYGFKGTIEAMASYSELOALANPVTNHVQSA...DHNC...DVNSLGLVSRKTAEAIDLKLMSTYIVALCOA 510  
 405 NTRLAIANIGKLMFAQISELVNDFYNNGLPSNLSGGRNPSLDYGFKGTIEAMASYSELOALANPVTNHVQSA...DHNC...DVNSLGLVSRKTAEAIDLKLMSTYIVALCOA 516  
 398 NARLAIANIGKLMFAQFSELVNEFYNNGLSNLAGSRNPSLDYGFKGTIEAMASYSELOALANPVTNHVQSA...DHNC...DVNSLGLVSRKTAEAIDLKLMSTYIVALCOA 509  
 395 NTRLAIANIGKLMFAQFSELVNDFYNNGLSPNLSGGRNPSLDYGFKGTIEAMASYSELOALANPVTNHVQSA...DHNC...DVNSLGLVSRKTAEAIDLKLMSTYIVALCOA 506

Figure 17: Multiple sequence alignment of the PTALs from plants and fungi, the monofunctional PALs in monocots, and the bacterial TALs/TAMs. The surrounding residues in

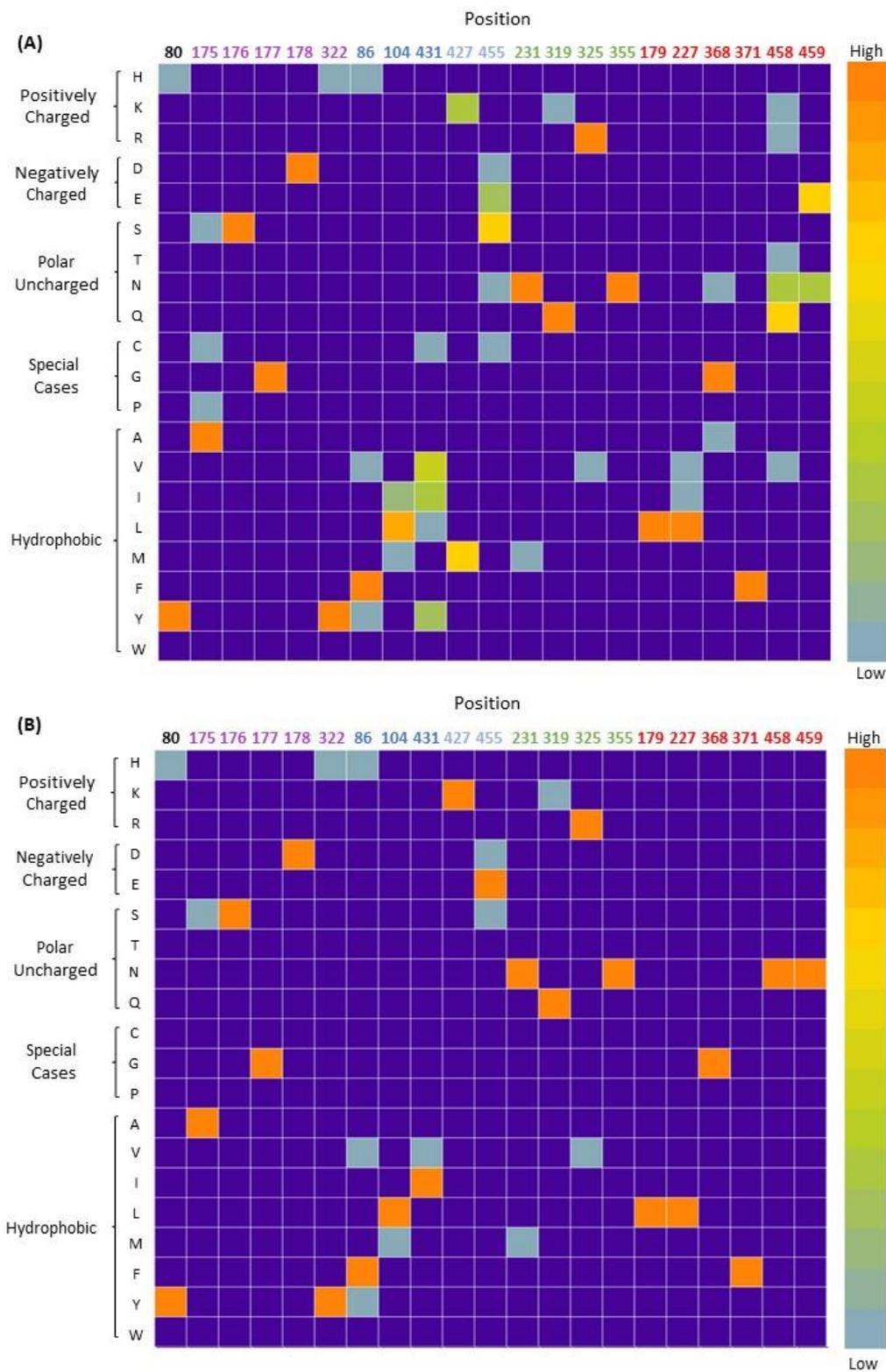
---

the active center, which are placed within 5 Å of the ligand backbone, are colored depending on their characterization of the side chain.

### 4.3.2 Conservation analysis

The highly conserved residues with lower mutation rates play a significant role in enzymes, participating the structure stability, cofactor formation, and substrate catalysis. The strong evidence is the disappearance of the enzyme with the varied amino acids at the position during the evolution in the early ages. Therefore, the determination of highly conserved residues is necessary, before creating the mutant library.

As shown in Figure 18, all of the residues, which located within 5 Å of the ligand backbone, was highly conserved in the enzyme with phenylalanine ammonia-lyase activity, whereas the residues at the position in the aromatic binding pocket (104 and 431) could be mutated in the other amino acids with hydrophobic side chain in HALs. These results might indicate that the residues in the aromatic binding pocket but not the carboxyl binding pocket are responsible for the substrate selectivity. It was consistent with the mutagenesis studies on substrate selectivity in MIO-dependent enzymes. A remarkable example was the acceptance of phenylalanine as substrate in the RsTAL by replacement of one histidine in the aromatic binding pocket with hydrophobic amino acid phenylalanine (Watts et al. 2006). The positions Asn458 and Gln459 allowed the amino acids with the similar in HAL, suggesting that these two residues might influence the activity of enzyme but not strictly required in the active site of TchPAM. Furthermore, the sequence alignment revealed two non-conserved residues Lys427 and Glu455.

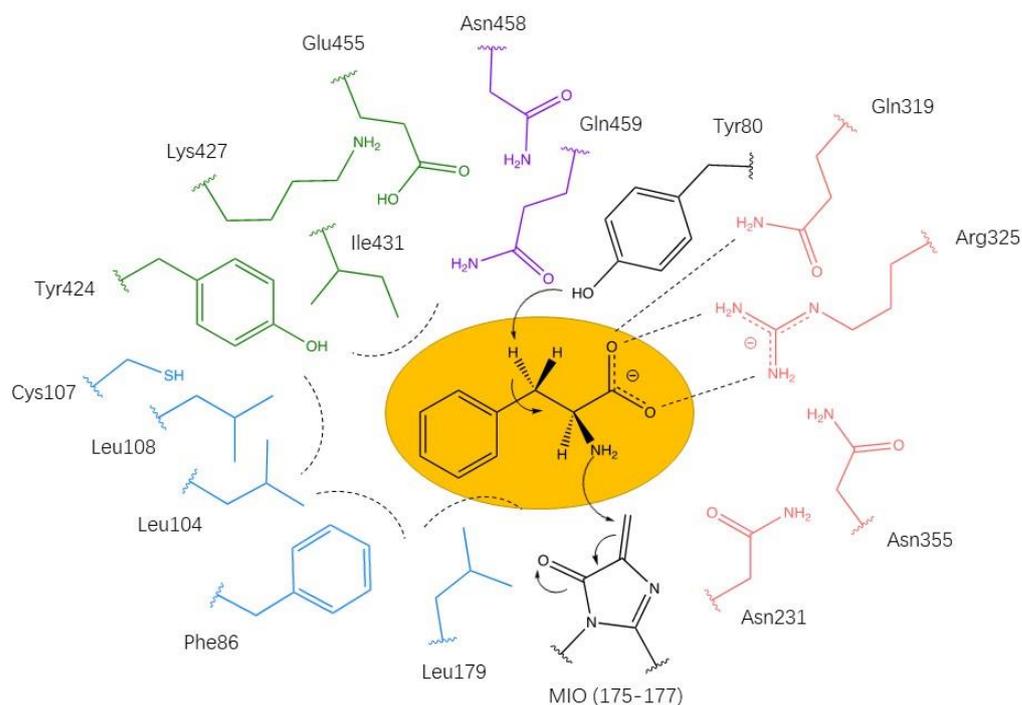


**Figure 18:** A heatmap for conservation analysis of the active site residues. Upon the appearance probability of the amino acids in various homologues, the corresponding squares are colored in red (hot, often appear) to light blue (cold, rarely appear). **(A)** The sequence alignment of homologues contained all the members from the MIO-dependent enzyme superfamily. **(B)** The sequence alignment of homologues consisted of PAM and PAL.

---

### 4.3.3 The selection of the designed residues in the active site of TchPAM

Combining the conservation analysis and the structural investigation on TchPAM, the residues surrounding the substrates can be divided into six main groups (Table 22). The first one contained the residue Tyr80, which was confirmed as a catalytic essential residue in all members of the MIO-dependent enzyme family (Röther et al. 2002)(Wybenga et al. 2014). The residues, which are responsible for MIO moiety formation, were summarized in the second part, including the tripeptide Ala175-Ser176-Gly177, the residue Asp178, and Tyr322. The residue Asp178 with a large side chain provided the mechanical compression to the MIO loop during the cyclization, which has been identified by the earlier structure study on HAL (Baedeker and Schulz 2002). While the latter residue formed a hydrogen bond towards the oxygen on the residue Ser176, promoting the exocyclic dehydration step (Sánchez-Murcia et al. 2016). The additional evidence was the lack of MIO in the mutant Tyr322Ala in TchPAM (Wybenga et al. 2014). The third part consisted of the residues Asn231, Gln319, Arg325, Asn355, and Asn458, which were classified in the carboxyl binding pocket since their locations were in close contact with the carboxylate group of the substrate. From a similar structural point of view, the residues with hydrophobic side chain, which were located at the aromatic side of the substrate, were considered as the residue in the aromatic binding pocket, containing the residue Phe86, Leu104, Leu108, Leu179, and Ile431. The residue Cys107 and Gln459 were also sorted into this aromatic binding pocket in the previous investigation due to their position (Wu, et al. 2012). However, these residues were grouped in the fifth part, together with the residues Lys427 and Glu455. The rest residues surrounding the ligand were Leu227, Gly368, and Phe371, which were highly conserved in homologous of TchPAM.



**Figure 19:** Schematic representation of the interactions between the natural substrate (*S*)- $\alpha$ -phenylalanine and residues in the active center of TchPAM (PDB ID: 4C5R). The color coding corresponds to the position: carboxyl binding pocket (red), aromatic binding pocket (blue), tyrosine towards pocket (green), essential functional residues (black), and all other residues without specified function (purple). The substrate (e.g., (*S*)- $\alpha$ -phenylalanine) is highlighted by a yellow background. The highly conserved residues are not shown in this Figure.

**Table 22:** A summary of dividing groups of the active site residues with their proposed functions.

<b>Group Number</b>	<b>Containing the Residues</b>	<b>Function</b>	<b>Reference</b>
1	Tyr80	Activation	(Röther et al. 2002)(Wybenga et al. 2014)
2	Ala175, Ser176, Gly177, Asp178, Tyr322	MIO moiety Formation	(Wybenga et al. 2014)(Baedeker and Schulz 2002)(Sánchez-Murcia et al. 2016)
3	Asn231, Gln319, Arg325, Asn458	Carboxylate Binding	(Wu, et al. 2012), this thesis
4	Phe86, Leu104, Leu108, Leu179, Ile431	Aromatic group Binding	(Wu, et al. 2012), this thesis
5	Cys107, Tyr424, Lys427, Glu455, Gln459	Influence on the neighboring Residues	This thesis
6	Leu227, Gly368, Phe371	unknown	

---

## 4.4 Computational enzyme design by PyRosetta

### 4.4.1 Preparation of the input files for ligand and MIO moiety

The input files for both ligand and MIO moiety were prepared by OpenBabel, including the params files and a rotamer library (Boyle et al. 2011). As a noncanonical backbone part in TchPAM, the MIO prosthetic group could not be identified and calculated by PyRosetta (Chaudhury et al. 2010). Therefore, the online service Rosie, which was coded based on PyRosetta, could not be used to design the residues in the active site of TchPAM (Moretti et al. 2018). During the initial attempt, the MIO moiety was treated as a normal cofactor, which was fixed on the backbone of the enzyme. In this case, the amino group of the ligand was docked always far from the MIO prosthetic group because of the clash penalty by calculation of PyRosetta. However, this amino group must react with the MIO moiety for catalysis. To solve this problem, two strategies had been formulated to create a suitable model for MIO prosthetic group. The MIO moiety was treated as the cofactor, which could flexibly move and rotate in the active site in a limited coordinate space. Otherwise, a theozyme was generated for *de novo* enzyme design by PyRosetta. This theozyme consisted of the catalytic essential residue Tyr80, and the ligand  $\beta$ -tyrosine, which was bound with the MIO moiety. Gaussian09 and GaussView06 were used to optimize and minimize the geometry of theozyme with quantum mechanic calculation (M. J. Frisch, 2016). In these two cases, the MIO moiety was not placed at the correct position in the active site.

The BCL version 3.2. with the additional license was used to generate the ligand conformer library (Kothiwale et al. 2015) (Allen 2002). Even though the MIO prosthetic group was put into the active site, the interaction between the alkene of the MIO moiety and the amino group of  $\beta$ -tyrosine was not calculated by PyRosetta. Therefore, the constraints file of the ligand  $\beta$ -tyrosine was mainly focused on the interaction of the surrounding residues without MIO moiety.

---

#### 4.4.2 Preparation of the input files for enzyme template

Two X-ray structures of TchPAM bound substrate was available in the PDB database: the one (PDB ID: 3NZ4) contained two monomeric units with cinnamic acid as bound ligand, while the other (PDB ID: 4C5R) provided the complete structure bound the (3S)-3-amino-2,2-difluoro-3-phenylpropanoic acid as a ligand (Berman et al. 2000) (Wybenga et al. 2014) (Feng et al. 2011). The active site of TchPAM consisted of the residues from three different monomeric units (chains). Therefore, the PDB structure of 4C5R was chosen as the template for computational enzyme design (Feng et al. 2011). The downloaded PDB file often included the extraneous molecules, which were not calculated by PyRosetta, such as water molecules. The enzyme structure should be optimized for reduction of the errors about the clashes by atom overlaps and minimization of the global energy, before the utilization as template for subsequent calculations.

### 4.4.3 Enzyme design

To create the mutant library, the computational program were performed several rounds with different limitations of residue variation, which were listed in Table 23.

**Table 23:** The summary of mutation limitation of different residue groups in the various program running. HP: The residues could be mutated in amino acids with a hydrophobic side chain. WL: The residues could be changed without any limitation. \*: In this round, the residue Leu179 was not mutated. +: Only residue Asn458 was selected for mutation. Considering the spontaneous mutation of residue Phe86 in the other members of the MIO-dependent enzymes, the residue Phe86 could be mutated only in tyrosine.

<i>Containing the Residues</i>	<i>Round Number</i>								
	1	2	3	4	5	6	7	8	
Tyr80	X	X	X	X	X	X	X	X	X
Ala175, Ser176, Gly177, Asp178, Tyr322	X	X	X	X	X	X	X	X	X
Asn231, Gln319, Arg325, Asn458	X	X	X	X	WL <sup>+</sup>				
Phe86, Leu104, Leu108, Leu179, Ile431	HP	WL	HP*	WL*	WL*	WL*	HP*	WL	WL
Cys107, Tyr424, Lys427, Glu455, Gln459	HP	WL	WL	WL	WL	WL	WL	WL	WL
Leu227, Gly368, Phe371	X	X	X	X	X	WL	WL	WL	WL

Depending on the different rules of the limitations for residue mutation, the residue specification files for enzyme design were prepared. Then, the mutant library was established.

All designed enzymes were filtered through the score function by PyRosetta and the structure control by PyMOL (Chaudhury et al. 2010) (Schrödinger, LLC 2015). At the final stage of the enzyme design, the score lists were created in an output file with the results library. This score list included all information about the energy calculation of all atomic interactions in the designed enzyme. Among them, `total_score` represented the protein stability. The designed enzyme (around -5500) was higher than the nature enzyme (around -7300) since the lower score means a more stable structure of the protein. The other important value in enzyme design was `Interface_delta_X` term, which represented the energy calculation of the interaction between the ligand and designed enzyme. As shown in Table 24, the mutant library containing 27 single point mutations was constructed for subsequently designed enzyme screening.

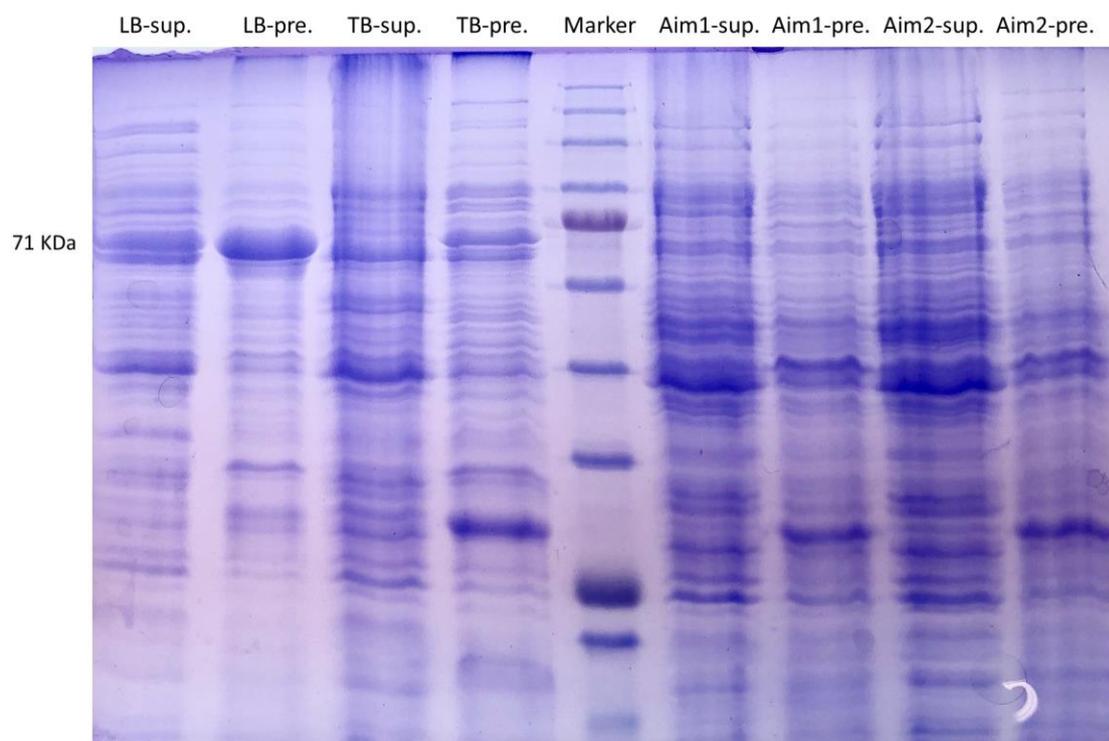
**Table 24:** List of the mutant library with different positions.

Position	Natural residues	Mutants
86	Phe	Tyr
104	Leu	Ser, Thr
108	Leu	Ser, Gln
179	Leu	Ser, Thr
424	Tyr	Cys, Asp, Asn, Gln, Met
427	Lys	Ala, Cys, Glu, Asn, Gln, Met
431	Ile	Val, Glu, Asn
455	Glu	Asn, Asp, His
458	Asn	Ala, Cys
459	Lys	Cys, Leu, His

#### 4.5 Optimization of enzyme expression

To improve the expression of the TchPAM, the cells were incubated in different mediums under various conditions. After around 24 h, the final OD<sub>600</sub> values were approximately 3 for the cells in LB-medium, while the *E. coli* with stronger growth in

TB-medium and auto-induction medium with a final  $OD_{600}$  of 7 and 8 respectively. As shown in Figure 20, the proportion of the desired enzyme in the expressed proteins was highest in the LB-medium.



**Figure 20: SDS-Gel for comparison of the expression in different mediums.** The band of TchPAM was observed between 63 kDa and 75 kDa. LB: The cells were incubated in the LB medium. TB: The cells were incubated in a TB medium. Aim: The cells were incubated in the auto-induction medium. Sup.: The supernatant of the enzyme solution after centrifugation, containing the soluble enzymes. Pre.: The precipitation of the enzyme solution after centrifugation.

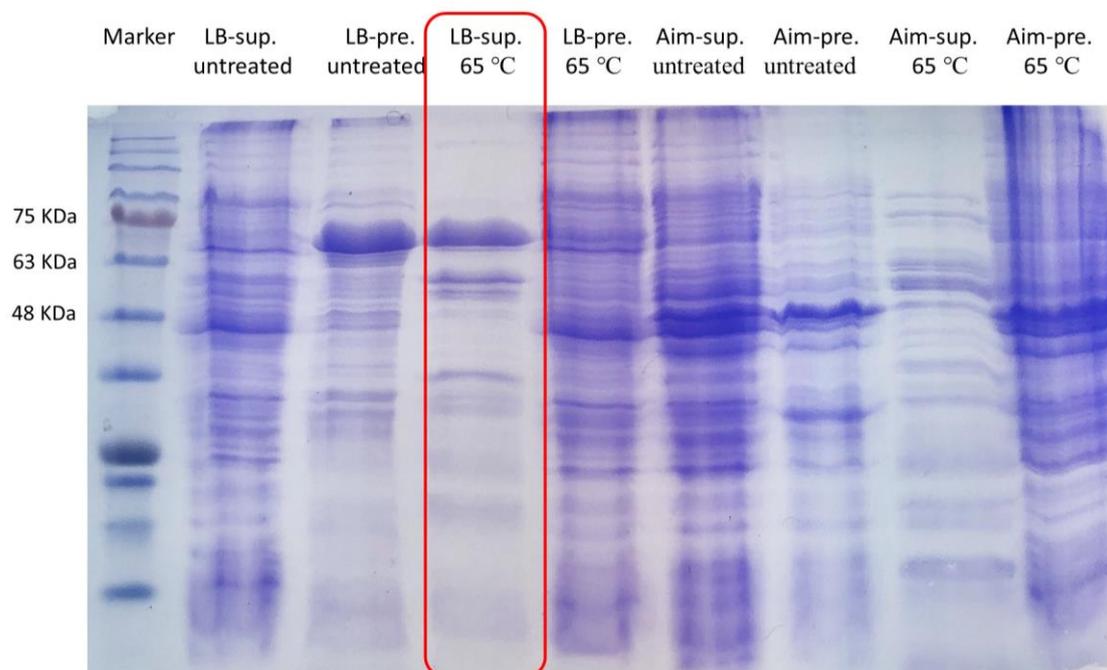
Therefore, the LB-medium was chosen for subsequent incubation and expression of the *E. coli*, which was optimized by varying the parameters, such as the concentration of the IPTG (from 0.1 to 1 mM), incubation temperature (16 °C, 18 °C, 22 °C, 25 °C, 30 °C, and 37 °C), and time (12 h, 15 h, 20 h, 22 h, 24 h), the reached  $OD_{600}$  for adding the IPTG (from 0.3 to 0.6). The concentration of IPTG did not significantly affect the number of grown cells and also the expressed proteins. More cells have been harvested at higher incubation temperatures, such as 30 °C and 37 °C. However, at these temperatures, the expected enzyme was insoluble, since the inclusion bodies were

---

formed. To gain more soluble enzymes, incubation at a lower temperature for a long time was carried out for the reduction of the rate of protein synthesis. Considering the efficiency, the culture of *E. coli* was incubated at 25 °C for 22 h. Similarly, IPTG was added at an earlier stage of incubation to avoid the formation of inclusion bodies. Combining the concentration of the expressed proteins and the proportion of the TchPAM, the most suitable condition was determined: The main culture of the *E. coli* was incubated in 100 ml LB-medium with 50 µg/ml kanamycin at 37 °C and 120 rpm until the OD<sub>600</sub> reached 0.3-0.4. After adding 0.1 mM IPTG, the shaking flasks were cooled down on the ice for 10 min and shaken for the other 22 h at 25 °C.

#### **4.6 Enzyme purification**

According to the previous study on TchPAM, the enzyme activity increased with enhancing temperature until 70 °C (Matthew M. Heberling et al. 2015). Therefore, the wild-type and designed mutants could be purified by heat treatment at 65 °C because of their great thermal stability. As shown in Figure 21, most of the undesired bands were removed after the heat treatment. However, there were still several undesired proteins with molecular weight in 48 – 63 kDa, which might affect the calculated total protein concentration, leading to a reduction of the calculation confidence of the specific activity during the comparison of the wild type and designed mutants. Moreover, the concentration of the desired enzyme was much lower than the incubation in LB-medium, even though the cells expressed more proteins by incubation in the auto-induction medium.



**Figure 21:** The SDS-Gel of the untreated and heat-treated wild type. The band of TchPAM was observed between 63 kDa and 75 kDa. LB: The cells were incubated in the LB medium. Aim: The cells were incubated in the auto-induction medium. Sup.: The supernatant of the enzyme solution after centrifugation, containing the soluble enzymes. Pre.: The precipitation of the enzyme solution after centrifugation. The wild type, which was purified through the heat treatment at 65 °C, was marked in a red frame.

Subsequently, the expressed enzyme was purified by the Äkta start protein purification system (GE Healthcare Life Sciences) with the HisTrap™ High-Performance column (1 ml). Even though the purified enzymes were successfully obtained, the duration of the purification by Äkta was at least one hour for just one sample. Considering the preparation of the samples in triplicate for wild type and the designed mutants (around 90 samples), it would spend a lot of time for the purification. Therefore, the His SpinTrap™ TALON® column was used for parallel purification of the samples of the wild type and mutants, by which could prepare various samples in one hour.

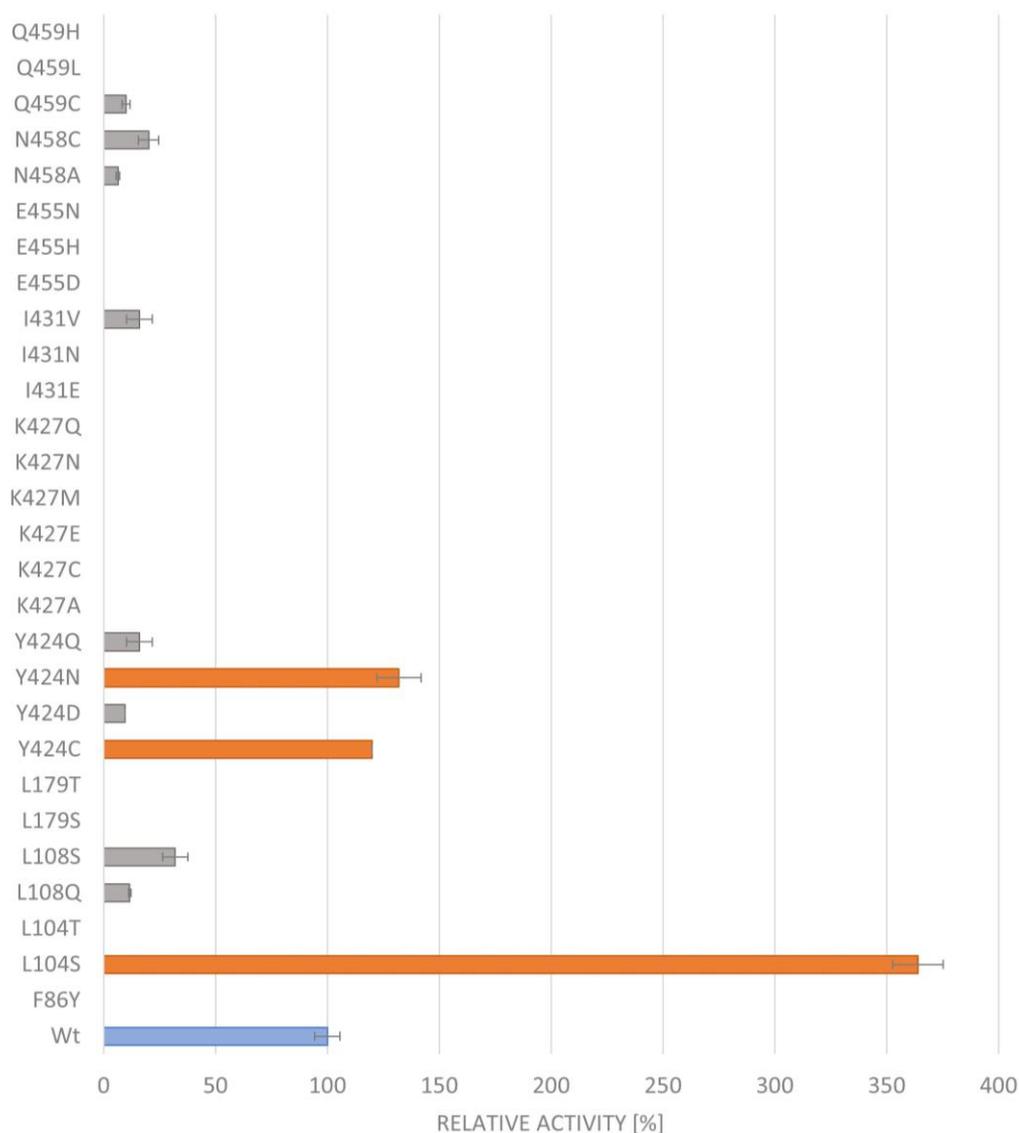
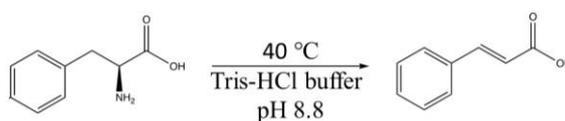
---

## 4.7 Comparison of relative activity in the ammonia elimination reaction

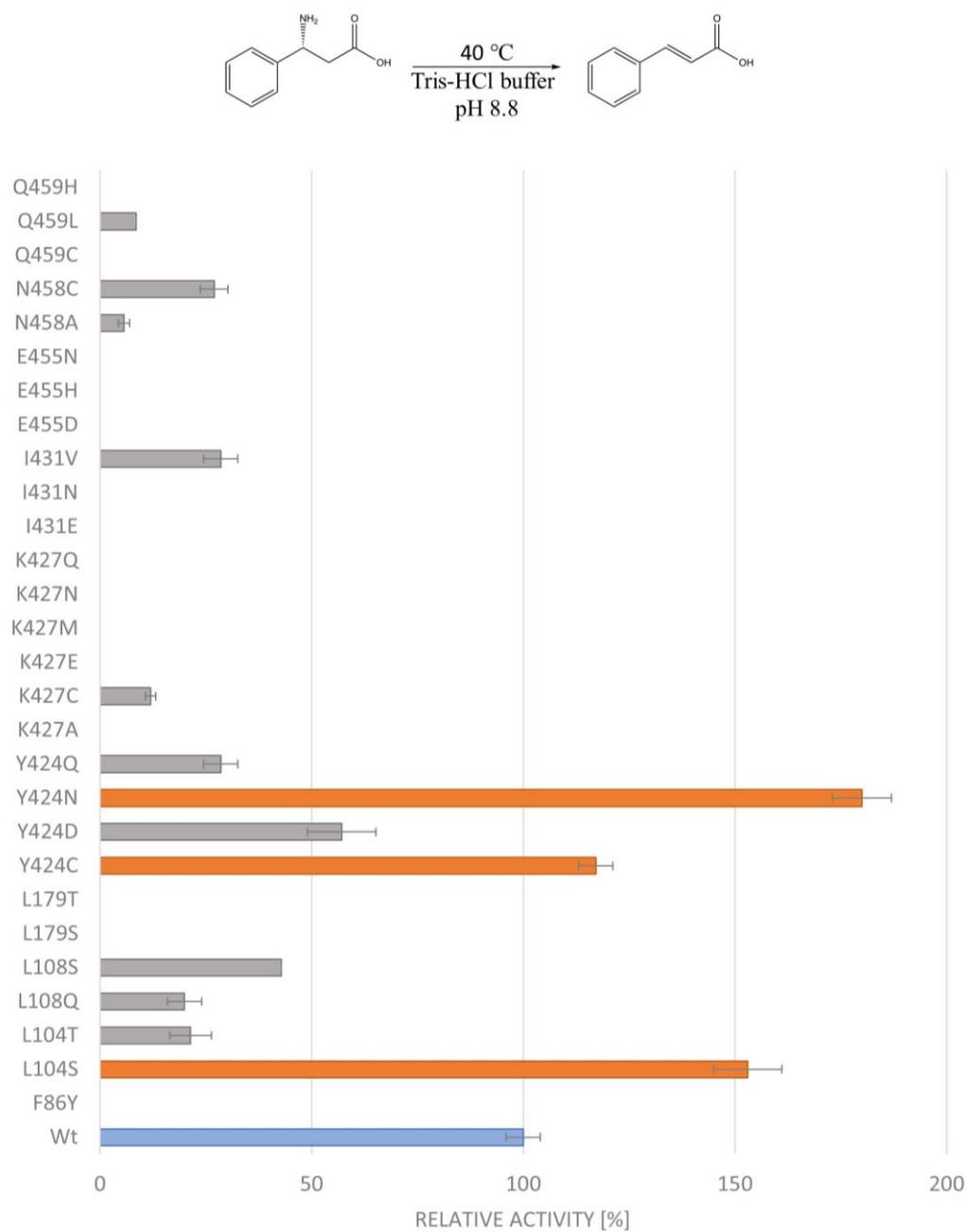
### 4.7.1 Ammonia elimination of phenylalanine

Initially, the activities of all selected mutants were compared to the activity of the wild type towards the enzymatic conversion of the phenylalanine in *trans*-cinnamic acid. For comparison of the substrate tolerance of the designed enzymes, the natural substrate (*S*)- $\alpha$ -phenylalanine and (*R*)- $\beta$ -phenylalanine were chosen for these measurements. As shown in Figure 22 and Figure 23, the mutant Leu104Ser with the highest activity towards (*S*)- $\alpha$ -phenylalanine (364 %) showed also an increased activity by using the (*R*)- $\beta$ -phenylalanine as substrate (153 %). Whereas, the mutant Tyr424Asn displayed a 180 % improvement in activity towards (*R*)- $\beta$ -phenylalanine compared to the wild type, as well as a similar enhancement of the activity with the (*S*)- $\alpha$ -phenylalanine (132 %). An analogous outcome was observed in the mutant Tyr424Cys with a lesser activity increasing towards both (*S*)- $\alpha$ -phenylalanine (120 %) and (*R*)- $\beta$ -phenylalanine (117 %). The mutant Tyr424Asp only retained the activity towards (*R*)- $\beta$ -phenylalanine of 57 %, but completely lost the activity to the (*S*)- $\alpha$ -phenylalanine. The other mutant, which was replaced the residue Tyr424 with glutamine, retained the activity with (*S*)- $\alpha$ -phenylalanine 16 %, while the remaining activity of this mutant with (*R*)- $\beta$ -phenylalanine were 29 %. A similar result was shown in Leu108Ser, Leu108Gln, Asn458Ala, and Asn458Cys, which have lower relative activities with both (*S*)- $\alpha$ -phenylalanine and (*R*)- $\beta$ -phenylalanine. Interestingly, the mutant Leu104Thr only retained the activity with the (*R*)- $\beta$ -phenylalanine of 21 %, but completely lost the activity towards its (*S*)- $\alpha$ -isomer. Except for the mutant Lys427Cys, which retained 12 % activity with (*R*)- $\beta$ -phenylalanine, the mutants by replacement of the residue Lys427 were not detected any activity towards phenylalanine. Similarly, the mutant Ile431Val retained the activity with the (*S*)- $\alpha$ -phenylalanine and (*R*)- $\beta$ -phenylalanine of 16 % and 29 % respectively. The other mutants Ile431Asn and Ile431Glu lost their activity with

the phenylalanine. Furthermore, the mutant Gln459Cys only retained the activity with (*S*)- $\alpha$ -phenylalanine while the mutant Gln459Leu only displayed the activity towards (*R*)- $\beta$ -phenylalanine. In addition, the variation of the residues Phe86, Leu179, and Glu455 led to the complete abolishment of the phenylalanine ammonia-lyase activity.



**Figure 22:** Comparison of the relative activity of the wild type and the designed mutants in the ammonia elimination of the (*S*)- $\alpha$ -phenylalanine. The wild type is colored in blue and the success mutants are colored in red.

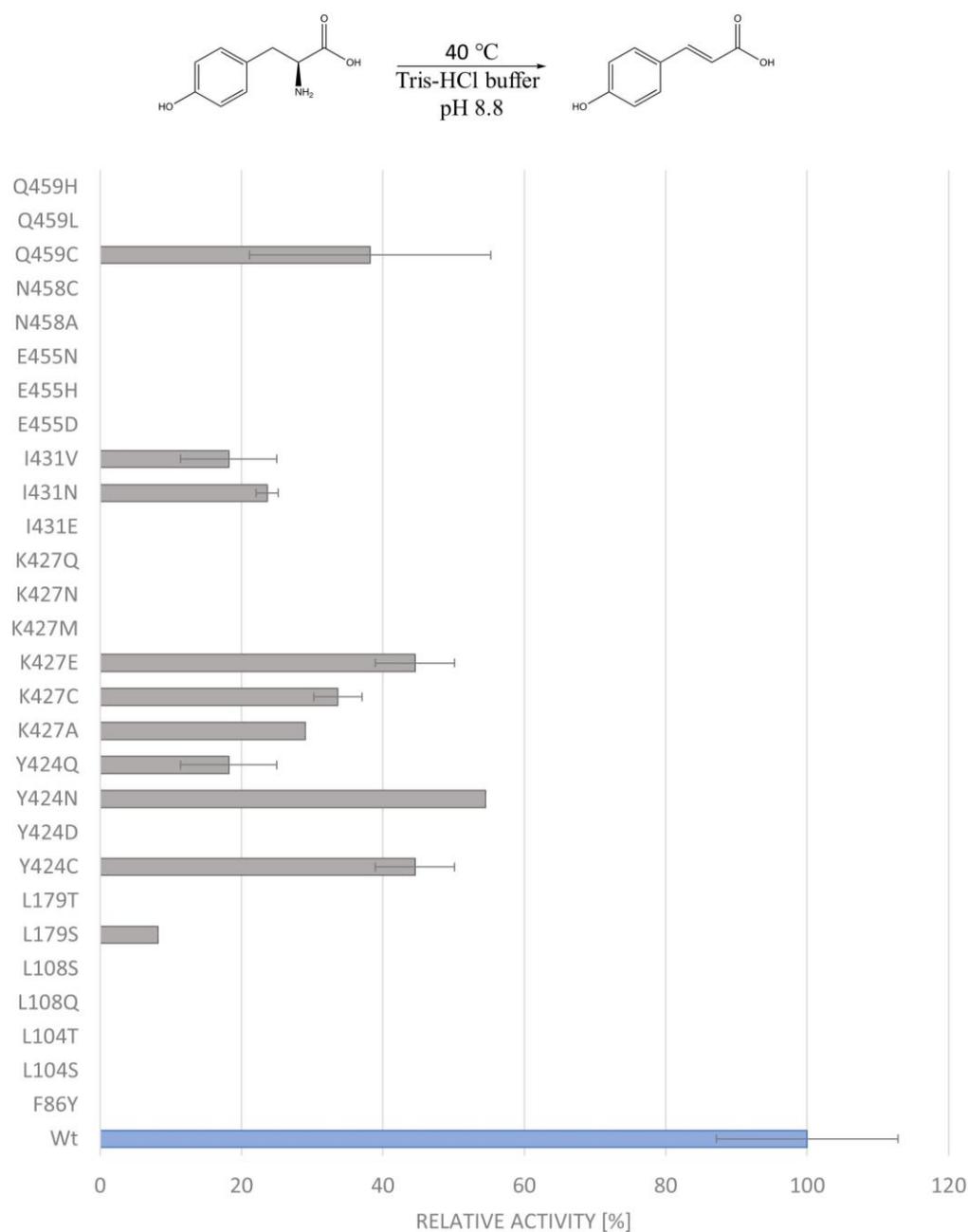


**Figure 23:** Comparison of the relative activity of the wild type and the designed mutants in the ammonia elimination of the *(R)*-β-phenylalanine. The wild type is colored in blue and the success mutants are colored in red.

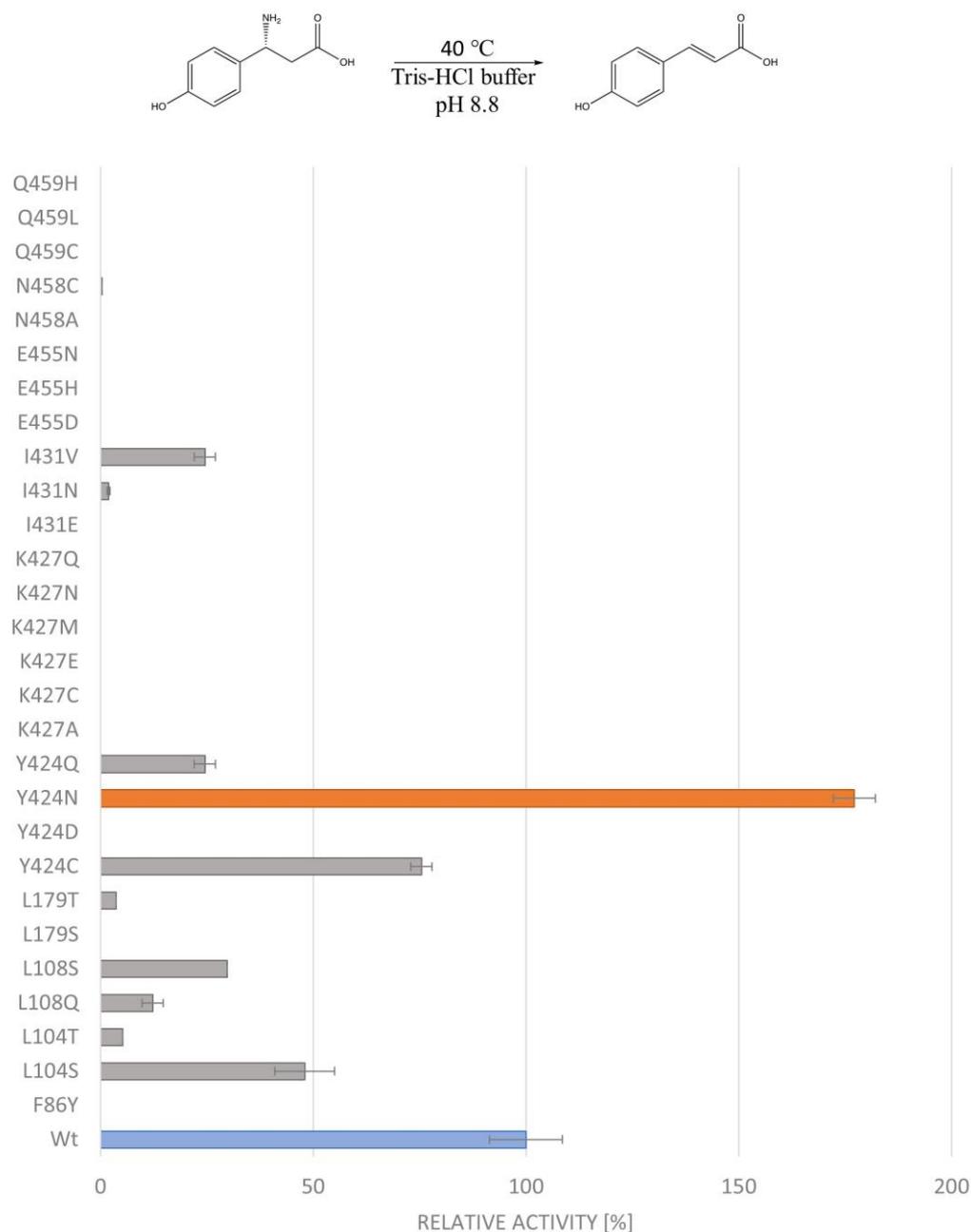
---

### 4.7.2 Ammonia elimination of tyrosine

To identify the substrate tolerance of the wild type and the designed mutants, the (*S*)- $\alpha$ -tyrosine and (*R*)- $\beta$ -tyrosine was used as the substrate in the measurement of the ammonia elimination reaction (Figure 24 and Figure 25). No mutant has shown the higher activity towards (*S*)- $\alpha$ -tyrosine compared to the wild type. Notably, in this case, the standard error values were also bigger than the results by using other substrates, since the specific activity of the wild type with (*S*)- $\alpha$ -tyrosine was relatively lower. However, the mutant Tyr424Asn has shown the highest activity towards (*R*)- $\beta$ -tyrosine (177 % of the wild type), which could be considered as the best-designed enzyme in the mutant library. Similarly, the mutant Tyr424Cys with the comparable activity to the wild type had also great potential for acceptance of the (*R*)- $\beta$ -tyrosine as substrate. The mutant Tyr424Gln retained 12 % activity towards (*S*)- $\alpha$ -tyrosine and 25 % activity with (*R*)- $\beta$ -tyrosine, while the mutant Tyr424Asp could not accept tyrosine as substrate. Unlike the phenylalanine, the mutant Leu104Ser displayed decreased activity towards (*R*)- $\beta$ -tyrosine (48 %) and completely lost the activity with (*S*)- $\alpha$ -tyrosine. The remaining activities of the mutants Leu104Thr, Leu108Ser, and Leu108Gln with (*R*)- $\beta$ -tyrosine were 5 %, 12 %, and 30 % respectively, despite the lack of the activity towards (*S*)- $\alpha$ -tyrosine. Both mutants Ile431Asn and Ile431Val had shown the activity with tyrosine. Interestingly, the mutants Leu179Ser and Gln459Cys only retained the activity towards (*S*)- $\alpha$ -tyrosine of 8 % and 38 % respectively, whereas, the mutants Leu179Thr and Asn458Cys displayed only the activity with (*R*)- $\beta$ -tyrosine of 4 % and 0.4 %. Except for the mutants Lys427Cys and Lys427Glu, which retained the activity with (*S*)- $\alpha$ -tyrosine of 34 % and 45 %, respectively, the other mutants with replaced residue Lys427 completely abolished the tyrosine ammonia-lyase activity. Compared to the phenylalanine, a similar outcome was observed in the results from the relative activity towards tyrosine in mutants Phe86 and Glu455, which completely lost the lyase-activity with tyrosine.



**Figure 24:** Comparison of the relative activity of the wild type and the designed mutants in the ammonia elimination of the (*S*)- $\alpha$ -tyrosine. The wild type is colored in blue.



**Figure 25:** Comparison of the relative activity of the wild type and the designed mutants in the ammonia elimination of the *(R)*-β-tyrosine. The wild type is colored in blue and the success mutants are colored in red.

#### 4.8 Kinetic parameters

As the best mutants of the library, the mutant Tyr424Asn with the highest activity towards *(R)*-β-tyrosine was chosen for subsequent measurement of the kinetic

parameters. To explain the variation of the active site in the designed enzyme, the other mutants with the replaced residue Tyr424, including Tyr424Cys, Tyr424Asp, and Tyr424Gln, were also analyzed their kinetic parameters. Furthermore, the mutant Leu104Ser, which displayed the highest activity towards phenylalanine, was also selected for kinetic analysis. To explore the influence of the residues in the aromatic binding pocket on the substrate selectivity, the mutants Leu108Ser and Ile431Val were also measured the kinetic parameters.

**Table 25:** Comparison of the kinetic parameters of the wild type and the mutants during the ammonia elimination of the (*S*)- $\alpha$ -phenylalanine. The best-designed mutants with the highest relative activity are colored in blue.

TchPAM	<i>(S)</i> - $\alpha$ -Phenylalanine		
	$K_M$ (mM)	$k_{cat}$ ( $s^{-1}$ )	$k_{cat}/K_M$ ( $mM^{-1} s^{-1}$ )
wt	0.032 $\pm$ 0.001	0.02 $\pm$ 0.001	0.625 $\pm$ 0.02
Y424C	0.048 $\pm$ 0.006	0.024 $\pm$ 0.002	0.506 $\pm$ 0.041
Y424D	n.d.	n.d.	n.d.
Y424N	0.075 $\pm$ 0.001	0.03 $\pm$ 0.002	0.407 $\pm$ 0.019
Y424Q	0.006 <sup>1</sup>	0.003 <sup>1</sup>	0.608 $\pm$ 0.095
L104S	0.856 $\pm$ 0.019	0.133 $\pm$ 0.001	0.156 $\pm$ 0.003
L108S	0.02 $\pm$ 0.003	0.006 $\pm$ 0.001	0.278 $\pm$ 0.004
<b>I431V</b>	<b>0.028 <math>\pm</math> 0.001</b>	<b>0.019<sup>1</sup></b>	<b>0.681 <math>\pm</math> 0.026</b>

The kinetic data have shown that the catalytic efficiency of the wild type was 0.625  $\pm$  0.02  $mM^{-1} s^{-1}$  during the ammonia elimination of the (*S*)- $\alpha$ -phenylalanine. Compared to the wild type, the mutants Tyr424Asn and Tyr424Cys had the lower catalytic efficiency with 0.506  $\pm$  0.041  $mM^{-1} s^{-1}$  and 0.407  $\pm$  0.019  $mM^{-1} s^{-1}$  respectively, because of the reduction of the binding affinity. The kinetic parameters of the mutant Ile431Val were comparable to the wild type. The catalytic efficiency of the mutant Tyr424Gln was also identified as the wild type and the mutant Ile431Val, albeit with a significantly lower turnover rate. Even though the  $K_M$ -value of the mutant Leu108Ser was comparable with the wild type, the catalytic efficiency of this mutant decreased by 0.278  $\pm$  0.004  $mM^{-1} s^{-1}$  due to the significantly low turnover rate.

Compared to the kinetic data in the ammonia elimination of (*S*)- $\alpha$ -phenylalanine, similar results came from the kinetic analysis by using the (*R*)- $\beta$ -phenylalanine as

substrate. The mutants Tyr424Asn and Tyr424Cys had shown the lower catalytic efficiency with  $0.320 \pm 0.044 \text{ mM}^{-1} \text{ s}^{-1}$  and  $0.109 \pm 0.01 \text{ mM}^{-1} \text{ s}^{-1}$  respectively, because of the improvement of the  $K_M$  values. Especially, the mutant Tyr424Asn displayed a 20-fold increased  $K_M$  value with  $1.266 \pm 0.108 \text{ mM}$ , compared to the wild type with  $0.062 \pm 0.012 \text{ mM}$ . An analogous outcome was observed in the mutant Leu104 with a significantly lower catalytic efficiency because of the higher  $K_M$  value.

Not only the mutant Ile431Val but also the mutant Leu108Ser displayed comparable kinetic parameters with the wild type. The catalytic efficiencies of the mutants Tyr424Asp and Tyr424Gln were identical with  $0.053 \pm 0.044 \text{ mM}^{-1} \text{ s}^{-1}$  due to their comparable  $K_M$  values and turnover rates.

**Table 26:** Comparison of the kinetic parameters of the wild type and the mutants during the ammonia elimination of the (*R*)- $\beta$ -phenylalanine.

TchPAM	<b>(<i>R</i>)-<math>\beta</math>-Phenylalanine</b>		
	<b><math>K_M</math> (mM)</b>	<b><math>k_{cat}</math> (<math>\text{s}^{-1}</math>)</b>	<b><math>k_{cat}/K_M</math> (<math>\text{mM}^{-1} \text{ s}^{-1}</math>)</b>
wt	$0.062 \pm 0.012$	$0.026 \pm 0.002$	$0.427 \pm 0.078$
Y424C	$0.110 \pm 0.017$	$0.035 \pm 0.004$	$0.320 \pm 0.044$
Y424D	$0.458 \pm 0.091$	$0.024 \pm 0.006$	$0.053 \pm 0.006$
Y424N	$1.266 \pm 0.108$	$0.138 \pm 0.002$	$0.109 \pm 0.01$
Y424Q	$0.422 \pm 0.073$	$0.022 \pm 0.005$	$0.053 \pm 0.006$
L104S	$1.311 \pm 0.308$	$0.098 \pm 0.019$	$0.077 \pm 0.019$
L108S	$0.039 \pm 0.003$	$0.013 \pm 0.001$	$0.337 \pm 0.035$
I431V	$0.076 \pm 0.013$	$0.027 \pm 0.001$	$0.357 \pm 0.059$

Since the (*S*)- $\alpha$ -tyrosine was not the natural substrate for the TchPAM, the  $K_M$ -value of the wild type was significantly high with  $2.435 \pm 0.160 \text{ mM}$ . Except for the mutants Tyr424Asp, Tyr424Gln, Leu104Ser, and Leu108Ser, which were not detected the activity towards (*S*)- $\alpha$ -tyrosine, the other mutants displayed the higher catalytic efficiency due to their lower  $K_M$  values, including the mutants Tyr424Asn, Tyr424Cys, and Ile431Val. Especially, the mutant Tyr424Asn and Tyr424Cys had shown a massive 7-fold reduction of the  $K_M$  values with  $0.347 \pm 0.065 \text{ mM}^{-1} \text{ s}^{-1}$  and  $0.330 \pm 0.024 \text{ mM}^{-1} \text{ s}^{-1}$  respectively.

**Table 27:** Comparison of the kinetic parameters of the wild type and the mutants during the ammonia elimination of the (*S*)- $\alpha$ -Tyrosine. The best-designed mutants with the highest relative activity are colored in blue.

TchPAM	<i>(S)</i> - $\alpha$ -Tyrosine		
	$K_M$ (mM)	$k_{cat}$ ( $s^{-1}$ )	$k_{cat}/K_M$ ( $mM^{-1} s^{-1}$ )
wt	2.435 $\pm$ 0.160	0.029 $\pm$ 0.005	0.011 $\pm$ 0.002
<b>Y424C</b>	<b>0.347 <math>\pm</math> 0.065</b>	<b>0.01 <math>\pm</math> 0.002</b>	<b>0.03 <math>\pm</math> 0.004</b>
Y424D	n.d.	n.d.	n.d.
<b>Y424N</b>	<b>0.330 <math>\pm</math> 0.024</b>	<b>0.007<sup>1</sup></b>	<b>0.021 <math>\pm</math> 0.001</b>
Y424Q	n.d.	n.d.	n.d.
L104S	n.d.	n.d.	n.d.
L108S	n.d.	n.d.	n.d.
<b>I431V</b>	<b>0.506 <math>\pm</math> 0.063</b>	<b>0.013 <math>\pm</math> 0.001</b>	<b>0.026 <math>\pm</math> 0.002</b>

Regarding (*R*)- $\beta$ -tyrosine, the mutants Tyr424Cys and Tyr424Asn displayed an approximately 2-fold improvement of the catalytic efficiency with  $0.281 \pm 0.023 \text{ mM}^{-1} \text{ s}^{-1}$  and  $0.234 \pm 0.013 \text{ mM}^{-1} \text{ s}^{-1}$  respectively, compared to the wild type with  $0.130 \pm 0.01 \text{ mM}^{-1} \text{ s}^{-1}$ . The mutant Tyr424Asn provided a comparable  $K_M$  value with the wild type, but a nearly 2-fold increased turnover rate. Whereas, the kinetic data have shown that the mutant Tyr424Cys with great binding affinity retained a similar turnover rate with the wild type. The mutant Tyr424Gln decreased the catalytic efficiency because of the lower turnover value, while the Tyr424Asp completely lost the activity with (*R*)- $\beta$ -tyrosine. In contrast with the other mutants, the mutant Ile431Val, which displayed a catalytic efficiency of  $0.137 \pm 0.003 \text{ mM}^{-1} \text{ s}^{-1}$ , shared similar kinetic parameters with the wild type. It was worth mentioning the 6-fold improvement of the  $K_M$  value with  $2.715 \pm 0.014 \text{ mM}$  of the mutant Leu104Ser, leading to a notable reduction of the catalytic efficiency. Whereas, the mutant Leu108Ser with a similar binding affinity compared to the wild type decreased the turnover value by  $0.021 \pm 0.002 \text{ s}^{-1}$ .

**Table 28:** Comparison of the kinetic parameters of the wild type and the mutants during the ammonia elimination of the (*R*)- $\beta$ -Tyrosine. The best-designed mutants with the highest relative activity are colored in blue.

TchPAM	<b>(<i>R</i>)-<math>\beta</math>-Tyrosine</b>		
	<b>K<sub>M</sub> (mM)</b>	<b>k<sub>cat</sub> (s<sup>-1</sup>)</b>	<b>k<sub>cat</sub>/K<sub>M</sub> (mM<sup>-1</sup> s<sup>-1</sup>)</b>
wt	0.465 $\pm$ 0.008	0.06 $\pm$ 0.001	0.130 $\pm$ 0.01
<b>Y424C</b>	<b>0.299 <math>\pm</math> 0.005</b>	<b>0.084 <math>\pm</math> 0.007</b>	<b>0.281 <math>\pm</math> 0.023</b>
Y424D	n.d.	n.d.	n.d.
<b>Y424N</b>	<b>0.480 <math>\pm</math> 0.015</b>	<b>0.112 <math>\pm</math> 0.003</b>	<b>0.234 <math>\pm</math> 0.013</b>
Y424Q	0.142 $\pm$ 0.001	0.014 $\pm$ 0.001	0.098 $\pm$ 0.008
L104S	2.715 $\pm$ 0.014	0.08 $\pm$ 0.016	0.03 $\pm$ 0.006
L108S	0.472 $\pm$ 0.054	0.021 $\pm$ 0.002	0.045 $\pm$ 0.001
<b>I431V</b>	<b>0.618 <math>\pm</math> 0.031</b>	<b>0.084 <math>\pm</math> 0.003</b>	<b>0.137 <math>\pm</math> 0.003</b>

#### 4.9 HPLC analysis for amination reaction

According to the results from the kinetic analysis, the mutants Tyr424Asn, Tyr424Cys, Leu108Ser, and Ile431Val were considered as the candidates for the production of the (*R*)- $\beta$ -tyrosine from the *trans-p*-hydroxycinnamic acid. To identify the substrate tolerance during the ammonia addition reaction, the specific activities of the wild type and the selected mutants were measured by HPLC. The candidate mutants Tyr424Asn, Tyr424Cys, and Ile431Val presented the tyrosine ammonia-lyase activity by using the *trans-p*-hydroxycinnamic acid as substrate. Strikingly, the mutants Tyr424Asn and Tyr424Cys could produce the (*R*)- $\beta$ -tyrosine with excellent enantiopurity ( $ee^{\beta}_R > 99\%$ ). The specific activities of the mutants Tyr424Asn and Tyr424Cys were  $0.104 \pm 0.002$  U/mg and  $0.141 \pm 0.001$  U/mg respectively (Peng et al. 2021). Whereas, the mutant Ile431Val was confirmed to produce both (*R*)- $\beta$ -tyrosine and (*S*)- $\beta$ -tyrosine. Interestingly, the enantioselectivity of this mutant was switched from the (*R*)- $\beta$ -aromatic amino acids to the (*S*)- $\beta$ -conformers. In addition, the mutant Leu108Ser could not accept the *trans-p*-hydroxycinnamic acid as substrate, which was similar to the wild type.

**Table 29:** Comparison of the specific activity and enantiomeric excess value of the wild type and the mutants Tyr424Asn, Tyr424Cys, Leu108Ser, and Ile431Val during the ammonia addition of the *trans-p*-hydroxycinnamic acid.

Enzyme	Specific Activity (U/mg)				ee <sup>β<sub>R</sub></sup>	ee <sup>β<sub>S</sub></sup>	Reference
	( <i>S</i> )-α-tyrosine	( <i>R</i> )-α-tyrosine	( <i>S</i> )-β-tyrosine	( <i>R</i> )-β-tyrosine			
Wt	n.d.	n.d.	n.d.	n.d.			(Wu et al. 2009)
Y424N	n.d.	n.d.	n.d.	0.104 ± 0.002	> 99%		This work
Y424C	n.d.	n.d.	n.d.	0.141 ± 0.001	> 99%		This work
L108S	n.d.	n.d.	n.d.	n.d.			This work
I431V	n.d.	n.d.	0.23 ± 0.006	0.04 ± 0.009		> 29%	This work

The conversion rates of the mutants Tyr424Asn and Tyr424Cys were 1.7 % and 1.3 % after 24 h, respectively. When the reaction was carried out for 48 h, the conversion rate of the mutant Tyr424Cys was improved to 2.3 %. However, it is the first time that enantiopure (*R*)-β-tyrosine has been produced from *trans-p*-hydroxycinnamic acid by using designed phenylalanine aminomutase (Peng et al. 2021). The mutant Ile431Val could produce the (*R*)-β-tyrosine in the conversion of 1.3 % and (*S*)-β-tyrosine of 2.5 % after 24 h. It was possible to achieve the total conversion of the β-tyrosine in 10 % after 3 days, albeit with the poor enantiomeric excess value.

**Table 30:** Comparison of the specific activity and enantiomeric excess value of the wild type and the mutants Tyr424Asn, Tyr424Cys, Leu104Ser, and Ile431Val during the ammonia addition of the *trans-p*-cinnamic acid.

Enzyme	Specific Activity (U/mg)		ee <sup>α<sub>R</sub></sup>	ee <sup>α<sub>S</sub></sup>
	( <i>S</i> )-α-phenylalanine	( <i>R</i> )-α-phenylalanine		
Wt	2.63 ± 0.48	n.d.		> 99%
Y424N	2.43 ± 0.18	n.d.		> 99%
Y424C	2.43 ± 0.18	n.d.		> 99%
L104S	n.d.	1.23 ± 0.1	> 35%	
I431V	2.24 ± 0.04	n.d.		> 99%

Similarly, the ammonia addition reaction of the wild type and mutants Tyr424Asn, Tyr424Cys, Leu104Ser, and Ile431Val were also measured by using the *trans*-cinnamic acid as substrate. Compared to the wild type, the mutants Tyr424Asn, Tyr424Cys, and Ile431 could also produce the (*S*)-α-phenylalanine with high enantiopurity from *trans*-

---

cinnamic acid. The wild type and the mutants Tyr424Asn, Tyr424Cys, and Ile431Val also shared the comparable specific activity with  $2.63 \pm 0.48$  U/mg,  $2.43 \pm 0.18$  U/mg,  $2.43 \pm 0.18$  U/mg, and  $2.24 \pm 0.04$  U/mg, respectively. Whereas, the mutation of the residue Leu104 in serine switched the (*S*)- $\alpha$ -selective TchPAM to the enzyme, which generated the (*R*)- $\beta$ -phenylalanine as the main product during the ammonia addition reaction of the *trans*-cinnamic acid. After 7 h, a tiny amount of the (*S*)- $\alpha$ -phenylalanine was observed in the detected samples.

During the ammonia addition of the *trans*-cinnamic acid for generation of the enantiopure (*S*)- $\alpha$ -phenylalanine, the conversion rates of the wild type and the mutants Tyr424Asn, Tyr424Cys, and Ile431Val were 32.2%, 29.8 %, 34.4 %, and 36.6 %, respectively. These values were lower than the result from the literature, which was around 50 % conversion of (*S*)- $\alpha$ -phenylalanine in wild type of TchPAM. Moreover, the mutant Leu104Ser could produce the (*R*)- $\alpha$ -phenylalanine in 6.9 % conversion and (*S*)- $\alpha$ -phenylalanine in 3 % conversion after 24 h.

---

## 5 DISCUSSION

### 5.1 Chemical synthesis of OH-PheDU

To prepare the substrate for the subsequent enzymatic reaction, OH-PheDU has been synthesized according to the protocol from Vuano et al (Vuano et al. 2000). This synthesis was recognized for the production of PheDU by heating an equimolecular mixture of cinnamic acid and urea at 250 °C. As the suggestion by Vuano et al., PheDU has been synthesized by an intramolecular condensation after a nucleophilic attack of the amino group from urea on  $\beta$ -position of cinnamic acid. The  $\beta$ -phenyl-propionic acid as an intermediate was produced before the cyclization. The reaction parameters such as reactants ratio, temperature, and reaction time, strongly affected the production of PheDU and  $\beta$ -phenyl-propionic acid (Vuano et al. 2000). To optimize the reaction time, Chang et al. have reported a similar synthesis process in a sealed tube (Chang and Jeung 1991). The yields of PheDU were achieved to around 35 % by both of these thermal synthesis processes (Chang and Jeung 1991) (Vuano et al. 2000). Therefore, in this project, the synthesis of OH-PheDU has been carried out under these two conditions. Subsequently, a microwave reactor has been also used for the improvement of the product yield.

However, only a few amounts of OH-PheDU could be identified. The produced styrene was indicated that cleavage of bond has occurred during the thermal reaction, which has been never mentioned in previous kinds of literature about the synthesis of PheDU (Chang and Jeung 1991) (Vuano et al. 2000). The electron-donating group (hydroxyl group) on para-substituted increases the electron density of the benzene ring and activates the carbon atom on the  $\beta$ -position (Deruiter 2005). The double bond of hydroxycinnamic acid is less stable in contrast with cinnamic acid. It might be the reason for the low production yield of OH-PheDU in the chemical synthesis. Even though the thermal reaction has been also carried out at a lower reaction temperature (120 °C), which was the lowest melting temperature of the substrate mixture, the

---

cleavage of the double bond of hydroxycinnamic acid could not be avoided. Furthermore, the synthesis has been also performed in methanol for reduction of the reaction temperature (from 50 °C to 80 °C). Unfortunately, hydroxycinnamic acid did not react with urea under this condition. In summary, it could be suggested that the bond cleavage might have occurred before the nucleophilic attack of the urea on hydroxycinnamic acid.

## 5.2 The origin of MIO-dependent enzymes

Bacteria compete with their neighbors in a challenging living community. Considering the density of microorganisms within the community and scarcity of resources, such as nutrients and space, ancient bacteria have evolved competitive phenotypes. These phenotypes can release small antibiotic compounds in the living environment to outcompete and displace the other microorganisms by poisoning them (Ghoul and Mitri 2016). To assemble the antibiotic molecules, some ancient bacteria employ TAM for conversion of proteinogenic amino acids in  $\beta$ -tyrosine, which is a common precursor of the enediyne antibiotics. These bacterial species are commonly isolated from soil, for instance, *Cupriavidus*, *Streptomyces*, and *Myxococcus*, suggesting that the common ancestor may acquire these genes from soil bacteria during the early stage of the colonization of a terrestrial environment.

Apart from the RsTAL from *Rhodobacter sphaeroides*, the other MIO-dependent enzymes, which can catalyze tyrosine for the production of secondary metabolites, have tyrosine aminomutase activity as a side activity. The enantioselectivity of these TAM is usually consistent with the configuration of the assembled  $\beta$ -tyrosine in the secondary metabolites. For example, the *Streptomyces globisporus* with (*S*)-selective TAM can form the (*S*)-3-chloro-5-hydroxy- $\beta$ -tyrosine, while the (*R*)-selective TAM in *Chondromyces crocatus* can release the cytotoxic chondramides, which contains (*R*)- $\beta$ -tyrosine (Rachid et al. 2007). In addition, the evolution of bacterial TAMs is independent, which is supported by the distribution of the enantioselectivity within the corresponding bacterial clades (Krug and Müller 2009). In the phylogenetic tree, the

---

most deeply branching line within the TAL/TAM bacterial clades is formed by RsTAL, suggesting that the genes for production of the antibiotic metabolites in TAM bacteria are acquired from the ancient bacteria with TAL for competition with the neighboring microorganisms in community.

The cytosolic HAL, which is encoded by the *hutH* gene, serves as a basal and core enzyme in the metabolic degradation pathway of histidine, which is widespread among the bacteria (Fuchs and Kane 1985). The bacteria without the *hutH* gene lost the ability to utilize histidine as carbon and nitrogen sources. However, the lack of the Hut pathway is not lethal for bacteria. According to the results from the phylogenetic analysis, the *hutH* gene may evolve in the ancestor of the *Streptomyces* bacteria and/or thermophilic or halophilic archaea, which acquire the HAL-like enzyme gene from the ancient TAL bacteria. The arrangement of the *hutH* gene is irregular in the clades of archaea and lower eukaryotes. However, the HAL is strictly required by all vertebrates, especially mammals. The lack of the HAL leads to a disease because of an insufficient capacity for degradation of histidine (Taylor et al. 1991) (Ghadimi, H., M. W. Partington 1962). The only example of protozoa with HAL is *D. discoideum*, comprising two homologs (Emiliani et al. 2009). The one appears in the eukaryotic HAL clade, while the other one has been placed in a lineage close to the cyanobacteria clade. Until recently, no Hut pathway is observed in the cyanobacteria. However, it is worth mentioning that these bacteria employ the HAL-like enzyme for the degradation of phenylalanine.

This HAL-like enzyme belongs to the MIO-dependent enzyme family, which is recognized as PAL. The bacterial PALs formed the deepest branching lineage in the PAL clades, indicating that the PALs in the plants and fungi may have evolved from the bacterial PAL. The ancestral fungi can form mutualistic associations with green algae or cyanobacteria. This symbiotic association may have occurred before the colonization of land, which is supported by the evolutionary analysis of *Glomeromycotina* (Ghadimi, H., M. W. Partington 1962). However, neither HAL nor PAL has been found in the

---

available genomic information in algae. On the other hand, the bacteria and fungi, which contain the *pal* gene, as well as the amoebae with HAL, can be isolated from the soil environments. It may indicate the co-evolutionary relationship between bacterial, fungi, and animals in the early soil environments. At that time, the corresponding genes could be transferred between them (Emiliani et al. 2009).

Generally, the multigene family of *pal* includes one to five members in most plant species. The expressions of these genes are different in various tissues or under numerous environmental stresses. For instance, PALs are encoded by *pal1*, *pal2*, *pal3*, and *pal4* in *Arabidopsis thaliana*. In stem tissue, *pal3* is expressed at a lower level compared to the others. Among these genes, *pal1* and *pal2* shares common promoter elements and mainly enrich in roots and stems, while *pal4* exists in seeds (Raes et al. 2003). Similarly, five different *pal* genes are expressed in *P. taeda*, which indicates that gene duplication can be traced back to gymnosperms (Bagal et al. 2012). In addition, the tandem arrangement of seven *pal* genes with four pseudogenes in two chromosomes from *Cucumis sativus* indicates that these genes are duplicated during the splitting of cucumber from the other dicots (Shang et al. 2012). Plants have gained the *pal* gene from a bacterial ancestor and expanded through gene duplication.

PTAL is a special PAL, especially in the Poaceae family. The expression of PTAL is associated with the synthesis process of lignin with more syringyl composition (Barros and Dixon 2020). Among eight *pal* genes in the *Brachypodium distachyon*, only one *pal* gene, which was named in BdPTAL1, can encode the phenylalanine ammonia-lyase with TAL activity (Barros et al. 2016). It could be assumed that these special *pal* genes have obtained the ability to express PALs, which can accept tyrosine as substrate, after the gene duplication during the separation of monocots and dicots. The phylogenetic analysis and multiple amino acid sequence alignment might indicate the origin of the genes of PTAL in monocots. Then fungi might gain these corresponding genes from plants through HGT.

The PAM catalyzes the rate-limiting step in the assembly process of the paclitaxel side

---

chain in yew species (Strom et al. 2012) (Steele et al. 2005). The PAMs can also be extracted from endophytic fungi, such as the *Penicillium* species (Soliman and Raizada 2013). However, no available data about plant PAL with PAM activity was found in the other gymnosperms, such as *P. taeda*. It seems like there are two possibilities of the PAM origin. The PAM in yew species might evolve from PAL in gymnosperms or from the endophytic fungi. Moreover, the endophytic fungi might obtain the corresponding genes from saprophytic or pathogenic fungi, especially *Pestalotiopsis malicola*, which is isolated from soil (Bi et al. 2011). The multiple sequence analysis of the gymnosperms and fungi suggests that the PAM in yew species has been probably obtained from PAL in gymnosperms through mutation and then transferred the corresponding genes to the endophytic fungi.

### **5.3 The designed TchPAM for enantiopure (*R*)- $\beta$ -tyrosine production: Tyr424Asn and Tyr424Cys**

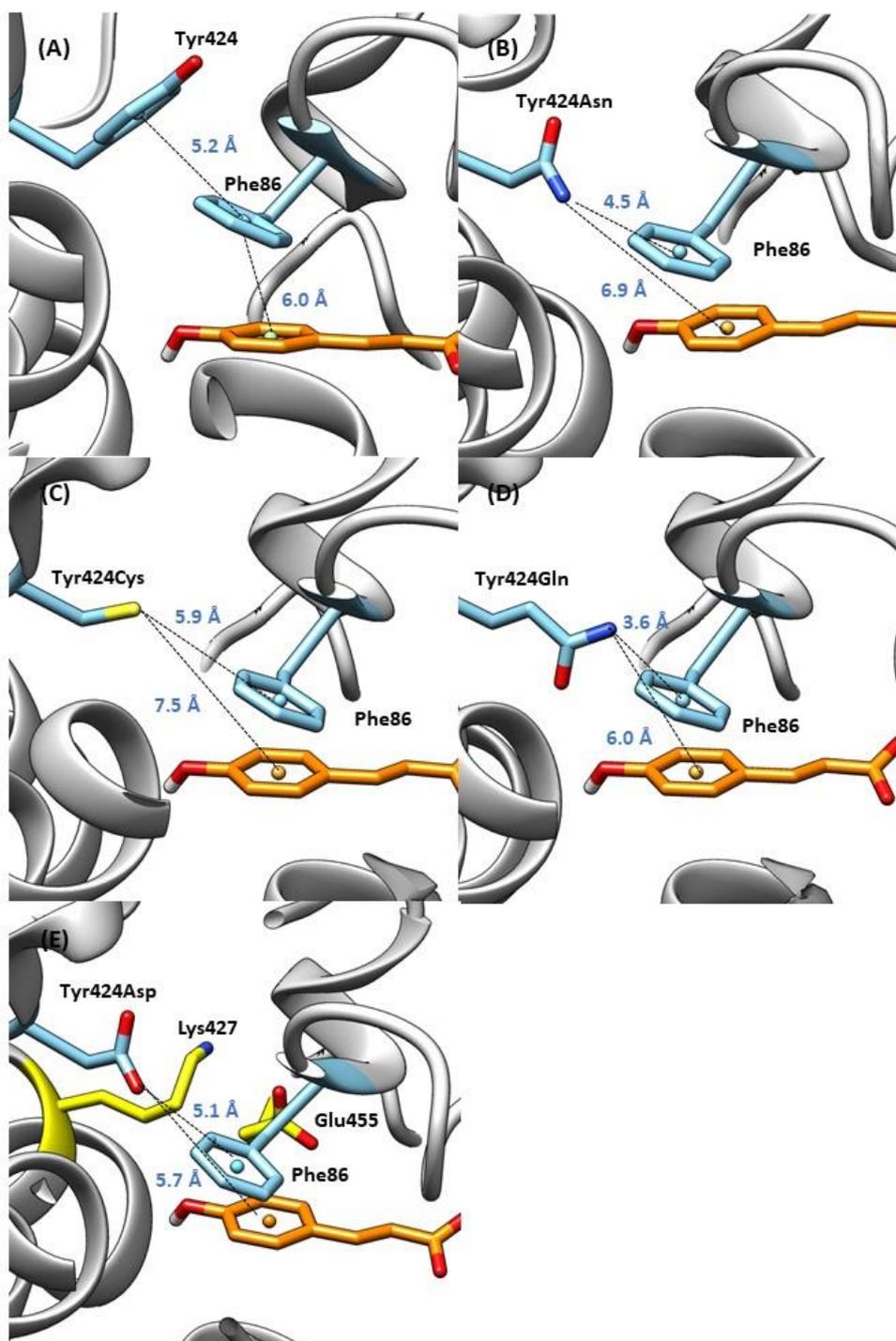
#### **5.3.1 Influence of the mutations on ammonia elimination reaction**

The best results of the mutant library, the mutants Tyr424Asn and Tyr424Cys enhance not only the tyrosine acceptance but also the activity with the natural substrates. They can also produce the (*R*)- $\beta$ -phenylalanine with high enantiopurity from the *trans*-cinnamic acid.

The residue Tyr424 may have no direct influence on the substrate positioning in the wild type, because of the large distance (8.1 Å) between the centroids of the phenolic ring of the residue and the aromatic ring of the substrate (Figure 26). Considering the position of the residue Phe86 with a benzene ring, it can be observed the sandwich coordinated aromatic rings in the active site of TchPAM, which also involve the aromatic ring of the substrate. As a highly conserved residue in the PAM and PAL, the residue Phe86 plays a key role in the substrate binding and activation, which is supported by the complete loss of activity in the mutant Phe86Tyr. The formation of this sandwich-like model may not be attributed to the  $\pi$ -stacking due to the large

---

separation between the ring planes (Janiak 2000)(Deng et al. 2020). However, during protein folding, the steric hindrance may change the position and orientation of the aromatic residues, which are placed in close contact with each other. The variation of the residue Tyr424 to a smaller cysteine also does not change the position of the aromatic ring of the substrate. However, the lack of the huge aromatic side chain at this position may provide more space to the neighboring residues, such as Phe86, Lys427, Ile431, and Glu455. It is consistent with the identical enzyme activity of the mutant Tyr424Cys, compared to the wild type. Whereas, in the mutant Tyr424Asn, a partial positively charged amide group is located above the aromatic ring of the substrate, which may generate an attractive force towards the ring center of the substrate (Petsko 1986)(Biot et al. 2002). Considering the distance between the amide group of the residue Tyr424Asn and the ring centroid of the binding substrate, the improvement of the enzyme activity in this mutant may be attributed to the prevention of the steric clash. In a similar structural point of view, a deviation is caused by the residue Tyr424Gln with a methylene longer side chain than asparagine, leading to a notable reduction of the enzyme activity with the phenylalanine, especially, the  $\alpha$ -regioisomer. In contrast with the asparagine, the residue Tyr424Asp with the negatively charged side chain generates a repulsive interaction with the ring center of the substrate, which may lead to deviation of the aromatic ring of the substrate (Philip et al. 2011). Moreover, the carboxyl group of the residue Tyr424Asp may form a salt bridge with the amino group of the neighboring residue Lys427, which orients towards the residue Gln455 to avoid the steric clash from the phenolic side chain of the residue Tyr424 in the wild type. Instead of the aromatic side chain, the smaller residues at position 424 may provide more space for the neighboring residues, leading to the varying of the conformation in the active site of the TchPAM. In addition, the residues Lys427 and Glu455 are confirmed to be the essential residues in the active site of the TchPAM, since all of the mutants at this two position completely lost their activity. Therefore, any minimal shift of these two residues may lead to a notable decrease of the enzyme activity.



**Figure 26:** The distance and the proposed interaction between the residue Tyr424 and the aromatic ring of the substrate in (A) the wild type, the mutants (B) Tyr424Asn, (C) Tyr424Cys, (D) Tyr424Gln, and (E) Tyr424Asp. The neighboring residue Arg427 and Glu455 are shown in (E).

### 5.3.2 Influence of the mutations on ammonia addition reaction

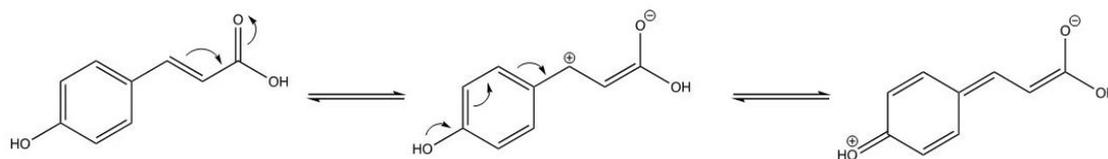
Upon the available sequence data from the SwissProt, only two (*R*)- $\beta$ -selective tyrosine aminomutases have been found: the CmdF from *Chondromyces crocatus* and the CmTAL from *Cupriavidus metallidurans*. Both enzymes can produce the (*R*)- $\beta$ -tyrosine and (*S*)- $\beta$ -tyrosine from the *trans-p*-hydroxycinnamic acid (Bairoch et al. 2004). Even though the early mutagenesis study on the CmdF reported the improvement of the enantioselectivity of the generated (*R*)- $\beta$ -tyrosine during the isomerization reaction by using the  $\alpha$ -tyrosine as the substrate. However, the attempt to turn TAM into TAL by site-directed mutagenesis was unsuccessful (Krug and Müller 2009). Similarly, the only one (*R*)-the  $\beta$ -selective example is OsTAM from *Oryza sativa*, which is also difficult to yield the tyrosine from *trans-p*-hydroxycinnamic acid (Yan et al. 2015). Therefore, the TchPAM, yielding the (*R*)- $\beta$ -phenylalanine with excellent enantiopurity, presents the great potential for generation of the (*R*)- $\beta$ -tyrosine (Wu et al. 2009).

**Table 31:** Comparison of the enantioselectivity of the tyrosine aminomutases from the SwissProt (Bairoch et al. 2004).

Enzyme Name	Organism	Enantioselectivity	Uniprot ID
CmdF	<i>Chondromyces crocatus</i>	( <i>R</i> )	Q0VZ68
CmTAL	<i>Cupriavidus metallidurans</i>	( <i>R</i> )	Q1LRV9
MfTAM	<i>Myxococcus fulvus</i>	( <i>S</i> )	B8ZV93
MxTAM	<i>Myxococcus sp.</i>	( <i>S</i> )	B8ZV94
RsTAL	<i>Rhodobacter sphaeroides</i>	( <i>S</i> )	Q3IWB0
SgTAM	<i>Streptomyces globisporus</i>	( <i>S</i> )	Q8GMG0

As mentioned in chapter 1.4.3, The (*S*)- $\alpha$ -phenylalanine and (*R*)- $\beta$ -phenylalanine are stabilized in the active site of TchPAM in different binding modes. While, when *trans*-

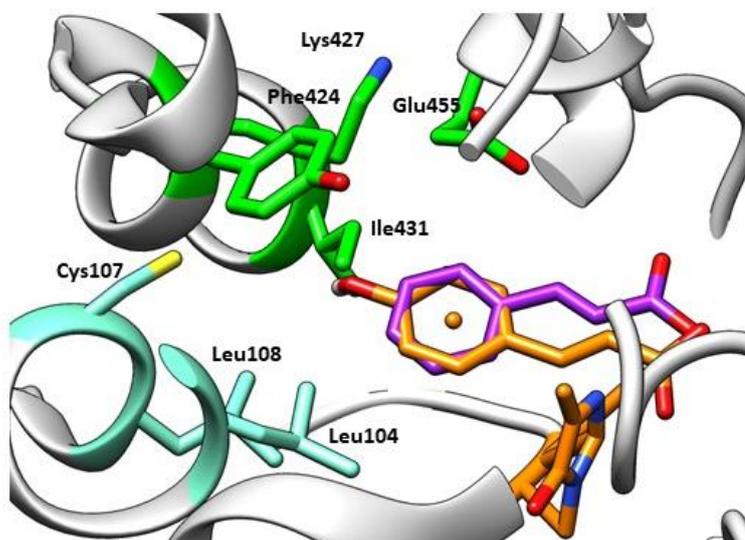
*p*-hydroxycinnamic enters the active site of the TchPAM, it seems to have only one available binding mode, involving the residues Asn231, Arg325, and Asn355, since the binding affinity of (*S*)- $\alpha$ -tyrosine is much lower than the phenylalanine and (*R*)- $\beta$ -tyrosine. Furthermore, the  $\beta$ -regioselectivity can be explained by the resonance of the *trans-p*-hydroxycinnamic acid with an electron-rich ring, which activates the  $\beta$ -position of the carbon atom on the backbone for nucleophilic attack (Figure 27) (Deruiter 2005). The *para*-substituted by an electron-donating group enhances the electron density of the benzene ring, promoting the formation of a hydrogen network by the carboxylate group of the substrate and the residues Asn231, Arg325, and Asn355 in the carboxylate binding site. Considering the *para*-substituent of the benzene ring, the *trans-p*-hydroxycinnamic acid may prefer to orient towards the residues Cys107, Tyr424, Lys427, Glu455, and Gln459 (Residue Group No.5), which contain the polar or charged side chain, instead of the aromatic binding pocket (Figure 29). The aromatic binding pocket has been reported by Wu et al. in 2012, involving the residues Phe86, Leu104, Cys107, Leu108, Leu179, and Gln459 (Wu, et al. 2012).



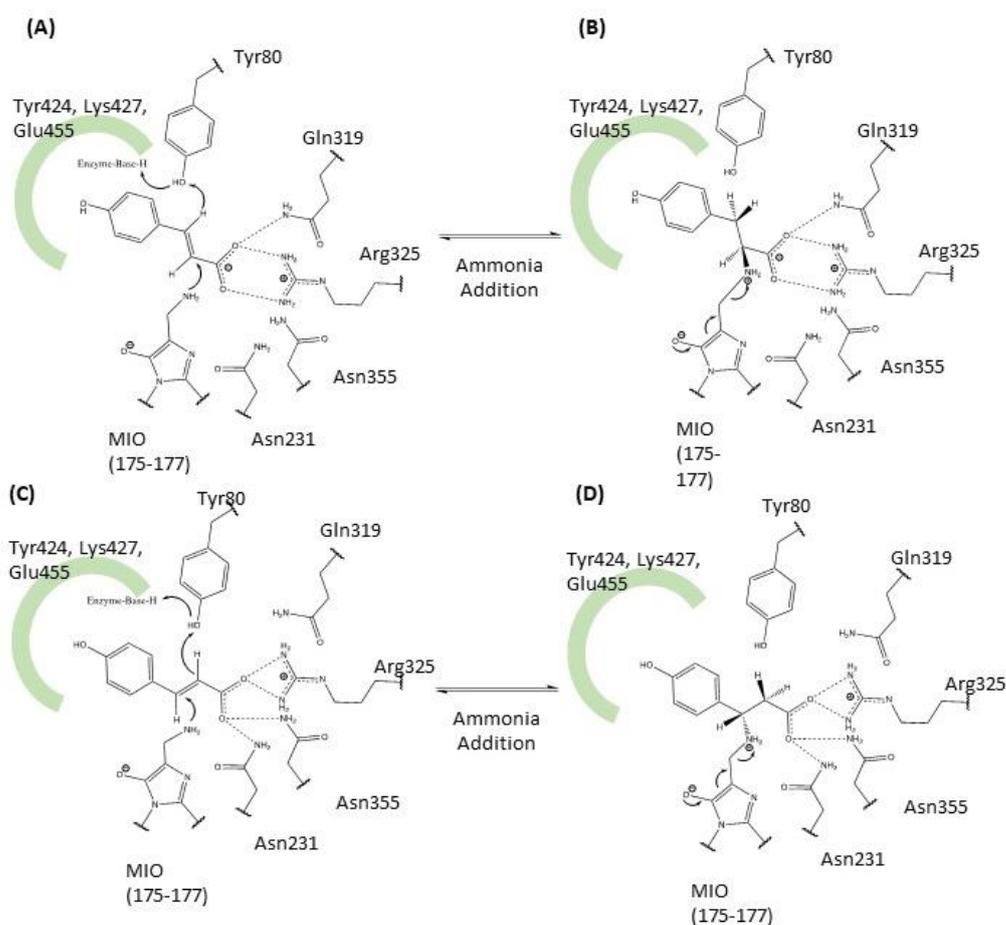
**Figure 27:** Resonance effect of the phenolic ring of the *trans-p*-hydroxycinnamic acid.

This binding mode is supported by the single-point mutagenesis results in this thesis (Figure 29). The variation at these positions, for example, Tyr424, to smaller residues is revealed to confer the acceptance of the *trans-p*-hydroxycinnamic acid. Among them, the mutants Tyr424Asn and Tyr424Cys do not show the reduction or switch the enantioselectivity, yielding the (*R*)- $\beta$ -tyrosine with high enantiomeric excess. These results strongly suggest the existence of the steric clash by the residue Tyr424 for the phenolic ring of the *trans-p*-hydroxycinnamic acid in the active center of TchPAM. It

also indicates that this phenolic ring orients the residues Tyr424, Lys427, and Glu455 (Figure 28).



**Figure 28:** The residues in the active center of TchPAM, which are placed close to the aromatic ring of the substrate. The residues in the aromatic binding pocket are colored in blue. The residues, including Tyr424, Lys427, Ile431, and Glu455, are colored in green. The natural substrate phenylalanine is colored in purple, while the color of the target substrate tyrosine and MIO moiety is orange.



**Figure 29:** The proposed mechanisms in the ammonia addition reaction by using the *trans-p*-hydroxycinnamic acid as substrate. The phenolic ring orients the binding pocket, including the residues Tyr424, Lys427, and Glu455, when the substrate enters in the active site of the TchPAM and binds in the (A) binding mode I, which promotes the amination on  $\alpha$ -position of the substrate, yielding the (*S*)- $\alpha$ -tyrosine (B). The substrate can be stabilized by the residues Asn231, Arg325, and Arg355 in the binding mode II (C) and then form the (*R*)- $\beta$ -tyrosine as the final product.

## 5.4 The designed TchPAM for (*S*)- and (*R*)- $\beta$ -tyrosine production:

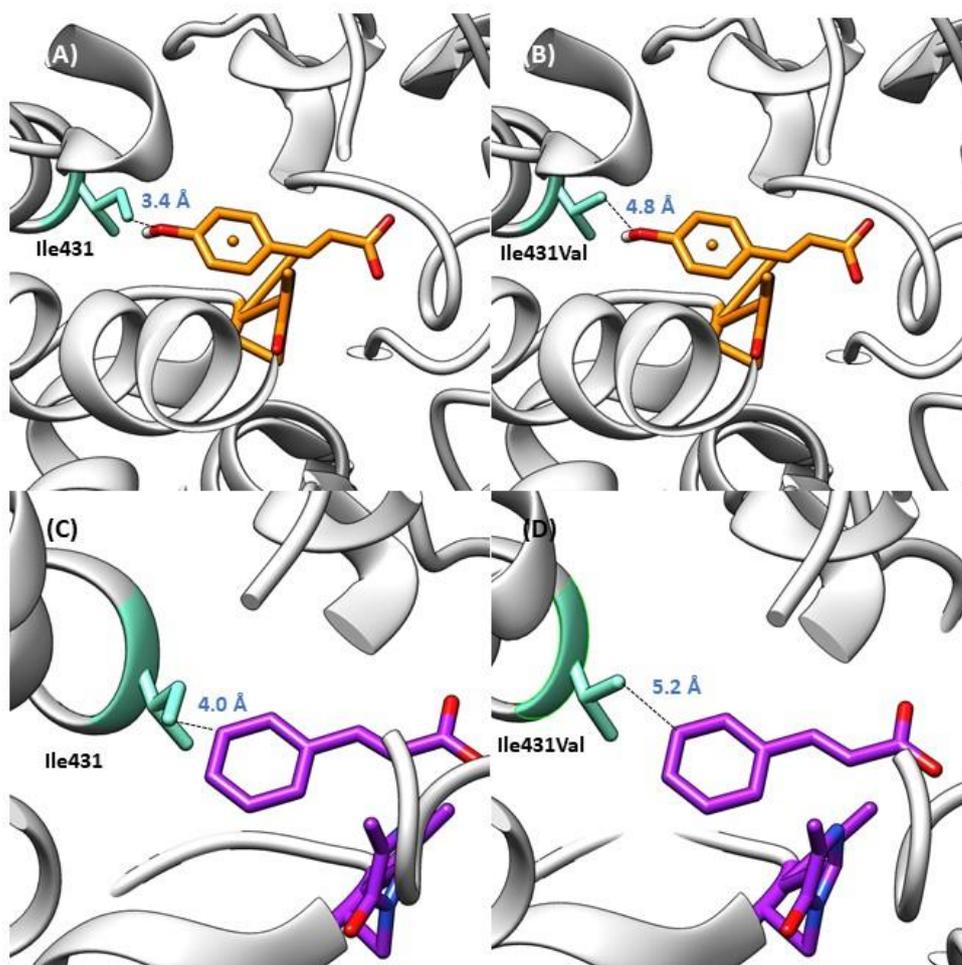
### Ile431

From a similar structural point of view, the mutation of the residue Ile431 to a shorter valine can also provide more space for the phenolic ring of the *trans-p*-hydroxycinnamic acid during the substrate binding, since the residue Ile431 is located in close contact with the residues in Residue Group No.5 (Figure 30). The mutant

---

Ile431Val can also catalyze the ammonia addition of *trans-p*-hydroxycinnamic acid to over 90%  $\beta$ -tyrosine, giving a conversion of 10% after 72 h, which backs up the hypothesis about the orientation of the phenolic ring. However, this mutant also gives an enantioselectivity switch from (*R*)- $\beta$ -tyrosine to (*S*)- $\beta$ -tyrosine. The residue Ile431 seems to strictly require a hydrophobic side chain for stabilization of the aromatic ring of the substrate since the mutants harboring variation to glutamic acid and asparagine completely abolish their activity with both phenylalanine and tyrosine, while the mutant Ile431Val retains partial ammonia-lyase activity. The catalytic efficiency and the binding affinity of phenylalanine and (*R*)- $\beta$ -tyrosine are identical in the mutant Ile431Val and the wild type, which suggests that the variation at this position with a smaller amino acid with similar characteristic will not change the binding position of these substrates. However, the  $K_M$  value of (*S*)- $\alpha$ -tyrosine in the mutant Ile431Val is much lower than the value in the wild type, indicating that the phenolic ring of (*S*)- $\alpha$ -tyrosine has a steric clash with the residue Ile431, which may be one reason of the poor acceptance of the (*S*)- $\alpha$ -tyrosine in the wild type. Furthermore, these results also indicate the difference in the binding position between the (*S*)- $\alpha$ -phenylalanine and tyrosine, which can be attributed to the interaction between the electron-rich ring of the tyrosine and the active site residues with the charged or partial charged side chains.

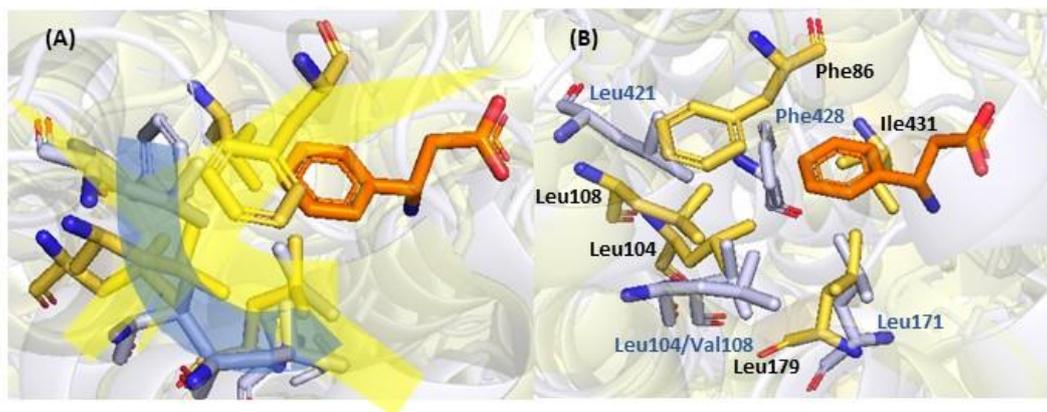
Even though the mutation of the residue Ile431 to valine decreases the London dispersion force towards the aromatic ring, this smaller hydrophobic residue can support the effective conformational change with the rotation around the  $C_1-C_\alpha$  bond, together with the other residues in the aromatic binding pockets (Brandl et al. 2001). However, the enantioselectivity switch from the (*R*)- $\beta$ -tyrosine to (*S*)- $\beta$ -tyrosine in the mutant Ile431Val indicates that the conformational alteration is unfavorable in the flexible aromatic binding pocket. Moreover, the replacement of the residue Ile431 by the other smaller hydrophobic amino acid has less influence on the substrate positioning of the benzene ring than the phenolic ring, since the mutant Ile431Val can not form any (*S*)- $\beta$ -phenylalanine as a product.



**Figure 30:** The distance and the proposed interaction between the residue Ile431 and the aromatic ring of the substrate tyrosine in (A) the wild type, the mutants (B) Ile431Val. The distance and the proposed interaction between the residue Ile431 and the aromatic ring of the substrate phenylalanine in (C) the wild type, (D) Ile431Val.

## 5.5 The proposed mechanisms of enantioselectivity in TchPAM

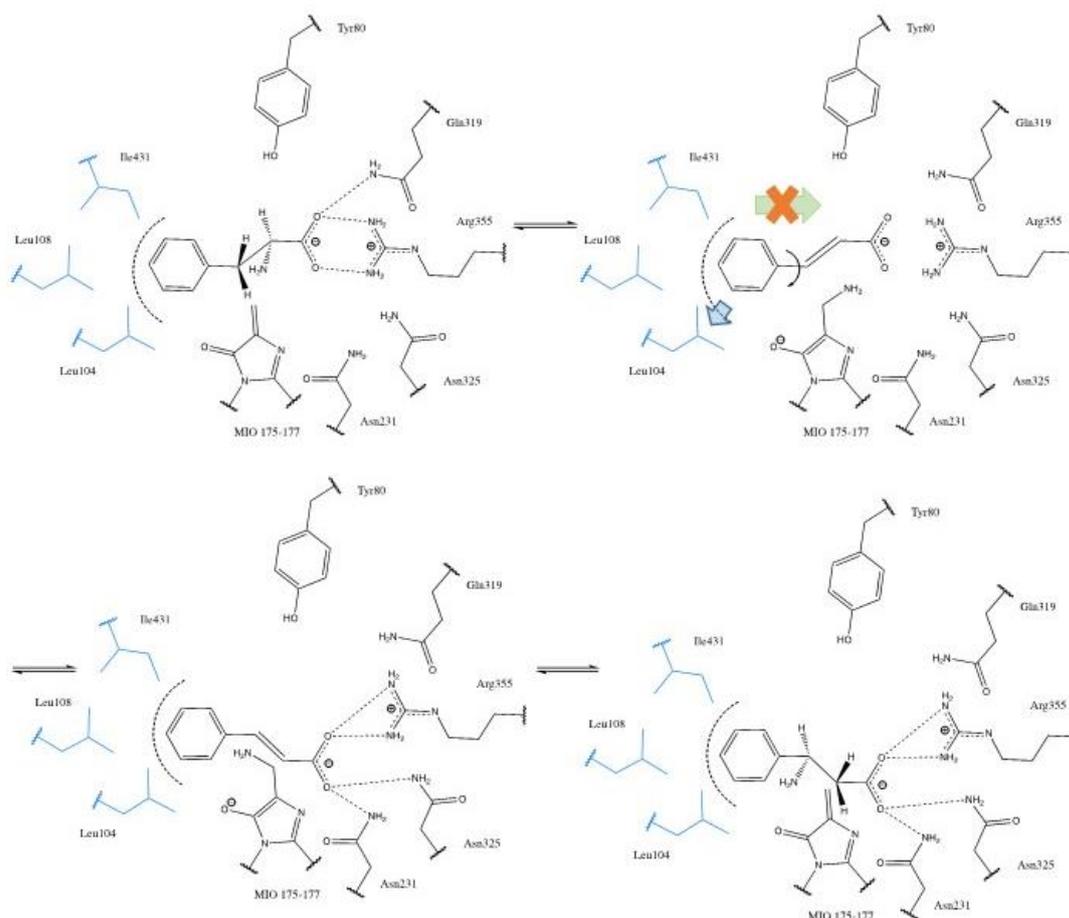
Compared to the aromatic binding pocket in the active site of the (*S*)- $\beta$ -selective PaPAM, the hydrophobic residues in TchPAM with the excellent enantioselectivity of (*R*)- $\beta$ -phenylalanine form an enclosing conformation and are located closer to the aromatic ring of the substrate, generating a stronger interaction between the substrate and the residues in the aromatic binding pocket. By stabilizing in this aromatic binding pocket, the movement of the backbone during the ammonia elimination and readdition is much harder than the more flexible aromatic binding pocket, like PaPAM (Figure 31).



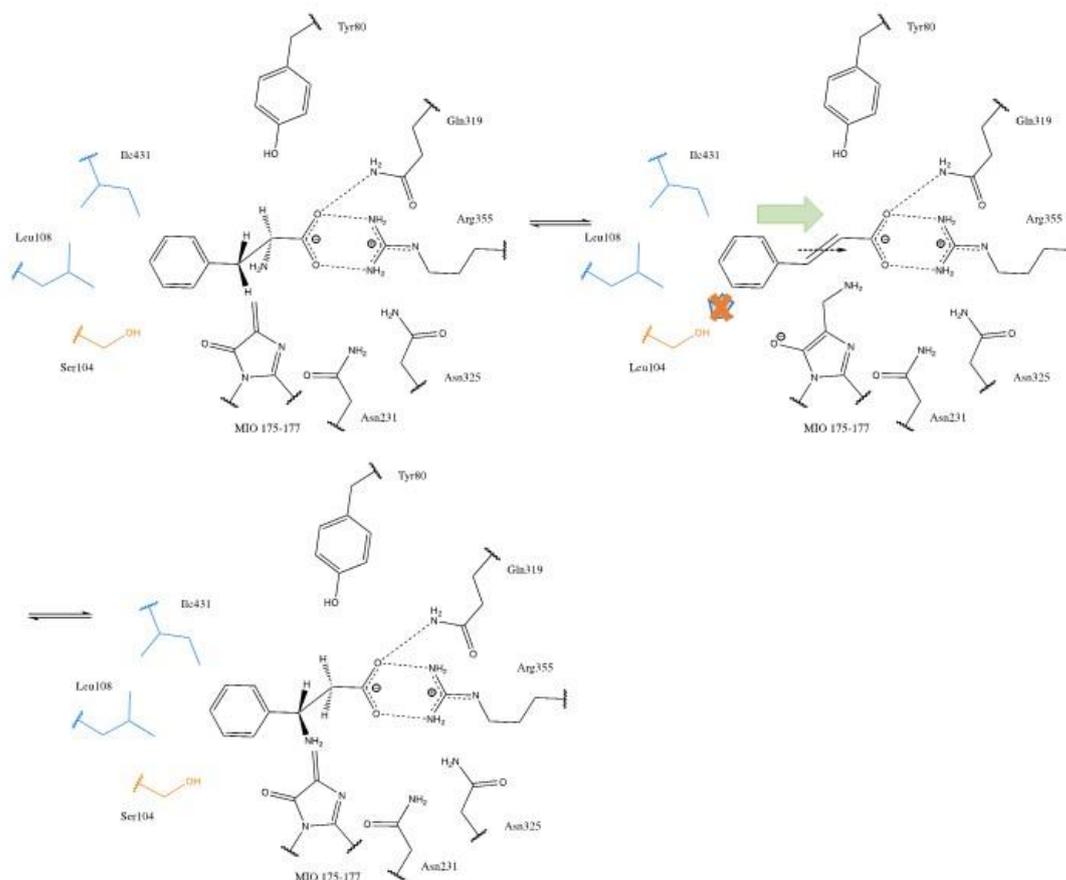
**Figure 31:** Overlay of the hydrophobic residues in the active site of (*R*)-selective TchPAM (yellow, PDB ID: 4C5R) and (*S*)-selective PaPAM (purple, PDB ID: 3UNV). The ligand is  $\beta$ -phenylalanine, which is colored in orange. **(A)**, the top view showed the enclosed aromatic binding pocket in TchPAM (yellow line) and the semi-enclosed aromatic binding pocket in PaPAM (blue line). **(B)**, the side view showed the involved hydrophobic residues (TchPAM: black, PaPAM: blue)

Combined with this comparison of the enzyme structures and the previous hypotheses from the other research group, a proposed isomerization mechanism has been established to explain the strict enantioselectivity of TchPAM. During the binding of the (*S*)- $\alpha$ -phenylalanine in the active site in TchPAM, the carboxylate of the substrate forms a bidentate salt bridge with the residue Arg325 (Shee et al. 2020). Meanwhile, compared to the (*S*)-selective PaPAM, the aromatic ring of the (*S*)- $\alpha$ -phenylalanine is stabilized by the hydrophobic residues Phe86, Leu104, Leu108, Leu179, and Ile431 with a stronger interaction, limiting the minimal shift of the substrate backbone. Therefore, it is difficult to readdition of the ammonia at the same stereoheterotopic face of the  $\beta$ -position on *trans*-cinnamic acid. For amination on the opposite face of the cinnamic acid intermediate, this intermediate has to rotate around the carbon bond and form a hydrogen network with the residue Asn231, Arg325, and Asn355. Considering the steric and torsional barrier, the single bond rotation around  $C_{\text{ipso}}-C_{\beta}$  may be favored in this case, compared to rotation around the  $C_1-C_{\alpha}$  bond or these two bonds together (Heberling et al. 2013). Whereas, the aromatic ring of the substrate can move quite flexibly in the aromatic binding pocket with a semi-enclosed conformation of the

PaPAM. Therefore, the difference of the residues in the aromatic binding pockets between the PaPAM and TchPAM plays an important role in the enantioselectivity (Figure 32 and Figure 33).



**Figure 32:** Proposed isomerization mechanism of (*R*)-selective TchPAM. When (*S*)- $\alpha$ -phenylalanine enters the active site of TchPAM, its carboxylate group is stabilized by residues Arg325 and Gln319. In the wild-type enzyme, the benzene ring is fixed in the aromatic binding pocket, in particular, with the residue Leu104, during the ammonia elimination to form the cinnamic intermediate. A single rotation around the  $C_{\text{ipso}}-C_{\beta}$  bond without breaking the interaction with residue Leu104 brings the  $\beta$ -position on unsaturated acid to a suitable position for readdition reaction by MIO-amine adduct, giving (*R*)- $\beta$ -phenylalanine as the product. Meanwhile, the carboxylate group from the intermediate is stabilized by residues Arg325, Asn231, and Asn355.



**Figure 33:** In the mutant Leu104Ser, the benzene ring can move flexibly in the aromatic binding pocket, due to the lack of hydrophobic residue at position 104, during the ammonia elimination to form the cinnamic intermediate. Thus, the minimal shift of the intermediate backbone led to an immediate readdition of ammonia on  $\beta$ -position, as well as the formation of (*S*)- $\beta$ -phenylalanine.

## 5.6 The proposed function of the surrounding residues

### 5.6.1 The highly conserved residues in the active site of TchPAM

#### Leu179: Not strictly required for MIO formation

The earlier structure study has been revealed that the neighboring residues of the tripeptide, which could be automatically folded in MIO moiety, might support the cyclization step of the loop with mechanical compression. The formation of the MIO moiety is blocked in the mutant Asp145Ala in PpHAL (the residue Asp178 in TchPAM) (Baedeker and Schulz 2002). Considering the high conservation of the residue Leu179

---

in TchPAM, it was also supposed to be an essential position for the generation of MIO moiety by providing the mechanical compression to the tripeptide loop or the residue Asp178. However, the mutants harboring variation to the amino acids with a hydroxyl group still could catalyze the natural substrate, albeit with much lower activity. These results suggest that the residue Leu179 does not influence the formation of MIO moiety in the mutants Leu179Thr and Leu179Ser. The residue Leu179 seemed to have a hydrophobic interaction with the aromatic ring of the substrate, affecting the substrate orientation and selectivity. However, it could be excluded the probability of varying the position of the residue Asp178 and prevent the cyclization of MIO moiety by a larger side chain at this position.

#### **Gln459: Key position for positioning of the amino group from substrates**

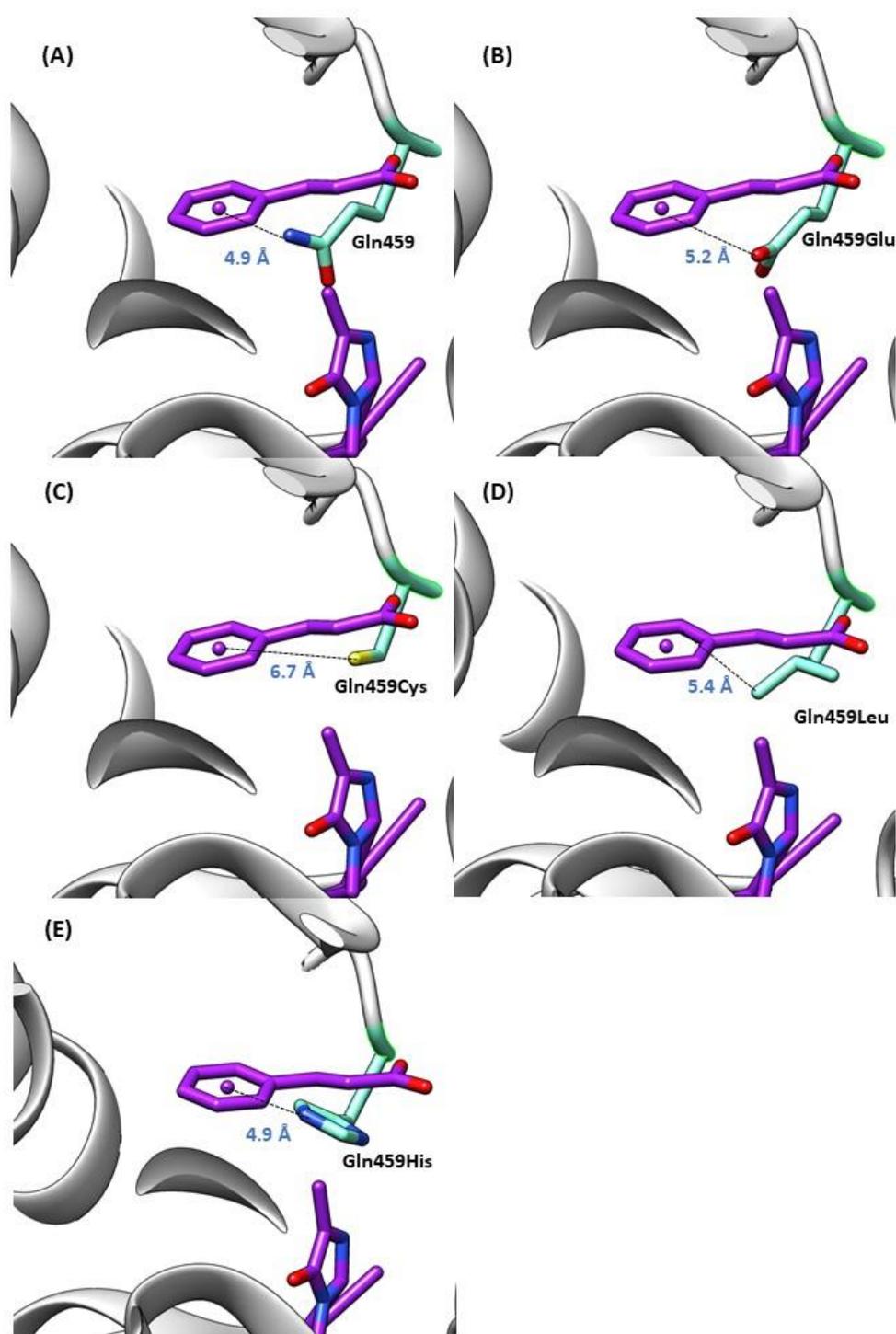
The residue Gln459 is replaced by glutamic acid in HAL (position Glu414), which has been reported to assist the dehydration step during the formation of MIO moiety (Figure 34) (Sánchez-Murcia et al. 2016). The hybrid quantum mechanics/molecular mechanics simulations on PpHAL pointed out the existence of a hydrogen network, which is formed by the carboxylate group of the residue Glu414 (Gln459 in TchPAM) and the hydroxyl groups of the residue Ser143 (Ser176 in TchPAM) and the highly conserved residue Tyr280 (Tyr322 in TchPAM) (Sánchez-Murcia et al. 2016). In contrast to the residue Tyr 322, the residue Gln459 was not strictly required for the post-translational modification in TchPAM, since the variants Arg325Lys/Gln459Glu remained the 30% of the activity with  $\beta$ -phenylalanine in the earlier mutagenesis study (Wu, et al. 2012). In this case, the negatively charged glutamic acid, which is located below the substrate, generates a cation- $\pi$  interaction with the aryl ring center, driving a deviation of the substrate-binding position (Philip et al. 2011). Whereas, in the wild type the formamide group of the residue Gln459 orients the edge of the aromatic ring, pushing it away from the Residue Group No.5. (Petsko 1986). Instead of this interaction, the residue Gln459Leu may provide an attractive force toward the aromatic ring of the substrate due to its hydrophobic side chain at this position, leading to a notable reduction of

---

activity. While the loss of the partial activity in the mutant Gln459Cys is caused by the variation of the residue to cysteine with a small side chain. It may not affect the aromatic ring of the substrate because of the huge distance between the side chain and substrate. In contrast with the mutant Gln459Cys, the mutant Gln459His completely loses its activity due to the steric and electronic hindrance by the large side chain of histidine. Besides, the side chain may also affect the MIO prosthetic group formation in the mutant Gln459His, considering the short distance between the histidine and the MIO moiety.

The tyrosine with the electron-rich aromatic ring has a stronger interaction with the side chains of the residue Gln459, compared to the benzene ring of the natural substrate. Therefore, the mutant Gln459Leu completely abolishes the activity with tyrosine, which may be attributed to a larger deviation of the phenolic ring, during the binding of the substrate in the active site of TchPAM. Additional supporting evidence comes from the mutant Gln459Cys, the only one that retained specific activity with  $\alpha$ -tyrosine of the mutants at position Gln459. In mutant Gln459Cys, the thiol group is far from the binding position of the aromatic ring of the substrate and has less influence than the other mutants (McGaughey et al. 1998)(Janiak 2000).

Interestingly, the early mutagenesis work on TchPAM by Wu et al. showed that the mutation at this position may change the substrate selectivity (Wu, et al. 2012). They have reported a variant Arg325Lys/Gln459Glu with a switch of substrate selectivity (Wu, et al. 2012). It is consistent with our results in this work: the mutant Gln459Leu shows preference toward the (*R*)- $\beta$ -phenylalanine, while the mutant Gln459Cys only accepts the  $\alpha$ -aromatic amino acids as substrate.



**Figure 34:** The distance and the proposed interaction between the residue Gln459 and the aromatic ring of the substrate in (A) the wild type, the mutants (B) Gln459Glu, (C) Gln459Leu, (D) Gln459Cys, and (E) Gln459His.

#### **Asn458: The other carboxylate binding residues**

The amide group is strictly required at position Asn458, even though the HALs exhibit

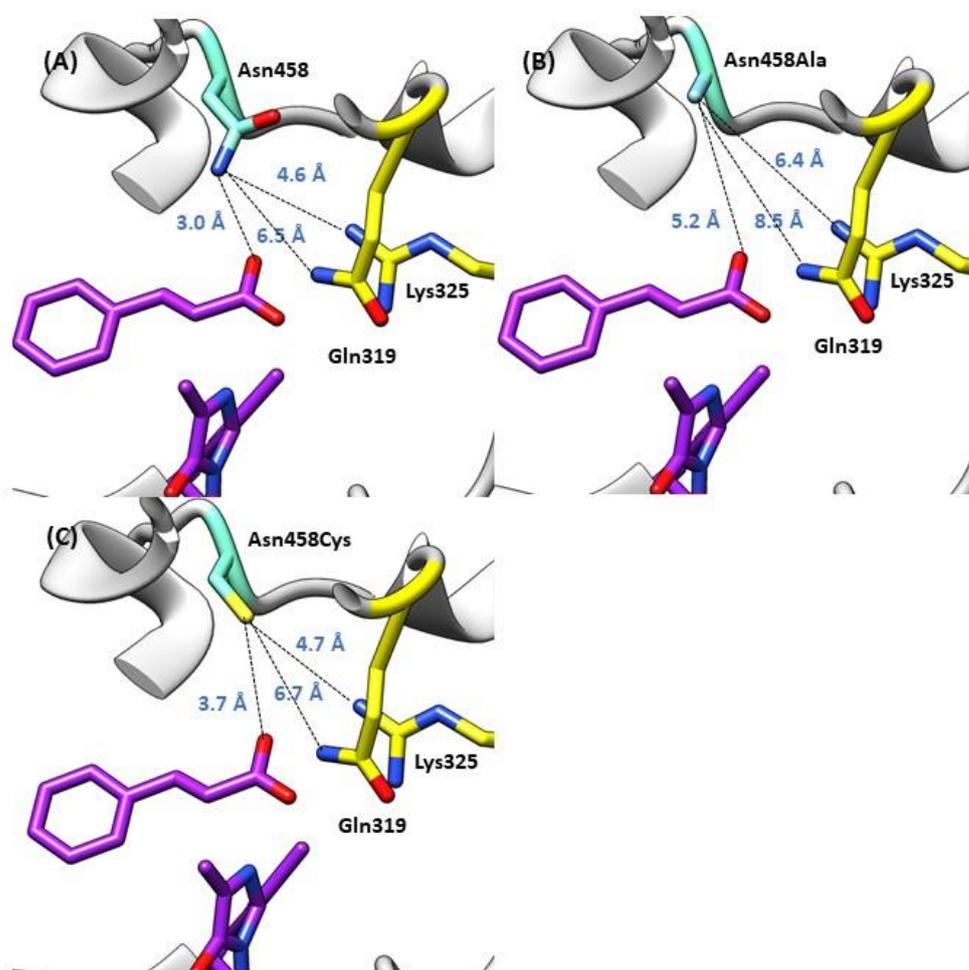
---

the glutamine with one methylene longer side chain instead of the asparagine in the TchPAM. According to the conservation analysis, more than 95% homologues have the residue asparagine at this position (in capital 4.3.2.). However, the members of the MIO-dependent enzyme family can also allow the arginine, threonine, lysine, and valine. Considering the location of the residue Asn458, the formamide group not only reacts with the carboxylate group of the substrate but also affects the residues in the carboxyl binding pocket, especially, the residue Arg325 (Figure 35).

In the mutant Asn458Cys, the thiol group may form a hydrogen bond towards the carboxylate group of the substrate instead of the salt bridge. However, the distance between the thiol group of the mutant Asn458Cys and the substrate is too large to generate such an interaction, leading to a decreasing activity with phenylalanine in the mutant Asn458Cys. Whereas, the mutant Asn458Ala with the breaking of the salt bridge retains only 6% activity with phenylalanine, in both  $\alpha$ -,  $\beta$ -regioisomers, supporting the importance of the interaction between the residue Asn458 and the carboxylate group of the substrate. The early mutagenesis study from the other group has reported two mutants Asn458Leu and Asn458Phe without any reduction of the phenylalanine ammonia-lyase activity and suggested the improvement of the preference of the  $\beta$ -amino acids due to the lack of the interaction between the amide group and the carboxylate group of the substrate (Zhu et al. 2018). Therefore, the steric hindrance but not the electronic effect at position 458 may be the reason for varying the position of the substrates.

The other reason for the decreasing of the activity in these mutants involves the interaction between the residues Asn458 and the residue Arg325. The latter residue is considered as the key position for binding of the carboxylate group of the substrate, affecting the activity and substrate selectivity (Wanninayake 2013). In the mutant Asn458Cys, the thiol group can form a salt bridge with the positively charged guanidino group of the residue Arg325, driving both side chains in close contact with each other. This deviation may break the bidentate salt bridge in the  $\alpha$ -phenylalanine binding mode,

which is formed by the carboxylate group of the substrate and residues in the carboxylate binding pocket, enhancing the  $\beta$ -selectivity (Wu, et al. 2012). It is consistent with the results of the measurement of the specific activity with  $\alpha$ -,  $\beta$ -phenylalanine in the mutant Asn458Cys. Whereas, the mutant Asn458Ala may decrease the steric restraint to the side chain of the residue Arg325, which may also change the position of the residue Arg325.



**Figure 35:** The distance and the proposed interaction between the residue Asn458, the carboxylate group of the substrate, and the amino group of the residues Gln319 and Arg325 in (A) the wild type, the mutants (B) Asn458Cys, (C) Asn458Ala.

Regarding tyrosine, the electron-rich ring leads to a stronger interaction with the residues, which is located near the aromatic ring. It may be a reason for the complete

---

loss of the enzyme activity with tyrosine in the mutant Asn458Ala and Asn458Cys (McGaughey et al. 1998)(Janiak 2000).

### 5.6.2 The residues in the aromatic binding pocket

#### **F86:**

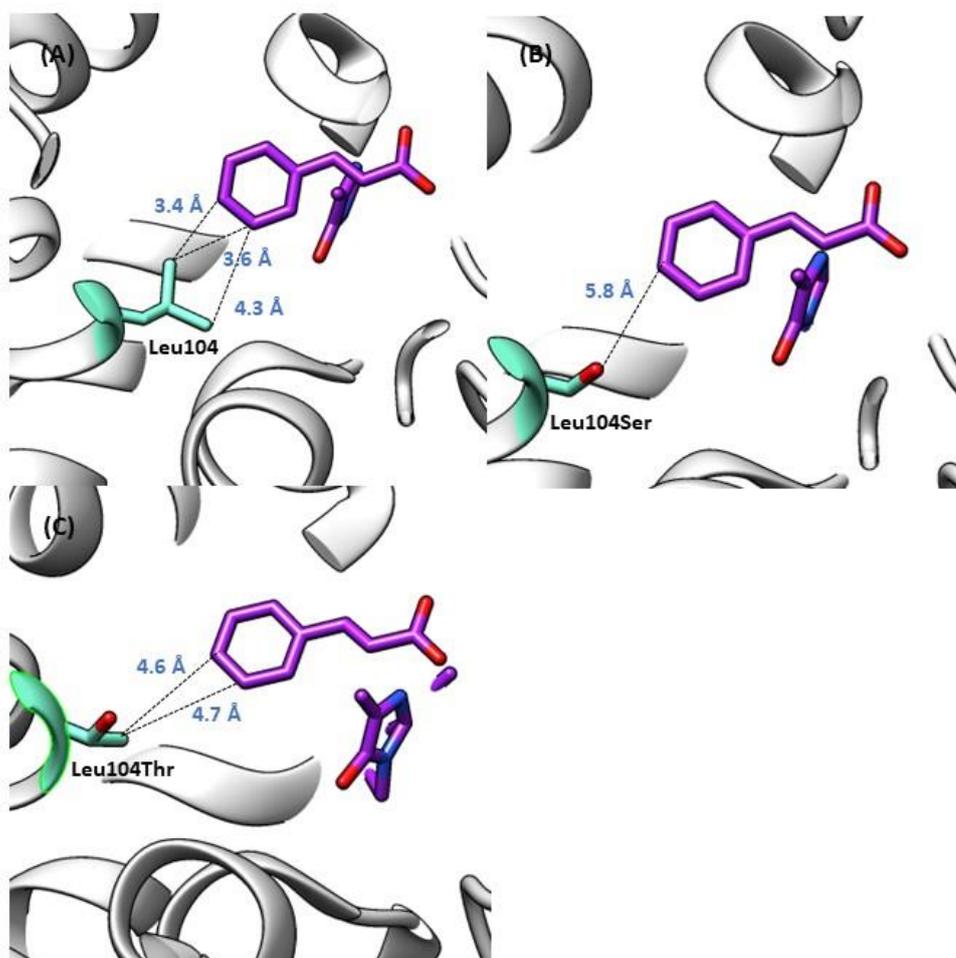
Even though the residue Phe86 is highly conserved in the MIO-dependent enzymes, which accept the phenylalanine as the substrate. It can be spontaneously mutated to tyrosine in the SgTAM, which is most investigated because of the corresponding biosynthesis pathway of the enediyne antitumor natural metabolite C-1027 (Christianson et al. 2007). Therefore, it seems to require an aromatic side chain at position 86 to stabilize the aromatic ring of the substrate through the  $\pi$ - $\pi$  interaction. The mutation of the residue Phe86 to tyrosine with an electron-rich ring decreases the distance between the aromatic ring of the substrate and the residues from the Residue Group No.5., which influences the orientation position of the phenolic ring of the *trans-p*-hydroxycinnamic acid (in capital 5.2). However, no activity has been detected in the mutant Phe86Tyr by using both phenylalanine and tyrosine as substrate, which may be attributed to the steric clash by the residues Ile431 and Tyr424 at this position. Furthermore, the reduction of the electrostatic repulsion by replacing the residue Phe86 with tyrosine may also affect the position and the orientation of the neighboring aromatic residue Tyr424, leading to a larger steric hindrance at this position.

#### **L104: The other key position for enantioselectivity in TchPAM**

Compared to the residue Ile431, the residue Leu104 orients towards the *meta*-position of the aromatic ring, but on the other side, which is supported by the previous structure studies (Wybenga et al. 2014)(Nagy et al. 2019). The additional evidence is the enhanced catalytic efficiency with higher binding affinity in the mutant Leu104Ala, compared to the wild type, when using the 3-methyl- $\alpha$ -phenylalanine as the substrate (Feng et al. 2011). It is consistent with the mutagenesis study on PcPAL, which has

---

suggested the orientation of the residue Leu134 (Leu104 in TchPAM) towards the *meta*- or *ortho*-position of the aromatic ring of the substrates (Nagy et al. 2019). The hydrophobic side chain at this position may have a great influence on the substrate binding, which is proved by the mutants Leu104Ser and Leu104Thr with the remarkable decreased binding affinity of all the tested aromatic amino acids. However, replacing the hydrophobic residue with the amino acids with a polar side chain has less influence on the substrate binding of the (*R*)- $\beta$ -tyrosine, compared to the natural substrate, which is additional evidence for the orientation of the phenolic ring of the substrate towards the residues Tyr424, Lys427, Ile431, and Glu455. In contrast with the phenolic ring, which may stabilize by residue Ile431, the benzene ring may bind in close contact with the residue Leu104. While the improvement of the  $K_M$  value of the (*R*)- $\beta$ -tyrosine may be attributed to the formation of the hydrogen bond between the phenolic ring of the substrate and the polar side chain at position 104. Moreover, the increased  $k_{cat}$  value in mutant Leu104Ser indicates that the more flexible aromatic binding pocket may promote the ammonia elimination reaction (Figure 36).



**Figure 36:** The distance and the proposed interaction between the residue Leu104, the aromatic group of the substrate phenylalanine in (A) the wild type, the mutants (B) Leu104Ser, (C) Leu104Thr.

According to the results of the ammonia addition reaction, the mutant Leu104Ser can produce the (*S*)- $\beta$ -phenylalanine, albeit with 7% conversion after 24 h. The variation of the residue Leu104 with serine leads to a reduction of the interaction between the residues in the aromatic binding pocket and the aryl ring of the *trans*-cinnamic acid, mitigating the energy requirement for the rotation around the C<sub>1</sub>-C <sub>$\alpha$</sub>  bond (Brandl et al. 2001). Moreover, by changing the enclosed aromatic binding pocket to a relatively flexible one in the mutant Leu104Ser, it is possible for reamination on the  $\beta$ -position of the *trans*-cinnamic acid without any conformational change, leading to produce the (*S*)- $\beta$ -phenylalanine.

---

**L108:**

The previous site-directed mutagenesis studies have shown that the residue Leu108 has a great potential for expanding the substrate scope (Watts et al. 2006). One remarkable example of the substrate selectivity switch is the mutation of two residues Leu138 (Leu108 in TchPAM) and Gln488 in the active site of PcPAL, which successfully changed the substrate selectivity from phenylalanine to histidine (Röther et al. 2002). In contrast with the residue Leu104 and Leu431, the residue Leu108 with hydrophobic side chain has less influence on the binding position of the substrate due to the larger distance between the residue and bound substrate, which is confirmed by the X-ray crystal structure of TchPAM (Feng et al. 2011)(Wybenga et al. 2014). The variation of the residue Leu108 to glutamine leads to a partial loss of the activity with the phenylalanine and tyrosine. When using the phenylalanine, the amide group of the residue Leu108Gln orients the positively charged ring edge of the benzene ring of the substrate, pushing the substrate far from the original binding position. Strangely, the recent mutagenesis research on TchPAM from the other research group showed a mutant Leu108Glu with an identical specific activity but increased  $\beta$ -selectivity, compared to the wild type (Zhu et al. 2018). In this case, the carboxyl group of the residue Leu108Glu may generate an attractive force towards the benzene ring of the substrate. The other mutant Leu108Ser displays higher specific activity in the ammonia elimination reaction of the phenylalanine and tyrosine than the mutant Leu108Gln, indicating that not only the functional group but also the chain length of the residue Leu108 may affect the enzyme activity of TchPAM. Compared to the wild type, the partial reduction of the activity can be attributed to the lack of the London dispersion force, which is provided by the hydrophobic residue Leu108. Moreover, the hydroxyl group of the residue Leu108Ser may interact with the neighboring residue Cys107, which is also recognized as a substrate selectivity residue in MIO-dependent enzymes. The early mutagenesis study has revealed the mutant His89Phe (Cys107 in TchPAM) in RsTAL, which confers phenylalanine acceptance (Watts et al. 2006). However, the

---

designed residues Cys107 are not found in the filtered mutant library by the computational enzyme design.

---

## 6 CONCLUSION AND OUTLOOK

The aim of this project was the construction of a strategy to produce valuable  $\beta$ -tyrosine with high enantiopurity from monolignol *trans-p*-hydroxycinnamic acid. The chemoenzymatic cascade based on a modified hydantoinase process was not achieved, since the produced yield of OH-PheDU was too low in the first chemical synthesis step. In the subsequent attempt to increase the isolated yield, the application of different methods and optimization reaction parameters were performed. However, the chemical synthesis of OH-PheDU from *trans-p*-hydroxycinnamic acid remains challenging.

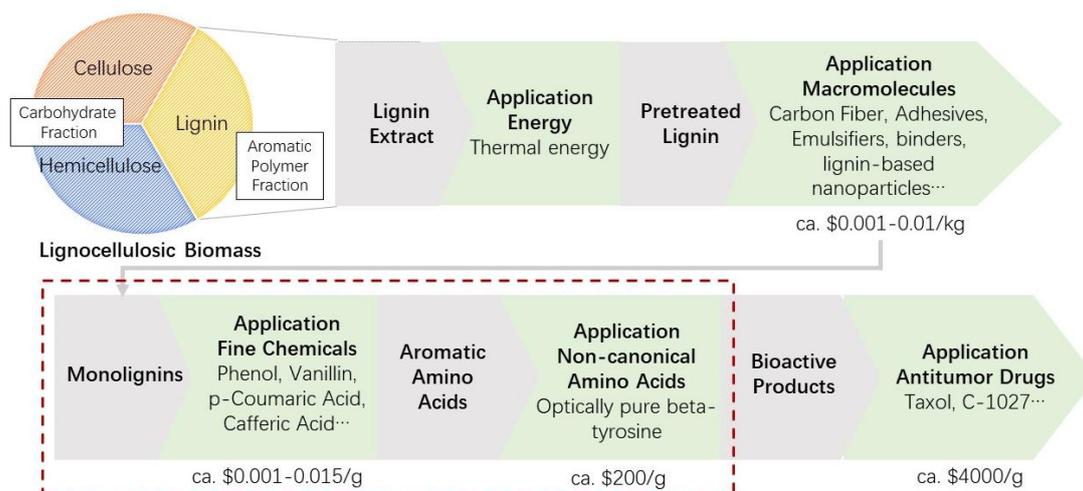
Phenylalanine aminomutase belongs to a superfamily with a similar structure, a characteristic highly electrophilic MIO prosthetic group in the active site, and an identical catalytic mechanism. As the most investigated representative of this superfamily, TchPAM has gained enormous attention. To improve the enantiopurity of produced  $\beta$ -phenylalanine or extend the substrate scope, several enzyme-engineering studies of TchPAM have been reported based on the random mutation (Wybenga et al. 2014)(Wu, et al. 2012)(Zhu et al. 2018). In contrast to PAM and PAL, the  $\beta$ -tyrosine-producing enzyme has not been extensively studied. Even though enantiopure  $\beta$ -tyrosine can also be utilized as a building block for pharmaceutical ingredients and drug candidates, for example, kedarcidin and C-1027 (Carretero-Molina et al. 2020)(Chang et al. 2016).

Considering the low enantioselectivity of bacterial TAM and TAL, it is difficult to produce (*R*)- $\beta$ -tyrosine with high enantiopurity by the natural enzyme (Rachid et al. 2007)(Krug and Müller 2009). Although the (*R*)-selective OsTAM from *Oryza sativa* can produce (*R*)- $\beta$ -tyrosine from (*S*)- $\alpha$ -tyrosine, this TAM is poorly accepted *trans-p*-hydroxycinnamic acid as substrate (Walter et al. 2016)(Yan et al. 2015). Therefore, in order to achieve the goal of the project, TchPAM was chosen as the template for computational enzyme design by PyRosetta to confer  $\beta$ -tyrosine acceptance. The highly conserved amino acid positions within TchPAM were determined through multiple

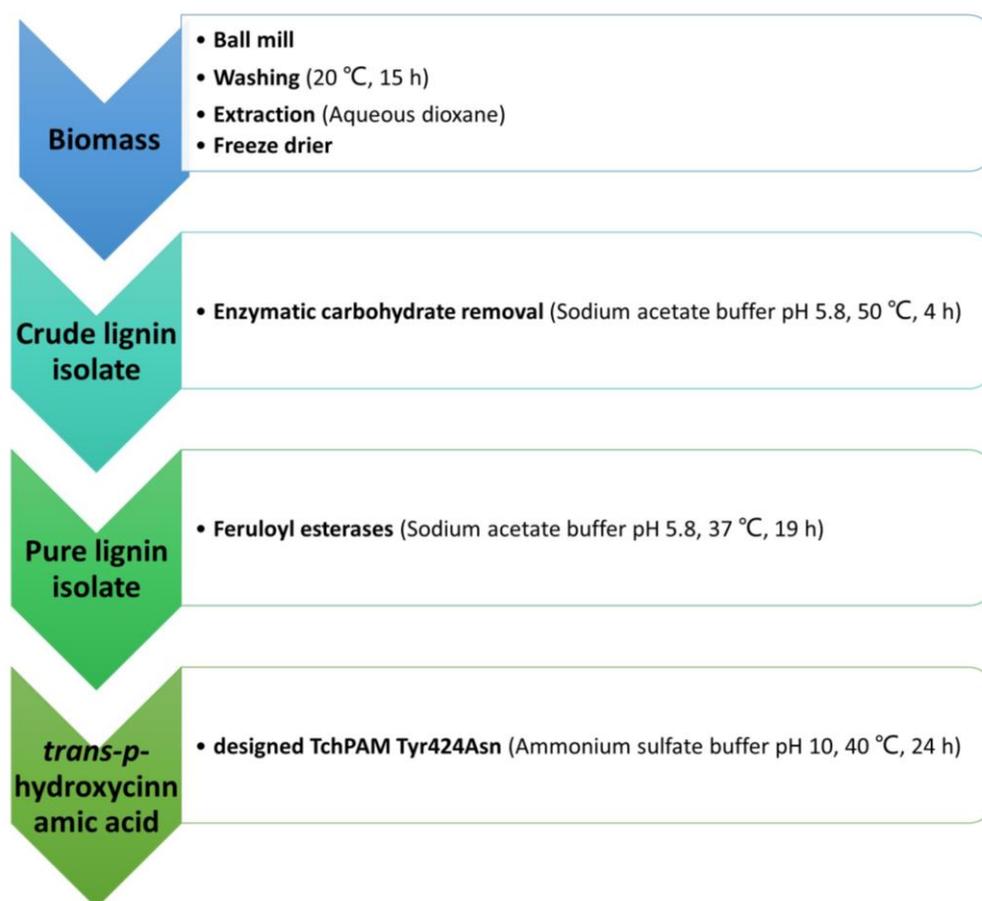
---

amino acid sequence analysis with the other members in the MIO-dependent enzyme family. The position Phe86, Leu104, Leu108, Leu179, Tyr424, Lys427, Ile431, and Glu455 were selected to mutate in residues with polar or electrically charged side chain since they are placed in close contact with the phenolic side of the ligand  $\beta$ -tyrosine. Meanwhile, the residue Asn458 and Gln459 were mutated without any limitation in computational design, regarding their potential of stable network formation with the substrate. According to the score function in PyRosetta and structure comparison by PyMOL, the designed enzymes were filtered. A mutant library with multiple position mutations was established for subsequent experimental screening. The characterization screening of twenty-eight designed enzymes with single point mutation afforded two variants Tyr424Asn and Try424Cys, which could catalyze the ammonia addition reaction of *trans-p*-hydroxycinnamic acid and yield enantiopure (*R*)- $\beta$ -tyrosine ( $ee_R > 99\%$ ). In addition, the single point mutation Ile431Val resulted in an enantioselectivity switch, giving (*S*)- $\beta$ -tyrosine with enantiomeric excess in 29%. Similarly, the variant Leu104Ser could also catalyze the conversion of *trans*-cinnamic acid into (*S*)- $\beta$ -phenylalanine. However, the conversion rate of the (*R*)- $\beta$ -tyrosine production was only 2% after 24 h by employing the designed enzyme Tyr424Asn. While the conversion rate was 7% after 72 h reaction by using the mutant Ile431Val to produce  $\beta$ -tyrosine from *trans-p*-hydroxycinnamic acid. Therefore, the ammonia addition reaction should be optimized to improve the yield of produced (*R*)- $\beta$ -tyrosine.

The enzymatic conversion with TchPAM variants Tyr424Asn and Tyr424Cys can connect with the lignin depolymerization upstream and the drug development downstream, developing efficient and straightforward pathways of lignin valorization (Figure 37). This green strategy can mitigate the problems of fossil resources and may reduce the cost of antitumor drugs through the generation of (*R*)- $\beta$ -tyrosine-involved agents from unsterilized lignin.



**Figure 37:** Flow chart of a strategy for lignin valorization. The main goal of this doctoral research is framed in a red dotted line.



**Figure 38:** A proposed pathway for multi-enzymatic synthesis of *(R)*- $\beta$ -tyrosine with high enantiopurity by starting from the pretreated lignin.

---

As an attractive example, combining the obtainment of *trans-p*-hydroxycinnamic acid from lignin by feruloyl esterases, a multi-enzymatic process with designed TchPAM Tyr424Asn or Tyr424Cys could be established. Recently, Underline et al. have reported that the feruloyl esterases from *Chrysosporium lucknowense* are able to release *p*-hydroxycinnamic acid from lignin isolates (Underlin et al. 2020). The biomass, such as corn stover and wheat straw, should be ball milled and washed with water at the beginning. To prepare lignin isolates, the wet residue was extracted by 80 % aqueous dioxane at room temperature under a nitrogen atmosphere. Then, the crude lignin isolates should be purified by enzymatic carbohydrate removal before utilization as the substrate for reaction with feruloyl esterases (Van Erven et al. 2017). After incubation in sodium acetate buffer (pH 5.8) at 37 °C for 19 h, *p*-hydroxycinnamic acid was produced (Underlin et al. 2020). The mutants Tyr424Asn or Tyr424Cys could be employed in the subsequent reaction in presence of ammonia under alkaline conditions. The ammonia addition reaction should be carried out at 40 °C for the other 24 h (Figure 38)(Peng et al. 2021).

Apart from the multi-enzymatic process, the mutants Tyr424Asn and Tyr424Cys from TchPAM could also be added in fermented culture for improvement of the value of final products in the lignin valorization. For example, the strain of *Pseudomonas putida* without the hydroxycinnamoyl-CoA hydratase-lyase gene has been found for loss of the ability to metabolize the ferulic acid and *p*-hydroxycinnamic acid, leading to an accumulation of these two aromatic compounds during the fermentation of sugarcane bagasse or miscanthus (Williamson et al. 2020). Through the development of fermentation with this strain of *P. putida* and the designed TchPAM, (*R*)- $\beta$ -tyrosine might be directly generated from the corresponding lignocellulosic biomass. The primary challenges of this trending topic are the optimization of the fermentation parameters for the involved microorganism and enzymes, as well as the development of a suitable environment for ammonia addition.

---

## 7 REFERENCES

- Allen FH. 2002. The Cambridge Structural Database: a quarter of a million crystal structures and rising. *Acta Crystallogr. B.* 58:380–388.
- Allwood EG, Davies DR, Gerrish C, Ellis BE, Bolwell GP. 1999. Phosphorylation of phenylalanine ammonia-lyase: Evidence for a novel protein kinase and identification of the phosphorylated residue. *FEBS Lett.* 457:47–52.
- AlQuraishi M. 2019. AlphaFold at CASP13. *Bioinformatics* 35:4862–4865.
- Amirav-Drory O, Debbi Y, Nevo R, others. 2015. Genome compiler.
- Asano Y, Kira I, Yokozeki K. 2005. Alteration of substrate specificity of aspartase by directed evolution. *Biomol. Eng.* 22:95–101.
- Attanayake G, Walter T, Walker KD. 2018. Understanding Which Residues of the Active Site and Loop Structure of a Tyrosine Aminomutase Define Its Mutase and Lyase Activities. *Biochemistry* 57:3503–3514.
- Auerbach VH, DiGeorge AM, Baldrige RC, Tourtellotte CD, Brigham MP. 1962. Histidinemia: a deficiency in histidase resulting in the urinary excretion of histidine and of imidazolepyruvic acid. *J. Pediatr.* 60:487–497.
- Azadi P, Inderwildi OR, Farnood R, King DA. 2013. Liquid fuels, hydrogen and chemicals from lignin: A critical review. *Renew. Sustain. Energy Rev.* 21:506–523.
- Azaryan L, Avetisyan S, Buyukyan N, Chachoyan A, Garibdzhanyan B. 2010. ChemInform Abstract: Synthesis and Antitumor Activity of 6-p-Alkoxyphenyl Derivatives of Hexahydropyrimidine-2,4-diones. *Cheminform* 28.
- Baedeker M, Schulz GE. 2002. Autocatalytic peptide cyclization during chain folding of histidine ammonia-lyase. *Structure* 10:61–67.
- Bagal UR, Leebens-Mack JH, Lorenz WW, Dean JFD. 2012. The phenylalanine ammonia lyase (PAL) gene family shows a gymnosperm-specific lineage. *BMC Genomics* 13 Suppl 3:1–9.

- 
- Bai X-C, McMullan G, Scheres SHW. 2015. How cryo-EM is revolutionizing structural biology. *Trends Biochem. Sci.* 40:49–57.
- Bairoch A, Boeckmann B, Ferro S, Gasteiger E. 2004. Swiss-Prot: juggling between evolution and stability. *Brief. Bioinform.* 5:39–55.
- Barros J, Dixon RA. 2020. Plant Phenylalanine / Tyrosine Ammonia-lyases Trends in Plant Science. *Trends Plant Sci.* 25:66–79.
- Barros J, Escamilla-Trevino L, Song L, Rao X, Serrani-Yarce JC, Palacios MD, Engle N, Choudhury FK, Tschaplinski TJ, Venables BJ, et al. 2019. 4-Coumarate 3-hydroxylase in the lignin biosynthesis pathway is a cytosolic ascorbate peroxidase. *Nat. Commun.* 10:1–11.
- Barros J, Serrani-Yarce JC, Chen F, Baxter D, Venables BJ, Dixon RA. 2016. Role of bifunctional ammonia-lyase in grass cell wall biosynthesis. *Nat. Plants* 2:1–9.
- Bender RA. 2012. Regulation of the Histidine Utilization (Hut) System in Bacteria. *Microbiol. Mol. Biol. Rev.* 76:565–584.
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. 2000. The Protein Data Bank. *Nucleic Acids Res.* [Internet] 28:235–242. Available from: <https://doi.org/10.1093/nar/28.1.235>
- Bi J, Ji Y, Pan J, Yu Y, Chen H, Zhu X. 2011. A new taxol-producing fungus (*Pestalotiopsis malicola*) and evidence for taxol as a transient product in the culture. *African J. Biotechnol.* 10:6647–6654.
- Biot C, Buisine E, Kwasigroch JM, Wintjens R, Rooman M. 2002. Probing the energetic and structural role of amino acid/nucleobase cation- $\pi$  interactions in protein-ligand complexes. *J. Biol. Chem.* 277:40816–40822.
- Blau N, Van Spronsen FJ, Levy HL. 2010. Phenylketonuria. *Lancet* 376:1417–1427.
- Boyle NMO, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR. 2011. Open Babel : An open chemical toolbox. :1–14.
- Brandl M, Weiss MS, Jabs A, Sühnel J, Hilgenfeld R. 2001. C-H $\cdots$  $\pi$ -interactions in proteins. *J. Mol. Biol.* 307:357–377.

- 
- Brown CS, Lichter-Konecki U. 2016. Phenylketonuria (PKU): A problem solved? *Mol. Genet. Metab. reports* 6:8–12.
- Carretero-Molina D, Ortiz-Lopez FJ, Martín J, González I, Sánchez-Hidalgo M, Román-Hurtado F, Díaz C, de la Cruz M, Genilloud O, Reyes F. 2020. Pentaminomycins F and G, first non-ribosomal peptides containing 2-pyridylalanine.
- Chang CY, Lohman JR, Cao H, Tan K, Rudolf JD, Ma M, Xu W, Bingman CA, Yennamalli RM, Bigelow L, et al. 2016. Crystal Structures of SgcE6 and SgcC, the Two-Component Monooxygenase That Catalyzes Hydroxylation of a Carrier Protein-Tethered Substrate during the Biosynthesis of the Eneidyne Antitumor Antibiotic C-1027 in *Streptomyces globisporus*. *Biochemistry* 55:5142–5154.
- Chang KL, Jeung YS. 1991. A Synthesis of 5, 6-Dihydrouracils in a Sealed-tube and Their Conformational Analysis. *Bull. Korean Chem. Soc.* 12:343–347.
- Chaudhury S, Lyskov S, Gray JJ. 2010. PyRosetta : a script-based interface for implementing molecular modeling algorithms using Rosetta. 26:689–691.
- Chauhan PS. 2020. Lignin nanoparticles: eco-friendly and versatile tool for new era. *Bioresour. Technol. Reports* 9.
- Chesters C, Wilding M, Goodall M, Micklefield J. 2012. Thermal bifunctionality of bacterial phenylalanine aminomutase and ammonia lyase enzymes. *Angew. Chemie - Int. Ed.* 51:4344–4348.
- Christianson C V., Montavon TJ, Van Lanen SG, Shen B, Bruner SD. 2007. The structure of L-tyrosine 2,3-aminomutase from the C-1027 enediyne antitumor antibiotic biosynthetic pathway. *Biochemistry* 46:7205–7214.
- Chung DDL, Chung D. 2012. Carbon fiber composites. Elsevier
- Cui Y, Wang Y, Tian W, Bu Y, Li T, Cui X, Zhu T, Li R, Wu B. 2021. Development of a versatile and efficient C–N lyase platform for asymmetric hydroamination via computational enzyme redesign. *Nat. Catal.* 4:364–373.
- Demuner IF, Colodette JL, Demuner AJ, Jardim CM. 2019. Biorefinery review:

- 
- Wide-reaching products through kraft lignin. *BioResources* 14:7543–7581.
- Deng JH, Luo J, Mao YL, Lai S, Gong YN, Zhong DC, Lu TB. 2020.  $\Pi$ - $\pi$  stacking interactions: Non-negligible forces for stabilizing porous supramolecular frameworks. *Sci. Adv.* 6:1–9.
- Deruiter J. 2005. Carboxylic Acid Structure and Chemistry : Part 1 .. H The dipolar nature of acids. :1–11.
- Dhankhar R, Gupta V, Kumar S, Kapoor RK, Gulati P. 2020. Microbial enzymes for deprivation of amino acid metabolism in malignant cells: biological strategy for cancer treatment. *Appl. Microbiol. Biotechnol.* 104:2857–2869.
- Dongre P, Driscoll M, Amidon T, Bujanovic B. 2015. Lignin-furfural based adhesives. *Energies* 8:7897–7914.
- Edwards PM. 2002. Origin 7.0: scientific graphing and data analysis software. *J. Chem. Inf. Comput. Sci.* 42:1270–1271.
- Ekielski A, Mishra PK. 2021. Lignin for bioeconomy: The present and future role of technical lignin. *Int. J. Mol. Sci.* 22:63.
- Emiliani G, Fondi M, Fani R, Gribaldo S. 2009. A horizontal gene transfer at the origin of phenylpropanoid metabolism: A key adaptation of plants to land. *Biol. Direct* 4:1–12.
- Van Erven G, de Visser R, Merckx DWH, Strotenberg W, de Gijsel P, Gruppen H, Kabel MA. 2017. Quantification of lignin and its structural features in plant biomass using  $^{13}\text{C}$  lignin as internal standard for pyrolysis-GC-SIM-MS. *Anal. Chem.* 89:10907–10916.
- Fache M, Boutevin B, Caillol S. 2016. Vanillin Production from Lignin and Its Use as a Renewable Chemical. *ACS Sustain. Chem. Eng.* 4:35–46.
- Feng L, Wanninayake U, Strom S, Geiger J, Walker KD. 2011. Mechanistic, mutational, and structural evaluation of a taxus phenylalanine aminomutase. *Biochemistry* 50:2919–2930.
- Filip A, Nagy EZA, Tork SD, B á n ó czi G, To\csa MI, Irimie FD, Poppe L, Paizs C,

- 
- Bencze LC. 2018. Tailored Mutants of Phenylalanine Ammonia-Lyase from *Petroselinum crispum* for the Synthesis of Bulky l-and d-Arylalanines. *ChemCatChem* 10:2627.
- Fu Y-M, Yu Z-X, Pelayo BA, Ferrans VJ, Meadows GG. 1999. Focal adhesion kinase-dependent apoptosis of melanoma induced by tyrosine and phenylalanine deficiency. *Cancer Res.* 59:758–765.
- Fuchs RL, Kane JF. 1985. In vivo synthesis of histidine by a cloned histidine ammonia-lyase in *Escherichia coli*. *J. Bacteriol.* 162:98–101.
- Fusetti F, Erlandsen H, Flatmark T, Stevens RC. 1998. Structure of tetrameric human phenylalanine hydroxylase and its implications for phenylketonuria. *J. Biol. Chem.* 273:16962–16967.
- Ge X, Chang C, Zhang L, Cui S, Luo X, Hu S, Qin Y, Li Y. 2018. Chapter Five - Conversion of Lignocellulosic Biomass Into Platform Chemicals for Biobased Polyurethane Application. In: Li Y, Ge X, editors. Vol. 3. Advances in Bioenergy. Elsevier. p. 161–213.
- Ghadimi, H., M. W. Partington and AH. 1962. Inborn error of histidine metabolism. *Pediatrics* 29:714–728.
- Ghaffar SH, Fan M. 2014. Lignin in straw and its applications as an adhesive. *Int. J. Adhes. Adhes.* 48:92–101.
- Ghoul M, Mitri S. 2016. The Ecology and Evolution of Microbial Competition. *Trends Microbiol.* 24:833–845.
- Gond SK, Kharwar RN, White JF. 2014. Will fungi be the new source of the blockbuster drug taxol? *Fungal Biol. Rev.* 28:77–84.
- Gupta S, Lau K, Harding CO, Shepherd G, Boyer R, Atkinson JP, Knight V, Olbertz J, Larimore K, Gu Z, et al. 2018. Association of immune response with efficacy and safety outcomes in adults with phenylketonuria administered pegvaliase in phase 3 clinical trials. *EBioMedicine* 37:366–373.
- Gutiérrez-Antonio C, Romero-Izquierdo AG, Gómez-Castro FI, Hernández S. 2021. 5

- 
- Production processes from lignocellulosic feedstock. In: Guti érez-Antonio C, Romero-Izquierdo AG, G ómez-Castro FI, Hern ández S, editors. *Production Processes of Renewable Aviation Fuel*. Elsevier. p. 129–169.
- Hanson KR, Havir EA. 1970. L-Phenylalanine ammonia-lyase: IV. Evidence that the prosthetic group contains a dehydroalanyl residue and mechanism of action. *Arch. Biochem. Biophys.* 141:1–17.
- Heberling Matthew M, Masman MF, Bartsch S, Wybenga GG, Dijkstra BW, Marrink SJ, Janssen DB. 2015. Ironing out their differences: dissecting the structural determinants of a phenylalanine aminomutase and ammonia lyase. *ACS Chem. Biol.* 10:989–997.
- Heberling MM, Wu B, Bartsch S, Janssen DB. 2013. Priming ammonia lyases and aminomutases for industrial and therapeutic applications. *Curr. Opin. Chem. Biol.* 17:250–260.
- Hofmarcher T, Lindgren P, Wilking N, J ö nsson B. 2020. The cost of cancer in Europe 2018. *Eur. J. Cancer* 129:41–49.
- Holme E, Lindstedt S. 1995. Diagnosis and management of tyrosinemia type I. *Curr. Opin. Pediatr.* 7:726–732.
- Islam MR, Hossain ME. 2021. Chapter 1 - Introduction. In: Islam MR, Hossain ME, editors. *Drilling Engineering. Sustainable Oil and Gas Development Series*. Gulf Professional Publishing. p. 1–16.
- Janiak C. 2000. A critical account on n-n stacking in metal complexes with aromatic nitrogen-containing ligands. *J. Chem. Soc. Dalton Trans.*:3885–3896.
- Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Ž ídek A, Potapenko A, et al. 2021a. Highly accurate protein structure prediction with AlphaFold. *Nature* 596:583–589.
- Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Ž ídek A, Potapenko A, et al. 2021b. Applying and improving AlphaFold at CASP14. *Proteins Struct. Funct. Bioinforma.*

---

89:1711–1721.

- Jun SY, Sattler SA, Cortez GS, Vermerris W, Sattler SE, Kang CH. 2018. Biochemical and structural analysis of substrate specificity of a phenylalanine ammonia-lyase. *Plant Physiol.* 176:1452–1468.
- Kataoka M, Miyakawa T, Shimizu S, Tanokura M. 2016. Enzymes useful for chiral compound synthesis: structural biology, directed evolution, and protein engineering for industrial use. *Appl. Microbiol. Biotechnol.* 100:5747–5757.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* 30:772–780.
- Khayyam H, Jazar RN, Nunna S, Golkarnarenji G, Badii K, Fakhrhoseini SM, Kumar S, Naebe M. 2020. PAN precursor fabrication, applications and thermal stabilization process in carbon fiber production: experimental and mathematical modelling. *Prog. Mater. Sci.* 107:100575.
- Khwanjaisakun N, Amornraksa S, Simasatitkul L, Charoensuppanimit P, Assabumrungrat S. 2020. Techno-economic analysis of vanillin production from Kraft lignin: Feasibility study of lignin valorization. *Bioresour. Technol.* 299:122559.
- Kohlmeier M. 2015. Chapter 8 - Amino Acids and Nitrogen Compounds. In: Kohlmeier M, editor. *Nutrient Metabolism (Second Edition)*. Second Edi. San Diego: Academic Press. p. 265–477.
- Kootstra A. 1994. Protection from UV-B-induced DNA damage by flavonoids. *Plant Mol. Biol.* 26:771–774.
- Kothiwale S, Mendenhall JL, Meiler J. 2015. BCL::Conf: small molecule conformational sampling using a knowledge based rotamer library. *J. Cheminform.* 7:47.
- Krug D, Müller R. 2009. Discovery of additional members of the tyrosine aminomutase enzyme family and the mutational analysis of CmdF. *ChemBioChem* 10:741–750.

- 
- Kvittingen EA. 1986. Hereditary tyrosinemia type I--an overview. *Scand. J. Clin. Lab. Invest. Suppl.* 184:27–34.
- Van Lanen SG, Oh T, Liu W, Wendt-Pienkowski E, Shen B. 2007. Characterization of the maduropeptin biosynthetic gene cluster from *Actinomadura madurae* ATCC 39144 supporting a unifying paradigm for enediyne biosynthesis. *J. Am. Chem. Soc.* 129:13082–13094.
- Lefort V, Longueville JE, Gascuel O. 2017. SMS: Smart Model Selection in PhyML. *Mol. Biol. Evol.* 34:2422–2424.
- Letunic I, Bork P. 2019. Interactive Tree of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* 47:256–259.
- Li M, Li Y, Zhang W, Li S, Gao Y, Ai X, Zhang D, Liu B, Li Q. 2018. Metabolomics analysis reveals that elevated atmospheric CO<sub>2</sub> alleviates drought stress in cucumber seedling leaves. *Anal. Biochem.* 559:71–85.
- Li R, Wijma HJ, Song L, Cui Y, Otzen M, Tian Y, Du J, Li T, Niu D, Chen Y, et al. 2018. Computational redesign of enzymes for regio- and enantioselective hydroamination. *Nat. Chem. Biol.* 14:664–670.
- Lu H, Cornell A, Alvarado F, Behm M, Leijonmarck S, Li J, Tomani P, Lindbergh G. 2016. Lignin as a binder material for eco-friendly Li-ion batteries. *Materials (Basel)*. 9:127.
- Eudmila, Hodásová et al. 2015. Lignin, potential products and their market value. *Wood Res* 60:973–986.
- M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ort WC. 2016. Gaussian 09, Revision A.02.
- MacDonald MJ, D’Cunha GB. 2007. A modern view of phenylalanine ammonia lyase. *Biochem. Cell Biol.* 85:273–282.
- Mahmood N, Yuan Z, Schmidt J, Xu C (Charles). 2016. Depolymerization of lignins

- 
- and their applications for the preparation of polyols and rigid polyurethane foams: A review. *Renew. Sustain. Energy Rev.* 60:317–329.
- Malik S, Cusidó RM, Mirjalili MH, Moyano E, Palazón J, Bonfill M. 2011. Production of the anticancer drug taxol in *Taxus baccata* suspension cultures: A review. *Process Biochem.* 46:23–34.
- Martău, Gheorghe Adrian, Lavinia-Florina Călinoiu and DCV. 2021. Bio-vanillin: Towards a sustainable industrial production. *Trends Food Sci. & Technol.*
- McGaughey GB, Gagné M, Rappé AK. 1998.  $\pi$ -Stacking interactions. Alive and well in proteins. *J. Biol. Chem.* 273:15458–15463.
- McGrogan BT, Gilmartin B, Carney DN, McCann A. 2008. Taxanes, microtubules and chemoresistant breast cancer. *Biochim. Biophys. Acta (BBA)-Reviews Cancer* 1785:96–132.
- McGuffin LJ, Bryson K, Jones DT. 2000. The PSIPRED protein structure prediction server. *Bioinformatics* 16:404–405.
- Meeks JC, Elhai J, Thiel T, Potts M, Larimer F, Lamerdin J, Predki P, Atlas R. 2001. An overview of the Genome of *Nostoc punctiforme*, a multicellular, symbiotic Cyanobacterium. *Photosynth. Res.* 70:85–106.
- Mikkonen KS. 2020. Strategies for structuring diverse emulsion systems by using wood lignocellulose-derived stabilizers. *Green Chem.* 22:1019–1037.
- Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, Lanfear R, Teeling E. 2020. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* 37:1530–1534.
- Moffitt MC, Louie G V, Bowman ME, Pence J, Noel JP, Moore BS. 2007. Discovery of two cyanobacterial phenylalanine ammonia lyases: kinetic and structural characterization. *Biochemistry* 46:1004–1012.
- Moretti R, Lyskov S, Das R, Meiler J, Gray JJ. 2018. Web-accessible molecular modeling with Rosetta: The Rosetta Online Server that Includes Everyone (ROSIE). *Protein Sci.* 27:259–268.

- 
- Müller MM. 2018. Post-Translational Modifications of Protein Backbones: Unique Functions, Mechanisms, and Challenges. *Biochemistry* 57:177–185.
- Nagy EZA, Tork SD, Lang PA, Filip A, Irimie FD, Poppe L, Toşa MI, Schofield CJ, Brem J, Paizs C, et al. 2019. Mapping the Hydrophobic Substrate Binding Site of Phenylalanine Ammonia-Lyase from *Petroselinum crispum*. *ACS Catal.* 9:8825–8834.
- Norval M. 2001. Chapter 5 - Effects of solar radiation on the human immune system. In: Giacomoni PU, editor. Sun Protection in Man. Vol. 3. Comprehensive Series in Photosciences. Elsevier. p. 91–113.
- Parmeggiani F, Weise Nicholas J., Ahmed ST, Turner NJ. 2018. Synthetic and therapeutic applications of ammonia-lyases and aminomutases. *Chem. Rev.*
- Peng F, Aliyu H, Delavault A, Engel U, Rudat J. 2021. Computational-Designed Enzyme for  $\beta$ -Tyrosine Production in Lignin Valorization. *Catalysts* 11:1310.
- Perrakis A, Sixma TK. 2021. AI revolutions in biology: The joys and perils of AlphaFold. *EMBO Rep.* 22:e54046.
- Petsko GA and SB. 1986. Amino-aromatic interactions in proteins. *FEBS Lett.* 203:139–143.
- Philip V, Harris J, Adams R, Nguyen D, Spiers J, Baudry J, Howell EE, Hinde RJ. 2011. A Survey of Aspartate to Phenylalanine and Glutamate to Phenylalanine. :2939–2950.
- Pilbæk S, Tomin A, Røst J, Poppe L. 2006. The essential tyrosine-containing loop conformation and the role of the C-terminal multi-helix region in eukaryotic phenylalanine ammonia-lyases. *FEBS J.* 273:1004–1019.
- Plesu Popescu AE, Torralba J, Bonet J, Llorens J. 2021. Vanillin production from lignin: Rigorous process simulation results for ethyl acetate versus aliphatic-alcohol-specific process designs. *Clean. Eng. Technol.* 4:100133.
- Prasad V, De Jesús K, Mailankody S. 2017. The high price of anticancer drugs: origins, implications, barriers, solutions. *Nat. Rev. Clin. Oncol.* 14:381–390.

- 
- Punekar NS. 2018. *Enzymes: catalysis, kinetics and mechanisms*. Springer
- Rachid S, Krug D, Weissman KJ, Müller R. 2007. Biosynthesis of (R)- $\beta$ -tyrosine and its incorporation into the highly cytotoxic chondramides produced by *Chondromyces crocatus*. *J. Biol. Chem.* 282:21810–21817.
- Raes J, Rohde A, Christensen JH, Van De Peer Y, Boerjan W. 2003. Genome-Wide Characterization of the Lignification Toolbox in *Arabidopsis*. *Plant Physiol.* 133:1051–1071.
- Ratnayake ND, Wanninayake U, Geiger JH, Walker KD. 2011. Stereochemistry and mechanism of a microbial phenylalanine aminomutase. *J. Am. Chem. Soc.* 133:8531–8533.
- Reid BG, Flynn GC. 1997. Chromophore formation in green fluorescent protein. *Biochemistry* 36:6786–6791.
- Rich MH, O'Connor KE, Narancic T. 2021. Non-canonical (unusual) amino acids as a toolbox for antimicrobial and anticancer peptide synthesis. *Amin. Acids, Pept. Proteins* 43:44–63.
- Richter F, Leaver-Fay A, Khare SD, Bjelic S, Baker D. 2011. De novo enzyme design using Rosetta3. *PLoS One* 6:e19230.
- Ritter H, Schulz GE. 2004. Structural basis for the entrance into the phenylpropanoid metabolism catalyzed by phenylalanine ammonia-lyase. *Plant Cell* 16:3426–3436.
- Röther D, Poppe L, Morlock G, Viergutz S, Réty J. 2002. An active site homology model of phenylalanine ammonia-lyase from *Petroselinum crispum*. *Eur. J. Biochem.* 269:3065–3075.
- Rouse B, Matalon R, Koch R, Azen C, Levy H, Hanley W, Trefz F, De La Cruz F. 2000. Maternal phenylketonuria syndrome: congenital heart defects, microcephaly, and developmental outcomes. *J. Pediatr.* 136:57–61.
- Sadhukhan J, Ng KS, Hernandez EM. 2014. *Biorefineries and chemical processes: design, integration and sustainability analysis*. John Wiley & Sons

- 
- Sánchez-Murcia PA, Bueren-Calabuig JA, Camacho-Artacho M, Cortés-Cabrera Á, Gago F. 2016. Stepwise Simulation of 3,5-Dihydro-5-methylidene-4H-imidazol-4-one (MIO) Biogenesis in Histidine Ammonia-lyase. *Biochemistry* 55:5854–5864.
- Sarker U, Oba S. 2018. Augmentation of leaf color parameters, pigments, vitamins, phenolic acids, flavonoids and antioxidant activity in selected *Amaranthus tricolor* under salinity stress. *Sci. Rep.* 8:1–9.
- Schafhauser T, Kirchner N, Kulik A, Huijbers MME, Flor L, Caradec T, Fewer DP, Gross H, Jacques P, Jahn L, et al. 2016. The cyclochlorotine mycotoxin is produced by the nonribosomal peptide synthetase CctN in *Talaromyces islandicus* ('*Penicillium islandicum*'). *Environ. Microbiol.* 18:3728–3741.
- Schrödinger, LLC. 2015. The {PyMOL} Molecular Graphics System, Version~1.8.
- Schwede TF, Réy J, Schulz GE. 1999. Crystal structure of histidine ammonia-lyase revealing a novel polypeptide modification as the catalytic electrophile. *Biochemistry* 38:5355–5361.
- Shang QM, Li L, Dong CJ. 2012. Multiple tandem duplication of the phenylalanine ammonia-lyase genes in *Cucumis sativus* L. *Planta* 236:1093–1105.
- Sharma A, Shahzad B, Rehman A, Bhardwaj R, Landi M, Zheng B. 2019. Response of phenylpropanoid pathway and the role of polyphenols in plants under abiotic stress. *Molecules* 24:1–22.
- Shee PK, Yan H, Walker KD. 2020. Intermolecular Amine Transfer to Enantioenriched trans-3-Phenylglycidates by an  $\alpha/\beta$ -Aminomutase to Access Both anti-Phenylserine Isomers. *ACS Catal.* 10:15071–15082.
- Sievers F, Higgins DG. 2014. Clustal omega. *Curr. Protoc. Bioinforma.* 48:3–13.
- Soliman SSM, Raizada MN. 2013. Interactions between co-habiting fungi elicit synthesis of Taxol from an endophytic fungus in host *Taxus* plants. *Front. Microbiol.* 4:1–14.
- Sorokina KN, Taran OP, Medvedeva TB, Samoylova Y V, Piligaev A V, Parmon VN.

- 
2017. Cellulose biorefinery based on a combined catalytic and biotechnological approach for production of 5-HMF and ethanol. *ChemSusChem* 10:562–574.
- Steele CL, Chen Y, Dougherty BA, Li W, Hofstead S, Lam KS, Xing Z, Chiang SJ. 2005. Purification, cloning, and functional expression of phenylalanine aminomutase: The first committed step in Taxol side-chain biosynthesis. *Arch. Biochem. Biophys.* 438:1–10.
- Strom S, Wanninayake U, Ratnayake ND, Walker KD, Geiger JH. 2012. Insights into the mechanistic pathway of the *Pantoea agglomerans* phenylalanine aminomutase. *Angew. Chemie - Int. Ed.* 51:2898–2902.
- Tabe Y, Lorenzi PL, Konopleva M. 2019. Amino acid metabolism in hematologic malignancies and the era of targeted therapy. *Blood, J. Am. Soc. Hematol.* 134:1014–1023.
- Tabuchi T. 2020. Cancer and Socioeconomic Status. *Soc. Determ. Heal. Non-Communicable Dis. Case Stud. from Japan; Kondo, K., Ed:*31–40.
- Takkellapati S, Li T, Gonzalez MA. 2018. An overview of biorefinery-derived platform chemicals from a cellulose and hemicellulose biorefinery. *Clean Technol. Environ. policy* 20:1615–1630.
- Taylor RG, Levy HL, McInnes RR. 1991. Histidase and histidinemia. Clinical and molecular considerations. *Mol. Biol. & Med.* 8:101–116.
- Thomas, Brian, Denis J. Murphy and BGM. 2016. Encyclopedia of applied plant sciences. *Acad. Press.*
- Tian X, Fang Z, Smith RL, Wu Z, Liu M. 2016. Properties, chemical characteristics and application of lignin and its derivatives. In: Production of biofuels and chemicals from lignin. Springer. p. 3–33.
- Timokhin VI, Regner M, Motagamwala AH, Sener C, Karlen SD, Dumesic JA, Ralph J. 2020. Production of p-Coumaric Acid from Corn GVL-Lignin. *ACS Sustain. Chem. Eng.* 8:17427–17438.
- Treutter D. 2006. Significance of flavonoids in plant resistance: A review. *Environ.*

- 
- Chem. Lett.* 4:147–157.
- Tria FDK, Landan G, Dagan T. 2017. Phylogenetic rooting using minimal ancestor deviation. *Nat. Ecol. Evol.* 1:193.
- Underlin EN, Frommhagen M, Dilokpimol A, Van Erven G, de Vries RP, Kabel MA. 2020. Feruloyl esterases for biorefineries: subfamily classified specificity for natural substrates. *Front. Bioeng. Biotechnol.* 8:332.
- Vuano BM, Pieroni OI, Cabaleiro MC. 2000. Thermal reaction of cinnamic acid and of  $\beta$ -styrylphosphonic acid with urea. *J. Chem. Res.* 2000:318–320.
- Walker KD, Klettke K, Akiyama T, Croteau R. 2004. Cloning, heterologous expression, and characterization of a phenylalanine aminomutase involved in taxol biosynthesis. *J. Biol. Chem.* 279:53947–53954.
- Walter T, King Z, Walker KD. 2016. A Tyrosine Aminomutase from Rice (*Oryza sativa*) Isomerizes (S)- $\alpha$ - to (R)- $\beta$ -Tyrosine with Unique High Enantioselectivity and Retention of Configuration. *Biochemistry* 55:1–4.
- Wang J, Yuan B, Huang B. 2019. Differential heat-induced changes in phenolic acids associated with genotypic variations in heat tolerance for hard fescue. *Crop Sci.* 59:667–674.
- Wang K, Hou Q, Liu Y. 2013. Insight into the mechanism of aminomutase reaction: A case study of phenylalanine aminomutase by computational approach. *J. Mol. Graph. Model.* 46:65–73.
- Wang L, Gamez A, Sarkissian CN, Straub M, Patch MG, Han GW, Striepeke S, Fitzpatrick P, Scriver CR, Stevens RC. 2005. Structure-based chemical modification strategy for enzyme replacement treatment of phenylketonuria. *Mol. Genet. Metab.* 86:134–140.
- Wang L, Wang S, He Q, Yu T, Li Q, Hong B. 2012. Draft genome sequence of *Streptomyces globisporus* C-1027, which produces an antitumor antibiotic consisting of a nine-membered enediyne with a chromoprotein.
- Wanninayake U. 2013. A mechanistic investigation of mio-based aminomutases.

---

Michigan State University

- Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. 2009. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25:1189–1191.
- Watts KT, Mijts BN, Lee PC, Manning AJ, Schmidt-Dannert C. 2006. Discovery of a Substrate Selectivity Switch in Tyrosine Ammonia-Lyase, a Member of the Aromatic Amino Acid Lyase Family. *Chem. Biol.* 13:1317–1326.
- WHO. 2020. <https://www.who.int/news-room/fact-sheets/detail/cancer>.
- Willcott MR. 2009. MestRe Nova. *J. Am. Chem. Soc.* [Internet] 131:13180. Available from: <https://doi.org/10.1021/ja906709t>
- Williamson JJ, Bahrin N, Hardiman EM, Bugg TDH. 2020. Production of substituted styrene bioproducts from lignin and lignocellulose using engineered *Pseudomonas putida* KT2440. *Biotechnol. J.* 15:1900571.
- Wu B, Szymanski W, Wietzes P, de Wildeman S, Poelarends GJ, Feringa BL, Janssen DB. 2009. Enzymatic synthesis of enantiopure  $\alpha$ - and  $\beta$ -amino acids by phenylalanine aminomutase-catalysed amination of cinnamic acid derivatives. *ChemBioChem* 10:338–344.
- Wu B, Szymański W, Wybenga GG, Heberling MM, Bartsch S, Dewildeman S, Poelarends GJ, Feringa BL, Dijkstra BW, Janssen DB. 2012. Mechanism-inspired engineering of phenylalanine aminomutase for enhanced  $\beta$ -regioselective asymmetric amination of cinnamates. *Angew. Chemie - Int. Ed.* 51:482–486.
- Wybenga GG, Szymanski W, Wu B, Feringa BL, Janssen DB, Dijkstra BW. 2014. Structural investigations into the stereochemistry and activity of a phenylalanine-2,3-aminomutase from *Taxus chinensis*. *Biochemistry* 53:3187–3198.
- Yan J, Aboshi T, Teraishi M, Strickler SR, Spindel JE, Tung C-W, Takata R, Matsumoto F, Maesaka Y, McCouch SR, et al. 2015. The Tyrosine Aminomutase TAM1 Is Required for  $\beta$ -Tyrosine Biosynthesis in Rice. *Plant Cell*

---

27:1265–1278.

- Yousuf A, Pirozzi D, Sannino F. 2020. Chapter 1 - Fundamentals of lignocellulosic biomass. In: Yousuf A, Pirozzi D, Sannino F, editors. *Lignocellulosic Biomass to Liquid Biofuels*. Academic Press. p. 1–15.
- Zhou M, Wang D, Peng R, Yang D, Qiu X, Qian Y. 2019. Synthesis of a Hindered Amine-Grafted Lignin-Based Emulsifier and Its Application in a Green Emulsifiable Concentrate. *J. Agric. Food Chem.* 67:11129–11136.
- Zhu L, Ge F, Li W, Song P, Tang H, Tao Y, Liu Y, Du G. 2018. One step synthesis of unnatural  $\beta$ -arylalanines using mutant phenylalanine aminomutase from *Taxus chinensis* with high  $\beta$ -regioselectivity. *Enzyme Microb. Technol.*

## 8 APPENDIX

### 8.1 Figure legend

- Figure 1: lignocellulosic biomass is composed of carbohydrates, including cellulose and hemicellulose, and aromatic polymer: lignin. Lignin contains mainly three monolignol subunits: *p*-coumaryl (H), coniferyl (G), and sinapyl (S). .....3
- Figure 2: Example of bioactive compounds containing  $\beta$ -phenylalanine. The unit of  $\beta$ -phenylalanine is colored in blue (Parmeggiani, et al. 2018). .....5
- Figure 3: Example of bioactive compounds containing  $\beta$ -tyrosine. The unit of  $\beta$ -tyrosine is colored in blue (Parmeggiani, et al. 2018). .....6
- Figure 4: (A) The X-ray crystal structure of TchPAM (PDB ID: 4C5R) with four monomeric chains, which are colored in white, light blue, orange, and purple. The MIO prosthetic group is marked in a red sphere (Wybenga et al. 2014). (B) The monomer contains three main domains: N-terminal domain (orange-red), central core domain (yellow), and C-terminal domain (dark blue), The MIO prosthetic group is placed in the N-terminal domain (red) (Feng et al. 2011). (C) Chemical Structure of MIO prosthetic group. (D) Each active site is composed of the residues from three monomeric chains. The MIO prosthetic group, which is marked in red, is the moiety in chain A, is surrounded by the residues from chain A (light blue), B (orange), and D (purple). .....8
- Figure 5: (A) Comparison of the X-ray structure of PAL, TAL, and HAL. The color of the RtPAL from *Rhodotorula toruloides* (PDB entry 1T6J) is yellow, while the PpHAL from *Pseudomonas putida* (PDB entry 1GKM) and RsTAL from *Rhodobacter sphaeroides* (PDB entry 2O6Y) are colored in blue and pink, respectively. (B) The inner loop (blue) and out loop (red) are shown in RsTAL with the MIO prosthetic group.....9
- Figure 6: High reaction temperature and more hydrophilic residues led to a flexible lid-like inner loop and a shift in the aminomutase activity to lyase activity..... 11
- Figure 7: (A) Comparison of the tripeptide in PpHAL from *Pseudomonas putida* and AvGFP from *Aequorea victoria*. (B) The tripeptide for the formation of the MIO prosthetic group with its functional essential neighboring residues in PpHAL. .... 12
- Figure 8: The E1cB catalytic mechanism of the MIO-dependent enzymes with phenylalanine and cinnamic acid as accepted substrates ..... 14
- Figure 9: Responding to the unfavorable stress parameters such as drought, heavy metal, salinity, and UV radiation, several phenolic metabolites are accumulated by regulation of the corresponding enzymes in the phenylpropanoid pathway of

---

higher plants.....	16
Figure 10 Histidine pathway in mammals, <i>Pseudomonas</i> , and the other bacterial genera with various enzymes. ....	18
Figure 11: Flowchart of the <i>de novo</i> enzyme design (Richter et al. 2011). ....	22
Figure 12: The main goal of this doctoral research is the computational design of TchPAM for the production of value-added lignin-based aromatic amino acids, such as $\beta$ -tyrosine, from monolignol: <i>trans-p</i> -hydroxycinnamic acid.....	23
Figure 13: Chemical synthesis reactions and methods. (A) OH-PheDU was synthesized in a two-necked flask from $\beta$ -tyrosine. (B) OH-PheDU was synthesized in pressured glass tubes of different sizes from <i>p</i> -hydroxycinnamic acid. (C) The monitor of chemical synthesis in a microwave-assisted reactor at different times. ....	37
Figure 14: Eutectic point test for the mixture of <i>p</i> -hydroxycinnamic acid and urea in different proportions. ....	59
Figure 15: TLC measurement of OH-PheDU, <i>N</i> -carbamoyl- $\beta$ -tyrosine (HPS), and $\beta$ -tyrosine. These products were solved in water or methanol. (A) TLC with ninhydrin reagent. (B) TLC under UV light at 254 nm. ....	61
Figure 16: The phylogenetic tree of the MIO-dependent enzymes. The protein sequences were collected from the SwissProt database (Bairoch et al. 2004). The characterized aromatic amino acid ammonia-lyases were displayed in the square, and the characterized aromatic amino acid aminomutases were marked as the circle. The different substrate acceptance was sorted in various colors: histidine in yellow, phenylalanine in green, and tyrosine in purple. The color of the outer ring indicated the source organisms. ....	65
Figure 17: Multiple sequence alignment of the PTALs from plants and fungi, the monofunctional PALs in monocots, and the bacterial TALs/TAMs. The surrounding residues in the active center, which are placed within 5 Å of the ligand backbone, are colored depending on their characterization of the side chain. ....	68
Figure 18: A heatmap for conservation analysis of the active site residues. Upon the appearance probability of the amino acids in various homologues, the corresponding squares are colored in red (hot, often appear) to light blue (cold, rarely appear). (A) The sequence alignment of homologues contained all the members from the MIO-dependent enzyme superfamily. (B) The sequence alignment of homologues consisted of PAM and PAL. ....	70
Figure 19: Schematic representation of the interactions between the natural substrate ( <i>S</i> )- $\alpha$ -phenylalanine and residues in the active center of TchPAM (PDB ID: 4C5R). The color coding corresponds to the position: carboxyl binding pocket (red), aromatic binding pocket (blue), tyrosine towards pocket (green), essential functional residues (black), and all other residues without specified function	

(purple). The substrate (e.g., (*S*)- $\alpha$ -phenylalanine) is highlighted by a yellow background. The highly conserved residues are not shown in this Figure. .... 72

Figure 20: SDS-Gel for comparison of the expression in different mediums. The band of TchPAM was observed between 63 kDa and 75 kDa. LB: The cells were incubated in the LB medium. TB: The cells were incubated in a TB medium. Aim: The cells were incubated in the auto-induction medium. Sup.: The supernatant of the enzyme solution after centrifugation, containing the soluble enzymes. Pre.: The precipitation of the enzyme solution after centrifugation. .... 78

Figure 21: The SDS-Gel of the untreated and heat-treated wild type. The band of TchPAM was observed between 63 kDa and 75 kDa. LB: The cells were incubated in the LB medium. Aim: The cells were incubated in the auto-induction medium. Sup.: The supernatant of the enzyme solution after centrifugation, containing the soluble enzymes. Pre.: The precipitation of the enzyme solution after centrifugation. The wild type, which was purified through the heat treatment at 65 °C, was marked in a red frame. .... 80

Figure 22: Comparison of the relative activity of the wild type and the designed mutants in the ammonia elimination of the (*S*)- $\alpha$ -phenylalanine. The wild type is colored in blue and the success mutants are colored in red. .... 82

Figure 23: Comparison of the relative activity of the wild type and the designed mutants in the ammonia elimination of the (*R*)- $\beta$ -phenylalanine. The wild type is colored in blue and the success mutants are colored in red. .... 83

Figure 24: Comparison of the relative activity of the wild type and the designed mutants in the ammonia elimination of the (*S*)- $\alpha$ -tyrosine. The wild type is colored in blue. .... 85

Figure 25: Comparison of the relative activity of the wild type and the designed mutants in the ammonia elimination of the (*R*)- $\beta$ -tyrosine. The wild type is colored in blue and the success mutants are colored in red. .... 86

Figure 26: The distance and the proposed interaction between the residue Tyr424 and the aromatic ring of the substrate in (A) the wild type, the mutants (B) Tyr424Asn, (C) Tyr424Cys, (D) Tyr424Gln, and (E) Tyr424Asp. The neighboring residue Arg427 and Glu455 are shown in (E). .... 99

Figure 27: Resonance effect of the phenolic ring of the *trans-p*-hydroxycinnamic acid. .... 101

Figure 28: The residues in the active center of TchPAM, which are placed close to the aromatic ring of the substrate. The residues in the aromatic binding pocket are colored in blue. The residues, including Tyr424, Lys427, Ile431, and Glu455, are colored in green. The natural substrate phenylalanine is colored in purple, while the color of the target substrate tyrosine and MIO moiety is orange. .... 102

Figure 29: The proposed mechanisms in the ammonia addition reaction by using

the *trans-p*-hydroxycinnamic acid as substrate. The phenolic ring orients the binding pocket, including the residues Tyr424, Lys427, and Glu455, when the substrate enters in the active site of the TchPAM and binds in the (A) binding mode I, which promotes the amination on  $\alpha$ -position of the substrate, yielding the (*S*)- $\alpha$ -tyrosine (B). The substrate can be stabilized by the residues Asn231, Arg325, and Arg355 in the binding mode II (C) and then form the (*R*)- $\beta$ -tyrosine as the final product. .... 103

Figure 30: The distance and the proposed interaction between the residue Ile431 and the aromatic ring of the substrate tyrosine in (A) the wild type, the mutants (B) Ile431Val. The distance and the proposed interaction between the residue Ile431 and the aromatic ring of the substrate phenylalanine in (C) the wild type, (D) Ile431Val. .... 105

Figure 31: Overlay of the hydrophobic residues in the active site of (*R*)-selective TchPAM (yellow, PDB ID: 4C5R) and (*S*)-selective PaPAM (purple, PDB ID: 3UNV). The ligand is  $\beta$ -phenylalanine, which is colored in orange. (A). the top view showed the enclosed aromatic binding pocket in TchPAM (yellow line) and the semi-enclosed aromatic binding pocket in PaPAM (blue line). (B). the side view showed the involved hydrophobic residues (TchPAM: black, PaPAM: blue) ..... 106

Figure 32: Proposed isomerization mechanism of (*R*)-selective TchPAM. When (*S*)- $\alpha$ -phenylalanine enters the active site of TchPAM, its carboxylate group is stabilized by residues Arg325 and Gln319. In the wild-type enzyme, the benzene ring is fixed in the aromatic binding pocket, in particular, with the residue Leu104, during the ammonia elimination to form the cinnamic intermediate. A single rotation around  $C_{\text{ipso}}-C_{\beta}$  bond without breaking the interaction with residue Leu104 brings the  $\beta$ -position on unsaturated acid to a suitable position for readdition reaction by MIO-amine adduct, giving (*R*)- $\beta$ -phenylalanine as the product. Meanwhile, the carboxylate group from the intermediate is stabilized by residues Arg325, Asn231, and Asn355. .... 107

Figure 33: In the mutant Leu104Ser, the benzene ring can move flexibly in the aromatic binding pocket, due to the lack of hydrophobic residue at position 104, during the ammonia elimination to form the cinnamic intermediate. Thus, the minimal shift of the intermediate backbone led to an immediate readdition of ammonia on  $\beta$ -position, as well as the formation of (*S*)- $\beta$ -phenylalanine. .... 108

Figure 34: The distance and the proposed interaction between the residue Gln459 and the aromatic ring of the substrate in (A) the wild type, the mutants (B) Gln459Glu, (C) Gln459Leu, (D) Gln459Cys, and (E) Gln459His. .... 111

Figure 35: The distance and the proposed interaction between the residue Asn458, the carboxylate group of the substrate, and the amino group of the residues Gln319 and Arg325 in (A) the wild type, the mutants (B) Asn458Cys, (C) Asn458Ala.... 113

---

Figure 36: The distance and the proposed interaction between the residue Leu104, the aromatic group of the substrate phenylalanine in (A) the wild type, the mutants (B) Leu104Ser, (C) Leu104Thr.....	116
Figure 37: Flow chart of a strategy for lignin valorization. The main goal of this doctoral research is framed in a red dotted line. ....	121
Figure 38: A proposed pathway for multi-enzymatic synthesis of ( <i>R</i> )- $\beta$ -tyrosine with high enantiopurity by starting from the pretreated lignin. ....	121
Figure 39: NMR spectrum of OH-PheDU .....	146
Figure 40: Plasmid map of TchPAM. The gene is cloned into pET28a.....	147
Figure 41: The SDS-PAGE gel of crude protein lysate from WT, purified enzyme from WT, and mutants (Tyr424Asn, Tyr424Cys, Tyr424Gln, Ile431Val, Leu104Ser, and Leu108Ser) by using His SpinTrap column (GE Healthcare Life Sciences, UK) for protein purification. WT: wild type.....	148
Figure 42: The SDS-PAGE gel of crude protein lysate from WT, purified enzyme from WT, and mutants (Lys427Gln, Lys427Glu, Asn458Ala, and Asn458Cys) by using His SpinTrap column (GE Healthcare Life Sciences, UK) for protein purification.....	149
Figure 43: The SDS-PAGE gel of crude protein lysate from WT, purified enzyme from WT, and mutants (Leu179Ser, Ile431Glu, Lys427Met, Phe86Tyr, Lys427Cys, and Glu455Asp) by using His SpinTrap column (GE Healthcare Life Sciences, UK) for protein purification.....	149
Figure 44: The SDS-PAGE gel of crude protein lysate from WT, purified enzyme from WT, and mutants (Gln459His, Gln459Cys, Gln459Leu, and Tyr424Asp) by using His SpinTrap column (GE Healthcare Life Sciences, UK) for protein purification.....	150
Figure 45: The SDS-PAGE gel of crude protein lysate from WT, purified enzyme from mutants (Glu455Asn, Glu455His, Leu108Gln, Leu104Thr, Lys427Asn, and Ile431Asn) by using His SpinTrap column (GE Healthcare Life Sciences, UK) for protein purification.....	150

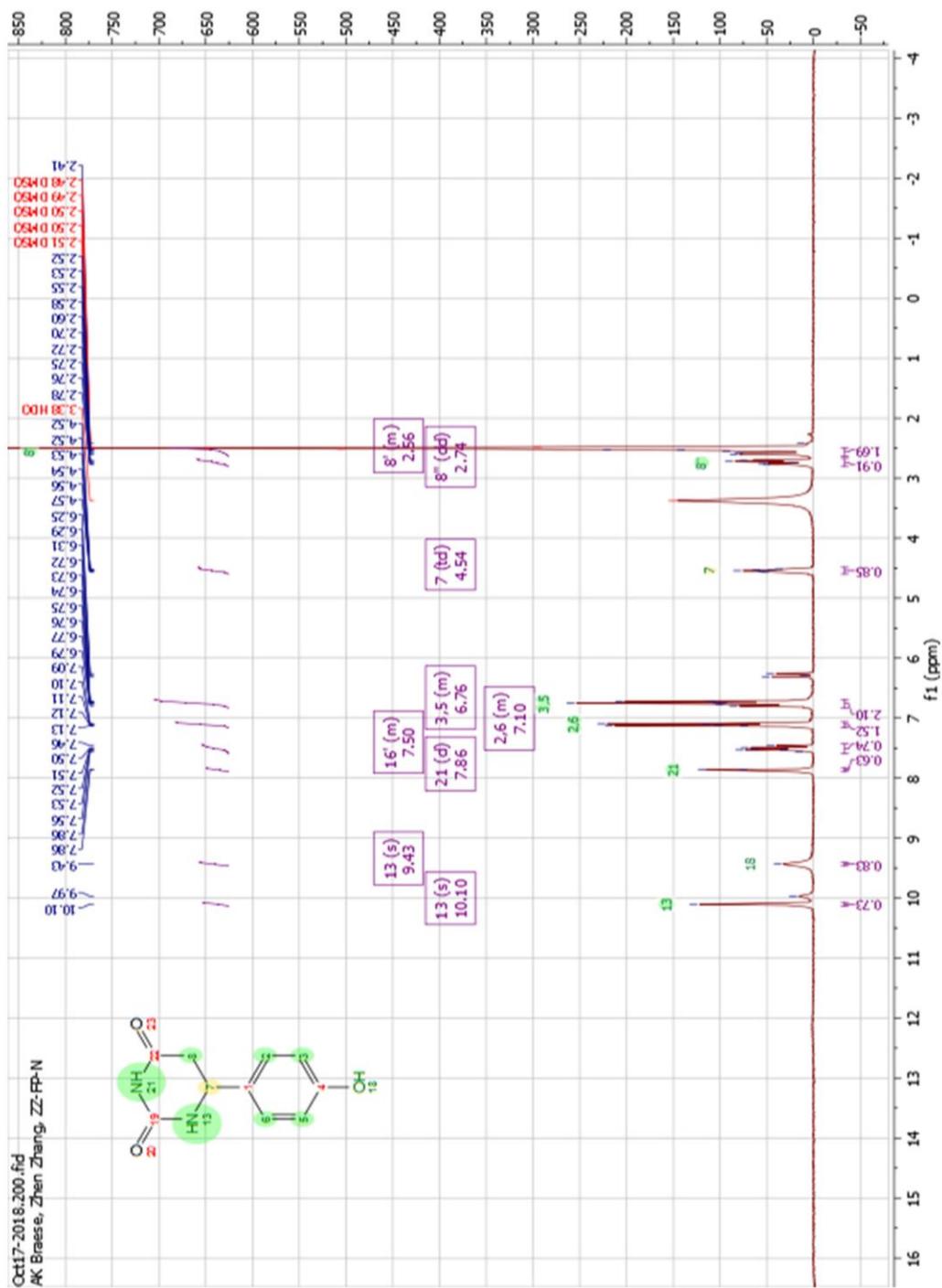
## 8.2 Table legend

Table 1: Distribution of the monolignol content in different types of lignin (Azadi et al. 2013). .....	4
Table 2: List of all chemicals used in this work.....	25
Table 3: Commercial enzymes, plasmids, and microorganisms .....	27
Table 4: The list of used commercial media .....	27
Table 5: Supplementary materials in media.....	27
Table 6: The list of the medium constitution .....	28
Table 7: Solution for protein purification .....	29
Table 8: Solution for thin-layer-chromatography .....	29
Table 9: Buffers and solution for agarose- and SDS-PAGE gels.....	30
Table 10: Solution for HPLC .....	31
Table 11: Primer list of variants Tyr424 by using KLD treatment.....	31
Table 12: Primer list of Touch-Down PCR products .....	31
Table 13: List of used software and web services.....	33
Table 14: Devices.....	33
Table 15: List of used supplies.....	35
Table 16: PCR batch with Q5® Hot Start High-Fidelity 2X Master Mix .....	49
Table 17: Touch-Down PCR protocol.....	50
Table 18: PCR protocol with Q5® Hot Start High-Fidelity 2X Master Mix.....	50
Table 19: Reagents for Kinase, Ligase, and DpnI (KLD) Treatment .....	52
Table 20: Solubility of OH-PheDU and <i>N</i> -carbamoyl- $\beta$ -tyrosine in different solvents .....	62
Table 21: The production yield of the OH-PheDU under different conditions.....	62
Table 22: A summary of dividing groups of the active site residues with their proposed functions. ....	73
Table 23: The summary of mutation limitation of different residue groups in the various program running. HP: The residues could be mutated in amino acids with a hydrophobic side chain. WL: The residues could be changed without any limitation. *: In this round, the residue Leu179 was not mutated. +: Only residue Asn458 was selected for mutation. Considering the spontaneous mutation of residue Phe86 in the other members of the MIO-dependent enzymes, the residue Phe86 could be mutated only in tyrosine. ....	76

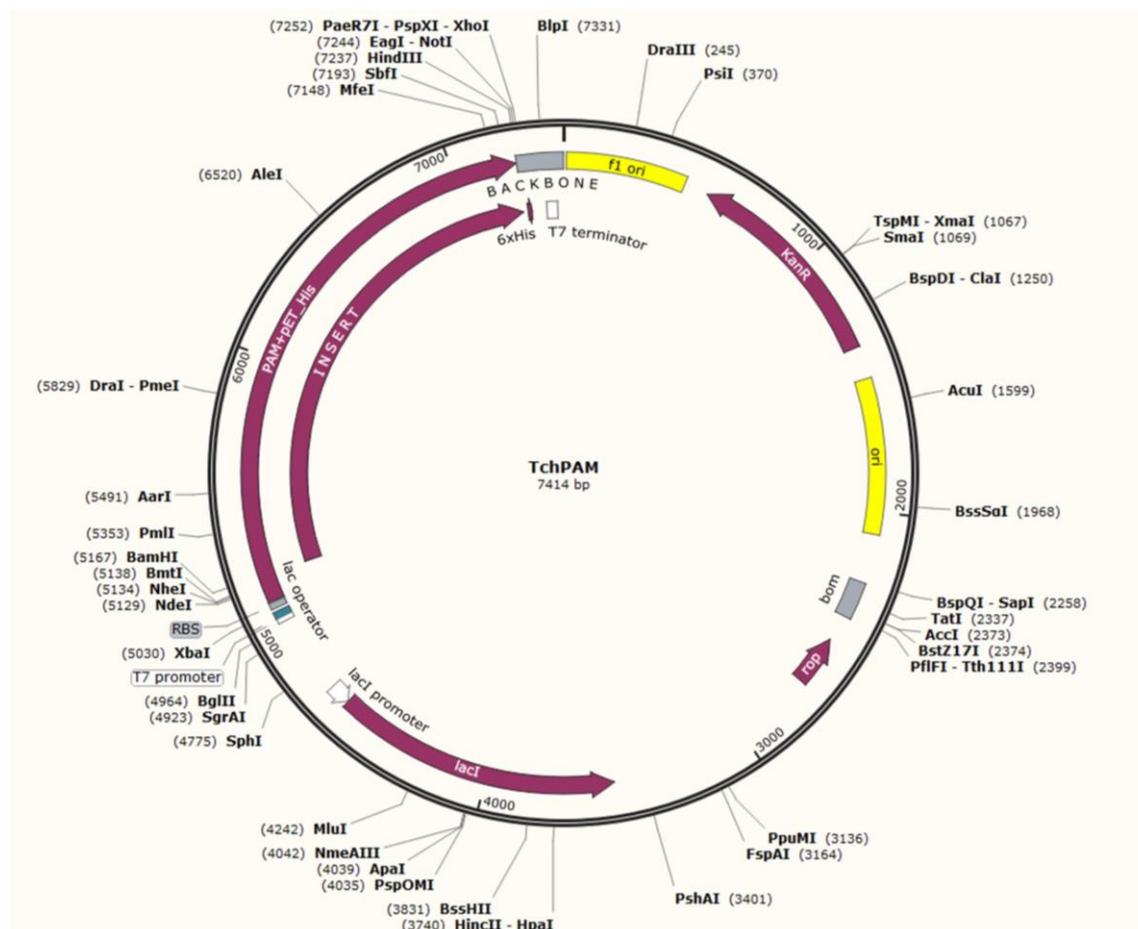
---

Table 24: List of the mutant library with different positions. ....	77
Table 25: Comparison of the kinetic parameters of the wild type and the mutants during the ammonia elimination of the ( <i>S</i> )- $\alpha$ -phenylalanine. The best-designed mutants with the highest relative activity are colored in blue. ....	87
Table 26: Comparison of the kinetic parameters of the wild type and the mutants during the ammonia elimination of the ( <i>R</i> )- $\beta$ -phenylalanine.....	88
Table 27: Comparison of the kinetic parameters of the wild type and the mutants during the ammonia elimination of the ( <i>S</i> )- $\alpha$ -Tyrosine. The best-designed mutants with the highest relative activity are colored in blue. ....	89
Table 28: Comparison of the kinetic parameters of the wild type and the mutants during the ammonia elimination of the ( <i>R</i> )- $\beta$ -Tyrosine. The best-designed mutants with the highest relative activity are colored in blue. ....	90
Table 29: Comparison of the specific activity and enantiomeric excess value of the wild type and the mutants Tyr424Asn, Tyr424Cys, Leu108Ser, and Ile431Val during the ammonia addition of the <i>trans-p</i> -hydroxycinnamic acid. ....	91
Table 30: Comparison of the specific activity and enantiomeric excess value of the wild type and the mutants Tyr424Asn, Tyr424Cys, Leu104Ser, and Ile431Val during the ammonia addition of the <i>trans-p</i> -cinnamic acid.....	91
Table 31: Comparison of the enantioselectivity of the tyrosine aminomutases from the SwissProt (Bairoch et al. 2004).....	100

## 8.3 NMR-spectra



## 8.4 Plasmid map

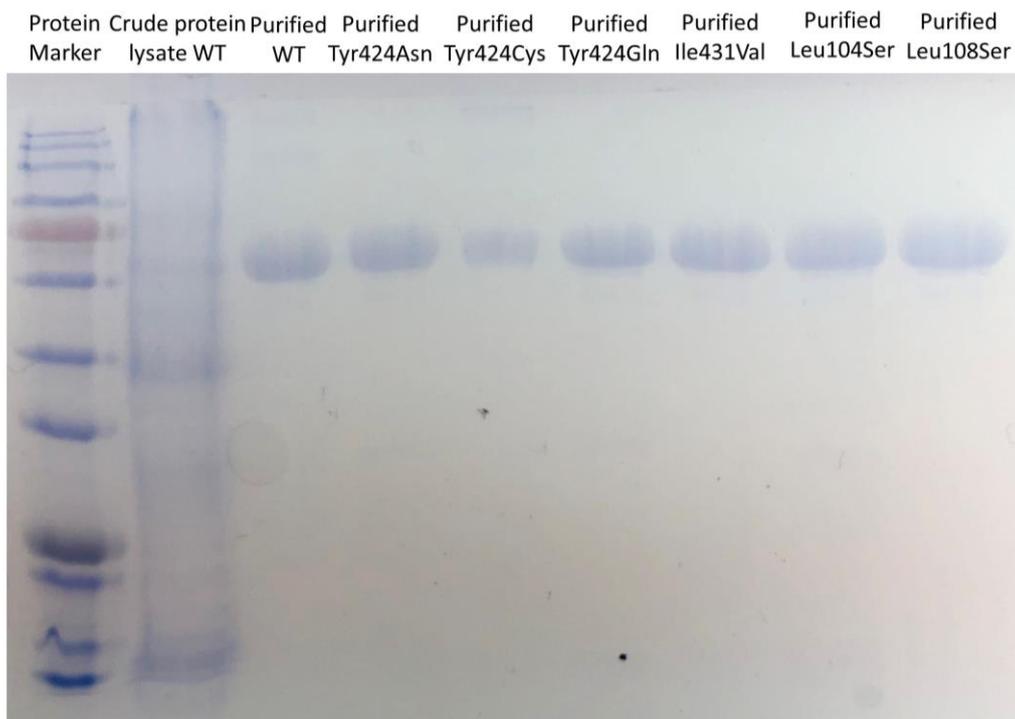


**Figure 40:** Plasmid map of TchPAM. The gene is cloned into pET28a

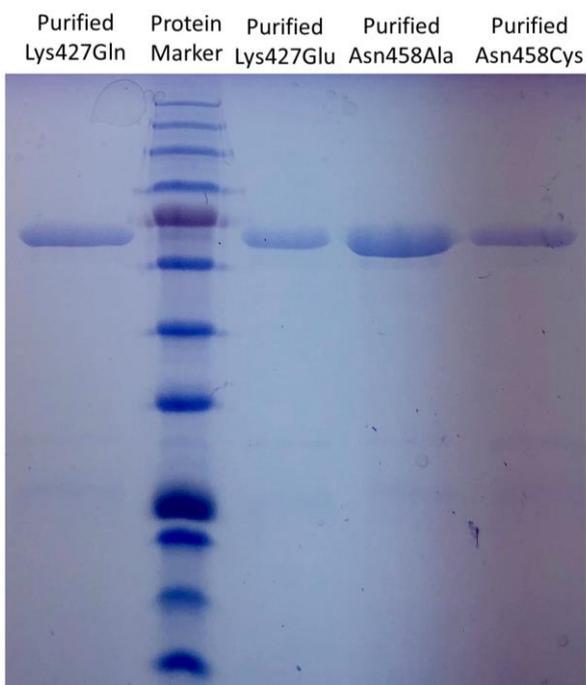
## 8.5 The amino acid sequence of wild type TchPAM

MGFAVESRSHVKDILGLINAFNEVKKITVDGTTTPITVAHVAALARRHDVKVAL  
 EAEQCRARVETCSSWVQRKAEDGADIYGVTTGFGACSSRRTNRLSELQESLIR  
 CLLAGVFTKGCAPSVDELPATATRSAMLLRLNSFTYGC SGIRWEVMEALEKLL  
 NSNVSPKVPLRGSVSASGDLIPLAYIAGLLIGKPSVIARIGDDVEVPAPALS RV  
 GLRPFKLQAKEGLALVNGTSFATAVASTVMYDANVLLLLVETLCGMFCEVIFG  
 REEFAHPLIHKVKPHPGQIESAELLEWLLRSSPFQELSREYYSIDKLLKPKQDR  
 YALRSSPQWLAPLVQTIRDATTTVETE VNSANDNPIIDHANDRALHGANFQGS  
 AVGFYMDYVRIAVAGLGKLLFAQFTELMIEYYSNGLPGNLSLGPDLSVDY GK  
 GLDIAMAAYSSELQYLANPVTTHVHSAEQHNQDINSLALISARKTEEALDILK  
 LMIASHLTAMCQAVDLRQLEEALVKVVENVVSTLADECGLPNDTKARLLYVA  
 KAVPVYTYLES PCDPTLPLLGLKQSCFDTILALHKKDGIETDTLVDR LAEF EK  
 RLSDRLENEMTAVRVLYEKKGHKTADNNDALVRIQGSKFLPFYRFVREELDTG  
 VMSARREQTPQEDV QKVFD AIADGRITVPLLHCLQGFLGQPNGCANGV

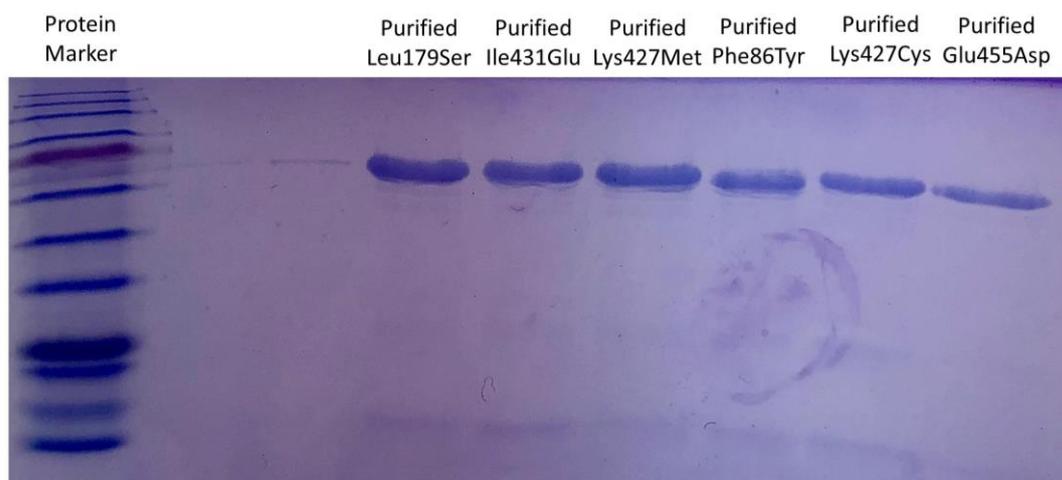
## 8.6 SDS-PAGE gels by His SpinTrap protein purification



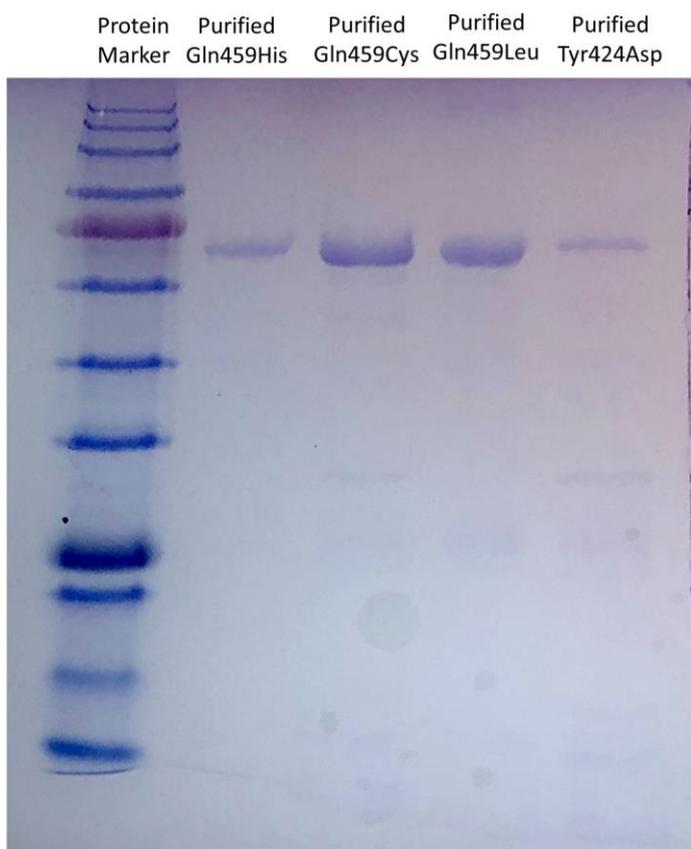
**Figure 41:** The SDS-PAGE gel of crude protein lysate from WT, purified enzyme from WT, and mutants (Tyr424Asn, Tyr424Cys, Tyr424Gln, Ile431Val, Leu104Ser, and Leu108Ser) by using His SpinTrap column (GE Healthcare Life Sciences, UK) for protein purification. WT: wild type.



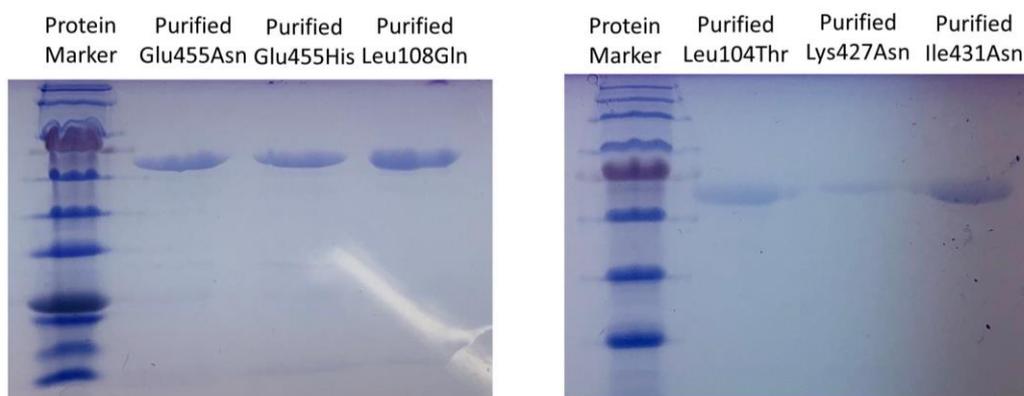
**Figure 42:** The SDS-PAGE gel of crude protein lysate from WT, purified enzyme from WT, and mutants (Lys427Gln, Lys427Glu, Asn458Ala, and Asn458Cys) by using His SpinTrap column (GE Healthcare Life Sciences, UK) for protein purification.



**Figure 43:** The SDS-PAGE gel of crude protein lysate from WT, purified enzyme from WT, and mutants (Leu179Ser, Ile431Glu, Lys427Met, Phe86Tyr, Lys427Cys, and Glu455Asp) by using His SpinTrap column (GE Healthcare Life Sciences, UK) for protein purification.



**Figure 44:** The SDS-PAGE gel of crude protein lysate from WT, purified enzyme from WT, and mutants (Gln459His, Gln459Cys, Gln459Leu, and Tyr424Asp) by using His SpinTrap column (GE Healthcare Life Sciences, UK) for protein purification.



**Figure 45:** The SDS-PAGE gel of crude protein lysate from WT, purified enzyme from mutants (Glu455Asn, Glu455His, Leu108Gln, Leu104Thr, Lys427Asn, and Ile431Asn) by using His SpinTrap column (GE Healthcare Life Sciences, UK) for protein purification.