



Anna Katharina Schweiger, BSc MSc

**Decarboxylation as Driving-Force:
Investigation of Structure-Function-Relationships in
Cofactor-Independent Decarboxylases and
Fe(II)-dependent Dioxygenases**

DOCTORAL THESIS

to achieve the university degree of
Doktorin der Naturwissenschaften (Dr.rer.nat.)

submitted to

Graz University of Technology

Supervisor

Univ.-Prof. Dr.rer.nat. Robert Kourist

Institute of Molecular Biotechnology

AFFIDAVIT

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present doctoral thesis.

Date, Signature

Danksagung

Zunächst möchte ich mich bei Prof. Robert Kourist für die Möglichkeit bedanken, die Stelle als Universitätsassistentin in seiner Gruppe anzutreten. Die vielfältigen Herausforderungen, vor die er mich gestellt hat, haben enorm zu meiner Weiterentwicklung als Wissenschaftlerin und Mensch beigetragen. Es war eine einzigartige Gelegenheit so nahe am Aufbau einer neuen Arbeitsgruppe mitwirken zu können.

Ich möchte mich außerdem bei Prof. Karl Gruber für die Begutachtung dieser Arbeit sehr herzlich bedanken. Des Weiteren danke ich Andrea Nigl, Dr. Daniel Kracher, Hanna Grimm und Leen Assil-Companiononi für das Korrekturlesen von Teilen dieser Arbeit.

Ich danke allen ehemaligen und aktuellen Mitgliedern dieser Arbeitsgruppe für die gemeinsame Zeit im Labor und auch außerhalb davon. Teil dieses dynamischen und internationalen Teams zu sein war eine echte Bereicherung. Namentlich erwähnen möchte ich an dieser Stelle Maria, Clemens und Simone für die wertvolle organisatorische Unterstützung im Labor und allen weiteren Drahtziehern im Hintergrund: Daniel, Leonie, Amneris und Svitlana. Vielen Dank Eveline, Dani und Eva für die bedeutsame organisatorische Hintergrundarbeit im Sekretariat.

Ein besonderer Dank geht an Hanna – mein norddeutsches Double und partner-in-crime der ersten Stunde. Egal ob in Forschung, Lehre oder auch privat: es ist von unschätzbarem Wert jemanden zu haben der einen ohne Worte versteht! Die gemeinsame Zeit mit all ihren Gesprächen wird mir sehr fehlen und ich wünsche dir alles Gute für dich und deinen wissenschaftlichen Nachwuchs.

Bedanken möchte ich mich auch bei meinen Bachelorstudenten, Annalena und Bernd, besonders aber bei meinen Masterstudentinnen, Steffi und Selina, für die großartigen Beiträge zu diesem herausfordernden Projekt. Danke für euren Fleiß und eure Geduld, falls es einmal nicht so einfach war.

Ein besonderer Dank ergeht an Assoc. Prof. Harald Pichler für die gemeinsame Zeit in den Praktika. Sein unermüdlicher Einsatz in der Lehre und seine fortwährende Unterstützung haben mich als Lehrende wesentlich mitgeprägt.

Ein großer Dank ergeht auch an meine ehemalige Arbeitsgruppe, die AKBR, welche mich wissenschaftlich grundlegend geprägt hat. Danke für die nach wie

vor aufrechten Kontakte, den guten Austausch und die freundliche Aufnahme zu den „AKBR Healthy Tuesdays“. Besonders hervorheben möchte ich hier die „elitäre Frühstücksrunde“ bestehend aus Anna, Julia und Rita. Auch wenn die Quantität etwas unter der Pandemie gelitten hat, so hat sich doch an der Qualität und unserer Ausdauer nichts geändert – danke für die unzähligen wunderbaren Stunden! An dieser Stelle möchte ich mich auch einmal mehr bei Kathrin bedanken, dass sie mir bereits während meiner Masterarbeit die Biokatalyse nähergebracht hat. Unsere gemeinsame Zeit werde ich immer in guter Erinnerung behalten!

Die wohl wichtigsten „Überbleibsel“ (im positivsten Sinne) meiner Studienzzeit sind ohne Zweifel Beate und Andi. Der bloße Gedanke an die Zeiten, die wir gemeinsam durchgestanden haben, erfüllt mich mit Freude und Trauer zugleich. Wenn auch unser gemeinsamer Weg vermeintlich hier in Graz ein Ende findet, so wird unsere Freundschaft weit darüber hinaus bestehen bleiben. Widerspruch zwecklos! Studium vergeht, Reisegruppe besteht!

Ich möchte mich auch bei allen lieben Freunden bedanken, die an dieser Stelle unerwähnt geblieben sind. Jeder einzelne, mehr oder weniger regelmäßige Kontakt ist unendlich wertvoll für mich!

Der größte Dank gilt meiner Familie, welche mich unterstützt, seitdem ich denken kann. Danke, dass ihr immer an mich glaubt, auch wenn es mir selber oft schwergefallen ist. Die gemeinsame Zeit zuhause war und ist für mich von unendlichem Wert, besonders um abzuschalten und auf andere Gedanken zu kommen. An dieser Stelle möchte ich mich auch bei meiner „zweiten“ Familie bedanken, für die Herzlichkeit und das Interesse, dass sie mir immer entgegenbringen. Ich freue mich auf Zeiten, wo man Fragen wie „Wann kommst du wieder?“ nicht mehr nur mit Verstand und Vernunft beantwortet, sondern wieder aus vollem Herzen.

Zuallerletzt möchte ich mich bei Felix bedanken, der es immer wieder schafft in jede noch so schwierige Situation etwas Leichtigkeit zu bringen. Bereits 41 % unseres bisherigen Lebens haben wir miteinander bestritten – aber niemand hätte gedacht, dass wir jemals ein Jahr beinahe ununterbrochen gemeinsam auf 38 m² verbringen werden. Also, was soll uns da noch aufhalten?

Danke!

The work presented in this thesis was conducted under the supervision of Univ.-Prof. Dr.rer.nat. Robert Kourist between March 2017 and March 2021 at the Institute of Molecular Biotechnology at Graz University of Technology.

» *Intelligence is the ability to adapt to change.* «

Stephen Hawking

Abstract

Biocatalysis is a rapidly developing field with a continuously growing number of successful synthetic applications in research and industry. Remarkable advances in protein and reaction engineering methodology combined with an increased availability of sequence and structural data in recent years have boosted the development of tailored biocatalysts to meet the requirements of a given process. Thereby, common limitations in biocatalysis like enzyme stability, substrate scope or substrate solubility can be overcome to fully benefit from the outstanding proficiency and selectivity of enzymes often unmet by chemical catalysts. Yet, the challenge remains to correlate the structure and function of an enzyme to enable predictability.

Within this work, various representative challenges faced *en route* to a broader synthetic applicability of enzymes were approached using the example of two cofactor-independent decarboxylases, phenolic acid decarboxylase from *Bacillus subtilis* (BsPAD) and arylmalonate decarboxylase from *Bordetella bronchiseptica* (AMDase), and the non-heme Fe(II)/ α KG dependent L-leucine dioxygenase from *Anabaena variabilis* (AvLDO).

Site-directed mutagenesis combined with characterization of the substrate scope of BsPAD revealed that substrate size is decisive for acceptance. Furthermore, it was proposed that the ability to precisely accommodate a given compound in the active site can result in rate enhancement compared to substrates with a less ideal fit. By introducing deep eutectic solvents (DES) as an alternative solvent system, the substrate load could be increased (300 mM) compared to aqueous systems (10 mM). DES further provided suitable conditions for both bio- and chemo-catalysts, thereby enabling a tandem reaction in continuous flow.

Involvement of alternative binding modes during decarboxylation of 2-methyl-2-vinylmalonic acid by the (S)-selective AMDase CLGIPL variant was investigated *via* rational design and substrate engineering. However, single amino acid exchanges did not result in significant improvement of product optical purity, thus underlining the importance of synergistic effects. Substrate engineering remained largely unsuccessful due to limitations of the synthetic strategy. Complementing

computational studies further provided evidence for a mechanism based on ground-state destabilization and highlighted the complex interplay within the active site.

As the determinants for substrate scope and selectivity of AvLDO are largely unknown, semi-rational design by site-saturation mutagenesis was the method of choice to expand the narrow substrate scope towards L-isoleucine. Active variants should be selected by growth complementation of an *E. coli* knock-out strain unable to produce succinate, which is in turn provided by substrate turnover of AvLDO. However, all attempts to confirm activity of the enriched variants remained unsuccessful. Particularly, enrichment of uncoupling variants could present a severe shortcoming of the applied selection approach.

The understanding of structure-function relationships of enzymes is often finite and consideration from various perspectives is needed to uncover molecular level determinants of enzyme activity and selectivity. Within this work, the added value of interdisciplinary approaches combining classical enzyme engineering with structural and computational analysis or substrate and reaction engineering was demonstrated.

Kurzfassung

Biokatalyse ist ein sich schnell entwickelndes Fachgebiet mit einer stetig wachsenden Anzahl an erfolgreichen synthetischen Anwendungen in Forschung und Industrie. Bemerkenswerte Fortschritte in der Methodik für Protein- und Reaktions-Engineering, gepaart mit einer erhöhten Verfügbarkeit von Sequenz- und Strukturdaten über die letzten Jahre, haben die Entwicklung von maßgeschneiderten Biokatalysatoren für bestimmte Prozesse vorangetrieben. Übliche Einschränkungen in der Biokatalyse, wie Enzymstabilität, Substratspektrum oder Löslichkeit der Substrate, können dadurch überwunden und die ausgezeichnete Leistungsfähigkeit und Selektivität von Enzymen, welche chemischen Katalysatoren oft überlegen sind, voll ausgenutzt werden. Die größte Herausforderung bleibt dennoch die Aufklärung von Struktur-Funktionsbeziehungen von Enzymen, um bessere Vorhersagen zu ermöglichen. In dieser Arbeit wurden repräsentative Herausforderungen am Weg zu einer breiteren synthetischen Anwendung von Enzymen anhand von zwei Cofaktor-unabhängigen Decarboxylasen, Phenolsäure-Decarboxylase aus *Bacillus subtilis* (BsPAD) und Arylmalonat-Decarboxylase aus *Bordetella bronchiseptica* (AMDase), und der nicht-häm Fe(II)/ α KG abhängigen L-Leucin Dioxygenase aus *Anabaena variabilis* (AvLDO) bearbeitet.

Ortsgerichtete Mutagenese und Charakterisierung des Substratspektrums von BsPAD haben gezeigt, dass die Größe des Substrates für dessen Umsetzung ausschlaggebend ist. Die präzise Platzierung einer Verbindung im aktiven Zentrum kann zu einer Erhöhung der Reaktionsrate, im Vergleich zu Substraten mit einer weniger optimalen Ausrichtung, führen. Durch die Einführung von „stark eutektischen Lösungsmitteln“ (DES) als alternatives Lösungsmittelsystem konnte die Substratkonzentration von 10 auf 300 mM im Vergleich zu wässrigen Lösungen erhöht werden. Dadurch wurden außerdem geeignete Reaktionsbedingungen für Bio- und Chemokatalyse geschaffen, wodurch Tandemreaktionen ermöglicht wurden.

Die Beteiligung von alternativen Bindungsmodi während der Umsetzung von 2-Methyl-2-vinylmalonsäure durch die (S)-selektive AMDase CLGIPL Variante wurde durch rationales Design und Substrat-Engineering untersucht. Vereinzelte

Substitutionen von Aminosäuren führten jedoch nicht zu einer erhöhten optischen Reinheit des Produktes. Dies hebt die Bedeutung von synergistischen Effekten hervor. Ergänzende computerbasierte Untersuchungen haben gezeigt, dass der Enzymmechanismus durch Destabilisierung des Grundzustandes getrieben wird und dass komplexe Wechselwirkungen im aktiven Zentrum eine wesentliche Rolle spielen.

Im Falle von AvLDO sind die bestimmenden Faktoren für Substratspektrum und Selektivität weitestgehend noch unbekannt. Semi-rationales Design durch Sättigungsmutagenese war daher die Methode der Wahl, um das eingeschränkte Substratspektrum in Richtung L-Isoleucin zu erweitern. Aktive Enzymvarianten sollen anschließend durch metabolische Komplementierung eines *E. coli* knock-out Stammes selektiert werden, welcher nicht in der Lage ist, Succinat zu produzieren. Dieses wird während Substratumsetzung durch AvLDO zur Verfügung gestellt. Es gelang jedoch nicht die Aktivität der angereicherten Varianten zu bestätigen. Insbesondere die Anreicherung von Entkopplungs-Varianten könnte eine gravierende Schwachstelle des angewendeten Selektionssystems darstellen.

Das Verständnis von Struktur-Funktionsbeziehungen von Enzymen ist oft begrenzt und es bedarf der Betrachtung aus unterschiedlichen Perspektiven, um die molekularen Einflussfaktoren auf Enzymaktivität und -selektivität aufzuklären. Der Mehrwert von interdisziplinären Ansätzen, welche klassisches Enzym-Engineering mit strukturellen und computerbasierten Analysen oder Substrat- und Reaktions-Engineering verbinden, wurde in dieser Arbeit gezeigt.

Publications

Parts of this work were published in the following articles:

Kourist, R.; Schweiger, A.; Büchsenschütz, H. Enzymatic Decarboxylation as a Tool for the Enzymatic Defunctionalization Organic Acids. In *Lipid Modification by Enzymes and Engineered Microbes*; Bornscheuer, U. T., Ed.; Elsevier, 2018; pp 89–118.

Schweiger, A. K.; Ríos-Lombardía, N.; Winkler, C. K.; Schmidt, S.; Morís, F.; Kroutil, W.; González-Sabín, J.; Kourist, R. Using Deep Eutectic Solvents to Overcome Limited Substrate Solubility in the Enzymatic Decarboxylation of Bio-Based Phenolic Acids. *ACS Sustainable Chem. Eng.* **2019**, 7, 16364–16370.

Grabner, B.; Schweiger, A. K.; Gavric, K.; Kourist, R.; Gruber-Woelfler, H. A Chemo-Enzymatic Tandem Reaction in a Mixture of Deep Eutectic Solvent and Water in Continuous Flow. *React. Chem. Eng.* **2020**, 5, 263–269.

Biler, M.;[‡] Crean, R. M.;[‡] Schweiger, A. K.; Kourist, R.; Kamerlin, S. C. L. Ground-State Destabilization by Active-Site Hydrophobicity Controls the Selectivity of a Cofactor-Free Decarboxylase. *J. Am. Chem. Soc.* **2020**, 142, 20216–20231.

A review covering parts of Chapter 4.1 is currently in preparation.

Table of Contents

Abstract.....	I
Kurzfassung.....	III
Publications	V
Table of Contents	VI
1 Introduction	1
2 Aim of this work.....	7
3 Phenolic acid decarboxylase from <i>Bacillus subtilis</i> (BsPAD).....	9
3.1 Theoretical background.....	9
3.1.1 Natural abundance of phenolic acids	9
3.1.2 Cofactor-independent phenolic acid decarboxylases.....	16
3.1.3 Recent applications.....	21
3.2 Aim of the project	26
3.3 Results and discussion	28
3.3.1 Generation of BsPAD I85V and I85L variants.....	29
3.3.2 Heterologous production and enzyme purification	30
3.3.3 Determination of specific activity of BsPAD and variants.....	32
3.3.4 Preparative-scale production of <i>p</i> -vinyl phenol derivatives	37
3.3.5 Immobilization of BsPAD.....	39
3.3.6 Applicability of BsPAD in deep eutectic solvents (DESs).....	42
3.4 Summary	49
3.5 Outlook.....	51
4 Arylmalonate decarboxylase from <i>Bordetella bronchiseptica</i> (AMDase)	52
4.1 Theoretical background.....	53
4.1.1 Discovery of AMDase.....	53
4.1.2 Inversion of enantioselectivity and introduction of racemase activity	56
4.1.3 Crystal structure and mechanism.....	57
4.1.4 Computational modelling on the enzyme mechanism.....	60
4.1.5 Substrate scope	62
4.1.6 Enzyme engineering	66
4.1.7 Recent applications.....	71
4.1.8 AMDases from other organisms.....	78
4.2 Aim of the project	82
4.3 Results and discussion	84
4.3.1 Selection of target residues for site-directed mutagenesis	84
4.3.2 Generation of AMDase CLGIPL variants	86
4.3.3 Heterologous production and purification of AMDase and variants	87

4.3.4	Substrate synthesis.....	89
4.3.5	Characterization of enantioselectivity of AMDase CLGIPL variants	102
4.3.6	Computational studies.....	112
4.4	Summary.....	119
4.5	Outlook.....	122
5	L-Leucine dioxygenase from <i>Anabaena variabilis</i> (AvLDO).....	123
5.1	Theoretical background.....	123
5.1.1	Classification of oxygenases.....	123
5.1.2	Structure and mechanism of Fe(II)/ α KG dependent dioxygenases	126
5.1.3	Fe(II)/ α KG dependent amino acid hydroxylases.....	130
5.1.4	Recent concepts	133
5.2	Aim of the project	138
5.3	Results and discussion	140
5.3.1	Selection of target residues for site-saturation mutagenesis	140
5.3.2	Preparation of an AvLDO library	143
5.3.3	Selection of AvLDO variants by enrichment culture.....	153
5.3.4	Rescreening of selected AvLDO variants	160
5.3.5	Heterologous production and purification of dioxygenases.....	168
5.3.6	Verification of the hydroxylating activity of AvLDO variants towards L-Ile.....	171
5.4	Summary.....	178
5.5	Outlook.....	181
6	Conclusion	182
7	Experimental section.....	184
7.1	Microbiological methods.....	184
7.1.1	Strains	184
7.1.2	Over-night cultures (ONCs).....	184
7.1.3	Glycerol stocks.....	185
7.2	Molecular biological methods.....	185
7.2.1	List of plasmids	185
7.2.2	Preparation of competent <i>E. coli</i> cells.....	186
7.2.3	Transformation of competent <i>E. coli</i> cells	187
7.2.4	Isolation of plasmid DNA.....	188
7.2.5	Gel extraction and PCR clean-up.....	188
7.2.6	Isolation of genomic DNA.....	189
7.2.7	Colony PCR (cPCR).....	189
7.2.8	Determination of DNA concentration.....	190
7.2.9	DNA sequencing and sequence analysis.....	190
7.2.10	Agarose gel electrophoresis.....	190
7.2.11	Mutagenesis.....	191
7.3	Biochemical methods	195

7.3.1	Heterologous production of recombinant enzymes.....	195
7.3.2	Preparation of cell-free extracts	196
7.3.3	Purification of enzymes	198
7.3.4	Determination of protein concentration	198
7.3.5	Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).....	199
7.4	Specific methods	200
7.4.1	Specific methods – <i>BsPAD</i>	200
7.4.2	Specific methods – <i>AMDase</i>	201
7.4.3	Specific methods – <i>AvLDO</i>	202
7.5	Analytical methods	206
7.5.1	Nuclear magnetic resonance (NMR).....	206
7.5.2	Gas chromatography (GC).....	207
7.5.3	High-performance liquid chromatography (HPLC)	209
7.5.4	Thin layer chromatography (TLC)	209
7.6	Chemical methods	210
7.6.1	General experimental aspects, materials and methods	210
7.6.2	Flash column chromatography.....	210
7.6.3	Experimental procedures	211
8	References.....	216
Appendix A	Abbreviations.....	i
Appendix B	Media, buffers and solutions	i
Appendix C	Primers	vi
Appendix D	DNA and protein sequences	viii
Appendix E	Agarose gels	x
Appendix F	SDS gels	xii
Appendix G	TLC.....	xxi
Appendix H	Raw data	xxii
Appendix I	GC supplementary data	xxiii
Appendix J	HPLC supplementary data	xxvii
Appendix K	NMR spectra	xxxii

1 Introduction

Enzymes are the catalysts of nature, which are, based on their primary structure constituted of 20 canonical L-amino acids, folded to complex three-dimensional structures with distinct functions. These structural scaffolds have been evolved by nature to ideally fit a given transformation in terms of catalytic efficiency, specificity and selectivity.¹ Despite their highly optimized roles in living organisms, enzymes are surprisingly promiscuous, which describes their ability to convert non-natural substrates (“substrate promiscuity”) or catalyze non-native reactions (“catalytic promiscuity”).²

For a long time, it was believed that the structure of an enzyme translates directly into its function, which was illustrated by the most famous “lock-and-key” principle introduced in 1894 by Emil Fischer, suggesting that structural interactions of enzyme and substrate are responsible for catalysis.³ Yet, this principle fails to explain, for instance, cooperative or allosteric effects, promiscuity of enzymes and their reaction mechanisms in general.⁴ Furthermore, such a tight interaction of enzyme and substrate would merely result in stabilization of the substrate ground state rather than catalytic transformation.⁵ Instead, enzymes selectively stabilize the transition state of a reaction thereby reducing the activation energy barrier for transformation by lowering the Gibbs free energy relative to the ground state ($\Delta\Delta G^\ddagger$, Figure 1a).² Thus, the enzyme active site needs to be complementary to the transition state rather than the substrate.⁵

The spatial arrangement of the overall enzyme and, particularly, the active site allows for selective recognition of substrates. Enzymes can be selective in various ways: (1) Most importantly, they are able to discriminate different functional groups, thereby allowing selective transformations of a certain functionality, while leaving the non-reacting groups unchanged. Through this *chemoselectivity*, side-reactivities as well as implementation of laborious protecting group strategies can be mostly prevented. (2) Through their *regioselectivity*, enzymes are further capable of distinguishing identical groups within the same substrate at different positions. (3) Enzymes can be also considered chiral catalysts due to the nature of their building blocks. Thus, they

are also able to recognize chiral compounds. This *enantioselectivity* allows enzymes to transform prochiral substrates into chiral products (desymmetrization) or to preferentially react with a certain enantiomer of a racemic mixture (kinetic resolution).²

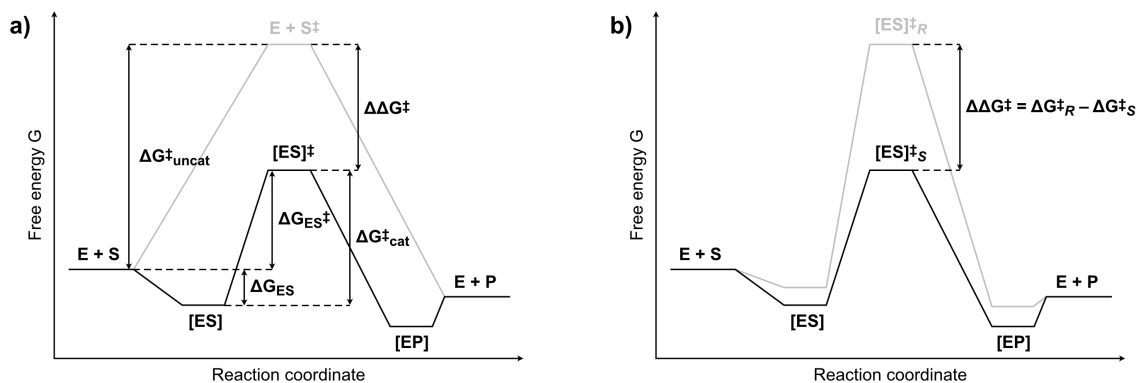


Figure 1. Energy profiles of enzyme catalyzed reactions. (a) Upon binding of the free substrate S, an enzyme-substrate complex [ES] is formed, which is stabilized compared to the free state (ΔG_{ES}). The catalyzed reaction proceeds *via* a transition state $[ES]^\ddagger$ and results in an enzyme-product complex [EP], where the product eventually dissociates from. In comparison to the uncatalyzed reaction (gray), the transition state $[ES]^\ddagger$ is energetically favored over the transition state S^\ddagger of the free substrate ($\Delta\Delta G^\ddagger$). (b) In the enzymatic conversion of chiral substrates, both enantiomers form distinct enzyme-substrate complexes $[ES]_R$ and $[ES]_S$ of different energy. Both complexes react *via* different transition state $[ES]_R^\ddagger$ and $[ES]_S^\ddagger$ and the net energy difference $\Delta\Delta G^\ddagger$ determines the enantioselectivity of the reaction.

While chemical catalysts are often unable to distinguish both enantiomers of a chiral molecule, the precisely defined active site allows enzymes to discriminate very well both structures by forming two distinct diastomeric enzyme-substrate complexes [ES] of different energy. This consequently results in both enantiomers reacting *via* energetically different transition states $[ES]_R^\ddagger$ and $[ES]_S^\ddagger$, where the net energy difference $\Delta\Delta G^\ddagger$ accounts for the observable enantioselectivity of a given transformation (Figure 1b).^{2,6}

Enzymes are well recognized for their ability to catalyze a great variety of different reactions under mild conditions in a highly efficient (typically 10^8 - 10^{10} fold rate enhancement compared to the uncatalyzed reaction) and selective manner, while being environmentally benign.^{1,2} This has also made them attractive targets for research and industry as substitutes for conventional chemical catalysts.

While earliest biotechnological “processes” in the production of fermented foods and beverages date back to ancient times,⁷ directed application of enzymes was only enabled by the several “waves” of biocatalysis which drove their broader

applicability forward. After realizing, over a century ago, that components of living cells are capable of mediating chemical reactions (first wave), these waves of extensive research in the field of biocatalysis have led to the development and advancement of crucial technologies including recombinant DNA technology (second wave, 1980s-1990s), protein engineering methods and high-throughput screening techniques used for “directed evolution” (third wave, late 1990s).⁸ Since then, advances in genome sequencing, availability of structural information and bioinformatic tools have boosted the identification of novel biocatalysts and development of metabolic engineering strategies, so that we might currently be already amidst the next, fourth, wave.⁹ Due to rapidly advancing protein engineering methodologies, biocatalysts can nowadays be adjusted to fit a predefined process *via* directed evolution. This is in sharp contrast to the traditional course of action, where industrial processes had to be adapted to fit a certain available (bio)catalyst.^{10,11}

Table 1. Required enzyme performance for the production of different chemicals. Table taken and adapted from Torrelo *et al.*¹²

	Product price range (Euro kg ⁻¹ _{product})	Allowable cost contribution of the catalyst (Euro kg ⁻¹ _{product})	Minimal productivity (kg _{product} kg ⁻¹ _{catalyst}) ^[a]	Minimal productivity (TTN) ^[b]
Pharma	> 100	10	20	4 000
Fine chemical	> 15	1.5	133	26 666
Specialty chemical	5	0.25	800	160 000
Bulk chemical	1	0.05	4 000	800 000

[a] Based on an average catalyst cost of 200 € kg⁻¹ (crude enzyme preparation). [b] Based on an average molecular weight of 40 kDa and a product molecular weight of 200 g mol⁻¹.

Despite the enormous progress made in the field of industrial biocatalysis, some preconceptions regarding enzymes still remain. The general opinion that enzymes are particularly expensive is still omnipresent, however the cost of a biocatalyst is strongly dependent on its formulation (*i.e.* crude, purified, immobilized, etc.) and scale of production. Most of all however, the product price range determines the legitimate cost contribution of a biocatalyst, which also

translates into minimal requirements in terms of catalyst productivity (Table 1).^{10,12} Implementation of a typically highly selective enzyme can further render protection and deprotection steps, required for chemical transformations, redundant and reduce side product formation, thus resulting in a simplified process.¹³ Other common restrictions like limited stability and imperfect activity or substrate scope can be tackled by protein engineering, whereas problems associated with substrate and product solubility or inhibition phenomena can be typically overcome by reaction engineering (e.g. in-situ product removal, alternative solvent systems, biphasic systems, etc.).^{12,13} It should be also noted that an enzymatic process is not intrinsically “green”, as, like in any other process, several factors including energy consumption or waste formation have to be considered. Several parameters, like the E-factor (which correlates the amount of product and waste produced), should be consulted to assess the sustainability of a process.^{10,12}

Due to the enormous selectivity and specificity offered by enzymes, their application in production of highly value-added compounds such as pharmaceuticals (low volume and high price) or other fine chemicals is obvious,^{14–16} yet biocatalytic bulk chemical production (high volume and low price) is also gaining ground, particularly in urge for more sustainable processes. Typically, feedstock accounts for the majority of overall production cost of bulk chemicals (70-80 %), thus requiring the biocatalyst to operate at high product concentrations for more efficient downstream processing and with high volumetric productivity and yield to minimize cost. Apart from that, the biocatalyst needs to be stable and robust and ideally reusable.¹¹ However, sustainability of a given process does not only concern the catalyst, but also the feedstock. That is, while the raw material obviously needs to be cheap, it should also be renewable and must be available without competing with food and feed supply (i.e. unrefined waste materials like lignocellulose).¹⁷ Impressive examples of enzymes established for bulk chemical production include immobilized glucose isomerase (GI) for the manufacture of high-fructose corn syrup (continuous process, approx. 1 kg L⁻¹ h⁻¹, approx. 0.05 g immobilized enzyme per kg of HFCS, 14 million t a⁻¹ dry weight)¹⁸ and nitrile hydratase (NHase) produced in *Rhodococcus rhodochrous* J1 for the production of acrylamide from acrylonitrile (> 99.99 % yield, approx. 2 kg L⁻¹ d⁻¹, 650 000 t a⁻¹),^{19,20} yet both processes rely

on refined feedstocks. Thus, valorization of unrefined feedstocks still presents a major challenge on the way to fully sustainable processes, which might particularly benefit from the extraordinary selectivity of enzymes due to unknown or fluctuating composition of these waste materials.¹⁷

To evolve a biocatalyst for a given application, a wide array of different enzyme engineering methodologies has been developed to date and, depending on the goal and available knowledge of the target enzyme, a suitable approach can be selected thereof (Figure 2).²¹

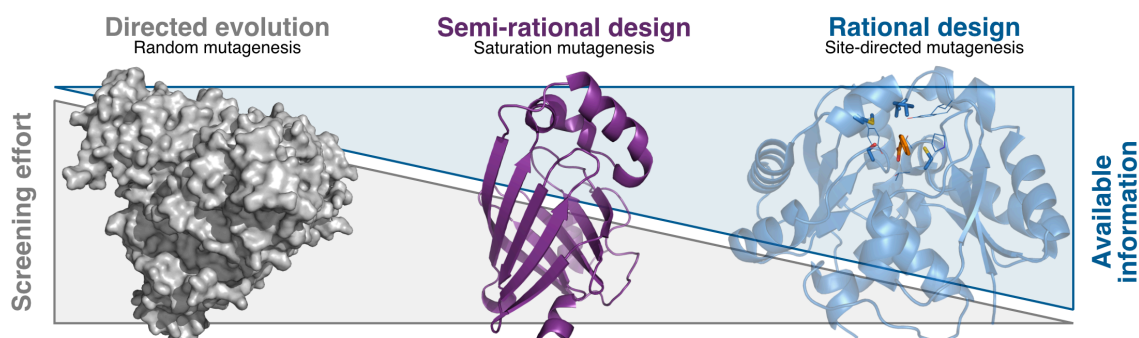


Figure 2. Overview of different protein engineering techniques, which require different levels of available information and screening effort. Figure adapted from Balke *et al.*²²

Clearly, directed evolution presents an extremely powerful and versatile approach, which virtually requires no prior knowledge of the enzyme structure, as mutagenesis is performed purely random (e.g. error-prone PCR, site-saturation mutagenesis, gene shuffling, etc.).^{23–25} The resulting high number of variants (typically 10^3 - 10^6) is further screened for the desired characteristics before additional rounds of mutagenesis and screening are performed to eventually obtain an improved biocatalyst. Thus, the success of directed evolution is determined in large part by the availability of an appropriate and effective high-throughput screening technique for the targeted enzyme characteristics.²⁶ Directed evolution has enabled successful improvement of a plethora of enzymes to fit a given process in terms of selectivity, stability or substrate scope.^{23,24,27} As already implied by the name, rational design presents the extreme opposite of directed evolution. Based on available sequence and structural information as well as on knowledge of structure-function relationships and crucial functional residues, targets for site-directed mutagenesis can be selected.²⁸

Semi-rational design can be seen as hybrid technology which combines both formerly mentioned approaches to create so-called focused or smart libraries.^{29,30} Only limited knowledge on structure and function of a target enzyme is necessary to select a restricted set of residues putatively important for the aspired characteristic for (typically random) mutagenesis. These smart libraries are much more likely to comprise variants with desirable characteristics.^{29,31} Also from a statistical perspective, by using a restricted amino acid alphabet, all relevant residue properties can still be covered in the library, however at a significantly reduced screening effort.³¹ Reetz and coworkers have made significant contributions to the field of semi-rational design, including methodologies such as CASTing (combinatorial active-site saturation test)³² or iterative saturation mutagenesis (ISM).³³ So far, numerous examples have been reported where enzyme activity, substrate specificity and enantioselectivity could be successfully improved.^{30,34–36}

The enormous advances in the field of biocatalysis in recent years were possible thanks to the progress made in bioinformatics, availability of structural data, enzyme engineering methodology including high-throughput screening techniques and gene sequencing and synthesis technology. It is nowadays possible to tailor a biocatalyst for a given application even at industrially suitable timelines.³⁷ Particularly, the combination of bio-bio- and bio-chemo-catalysis in cascades or pathway engineering can open remarkable opportunities for production of highly complex molecules. To established truly sustainable processes, a holistic perspective needs to be taken, including all steps of up- and downstream processing and potential issues connected to that. In this regard, the valorization of non-fossil and non-food carbon feedstocks (*e.g.* waste, gases, lignocellulose) will be a major challenge for future sustainable processes.^{37,38}

2 Aim of this work

Within this work, three different enzymes were investigated, particularly with regard to typical challenges faced on the way to a broader synthetical, if not industrial, applicability of these biocatalysts (see Chapter 1).

An intrinsic challenge in biocatalysis is the need to work with highly dilute reaction systems due to limited substrate solubility, as exemplified by *BsPAD* within this work (see Chapter 3). Phenolic acid decarboxylase from *Bacillus subtilis* (*BsPAD*) is an intriguing enzyme which catalyzes the decarboxylation of bio-based *para*-hydroxycinnamic acid derivatives without the aid of an external cofactor. The resulting 4-vinylphenol derivatives present interesting monomers for the production of functionalized bio-derived polymers.³⁹ While *BsPAD* catalyzes the decarboxylation at typical mild conditions (30°C, pH 6), chemical decarboxylation requires high temperatures in addition to chemical catalysts.^{39–41} Despite the robustness and high activity of *BsPAD*, its broad applicability is largely hampered by the low solubility of its substrates in aqueous medium (10 mM). Besides characterization of the substrate scope of *BsPAD* in classical aqueous environment, implementation a non-conventional solvent system – in particular deep eutectic solvents (DES) – was envisaged to overcome current limitations and to study the influence of different media on substrate scope.

While enzymes often show reactivities which do not have a chemical counterpart, the limited understanding of molecular level determinants of enzyme activity or selectivity hinders predictability and further development of a tailored catalyst. In the present work, this challenge is represented by *AMDase* (see Chapter 4). Arylmalonate decarboxylase from *Bordetella bronchiseptica* (*AMDase*) is an unique enzyme, which catalyzes the stereoselective decarboxylation of prochiral aryl- or alkenylmalonic acids without relying on a cofactor, thereby giving rise to chiral carboxylic acids in high optical purity. Intriguingly, *AMDase* is able to catalyze decarboxylation of highly unstable electron poor malonates with outstanding selectivity.^{42–45} Despite the stepwise progress made over the last years towards the utility of *AMDase* in an industrial context, some of the fundamental molecular drivers of *AMDase* selectivity and activity still remained

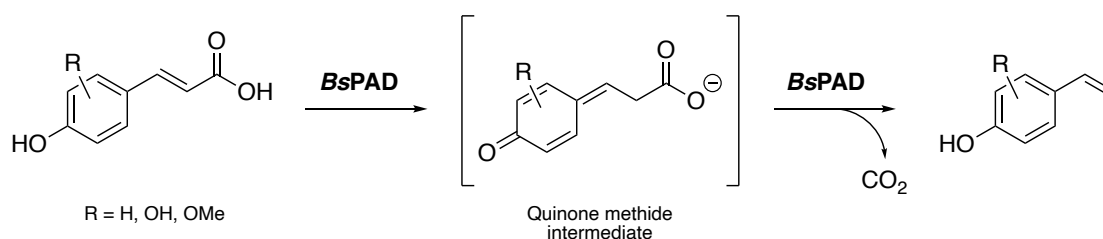
unclear.⁴⁶ We thus aimed at attaining a deeper understanding of these determinants by rational design of the (S)-selective AMDase CLGIPL variant to influence stereoselectivity during conversion of vinylmalonic acids, which was further complemented by computational studies to unravel crucial contributors of AMDase activity and selectivity.

Lastly, in many cases, a given enzyme certainly possesses interesting or desirable reactivities, but its applicability is limited by a narrow substrate scope or low activity. While a plethora of effective enzyme engineering methods has been developed until today, the success of enzyme evolution is closely linked to the availability of a proficient selection or screening system.²⁶ This major challenge is illustrated in the present work by using the example of AvLDO (see Chapter 5). L-leucine dioxygenase from *Anabaena variabilis* (AvLDO) is a non-heme Fe(II)/ α KG dependent dioxygenase which exhibits δ -specific hydroxylation activity towards L-leucine and L-norleucine and sulfoxidation activity towards L-methionine and ethionine. Enzymatic C-H functionalization is of particular interest, as it still presents a major challenge of chemical catalysis.⁴⁷ Several non-proteinogenic hydroxyamino acids or products thereof have been shown to be valuable chiral building blocks for pharmaceutical industry or natural product synthesis.^{47,48} Furthermore, the usability of amino acid dioxygenases in formation of diastereomerically pure hydroxylated amino acids from racemic *N*-acetylamino acids by combining enzyme catalyzed dynamic kinetic resolution and hydroxylation was demonstrated.⁴⁹ Yet, the substrate scope of AvLDO is strictly limited and the activity towards the accepted substrates low. In this work we aimed at extending the substrate scope of AvLDO towards L-isoleucine by semi-rational design *via* simultaneous saturation mutagenesis of several residues close to the putative substrate binding site. AvLDO variants able to convert L-isoleucine should be selected by complementing growth of *E. coli* 3Δ *sucA* through providing the coproduct succinate, thereby restoring the interrupted TCA cycle of this knock-out strain.⁵⁰

3 Phenolic acid decarboxylase from *Bacillus subtilis* (BsPAD)

Phenolic acid decarboxylase (PadC or BsPAD, EC 4.1.1.102), isolated from the soil bacterium *Bacillus subtilis*, is able to convert naturally abundant *p*-hydroxycinnamic acids into the corresponding 4-vinyl phenol derivatives *via* non-oxidative decarboxylation (Scheme 1). As the enzyme purely acts *via* acid-base catalysis without the need of an external co-factor, the substrate scope is strictly limited to *para*-hydroxylated compounds, thus enabling substrate anchoring and electron rearrangement towards a quinone methide intermediate. This highly unstable intermediate eventually collapses to release CO₂ and the corresponding 4-vinyl phenol derivative as decarboxylation product. The ability of PAD-expressing microorganisms to catalyze this specific reaction is thought to be closely connected to their natural habitat, where decaying biomass and, consequently, free phenolic acids are omnipresent.

Scheme 1. Overview on BsPAD catalyzed decarboxylation of *p*-hydroxy cinnamic acids proceeding *via* a quinone methide intermediate, which releases CO₂ upon collapse to eventually form 4-vinyl phenol derivatives.



3.1 Theoretical background

3.1.1 Natural abundance of phenolic acids

3.1.1.1 Biosynthesis of phenolic acids

Phenolic acids represent a subclass of plant secondary metabolites derived from the phenylpropanoid pathway. The pathway is fed by the primary metabolites phenylalanine or tyrosine and yields numerous aromatic secondary metabolites, which can be divided into five groups, which are monolignols (C6-C3 backbone),

flavonoids (C6-C3-C6 backbone), coumarins (C6-C3 backbone), stilbenes (C6-C2-C6 backbone) and phenolic acids, regarding their structural differences.

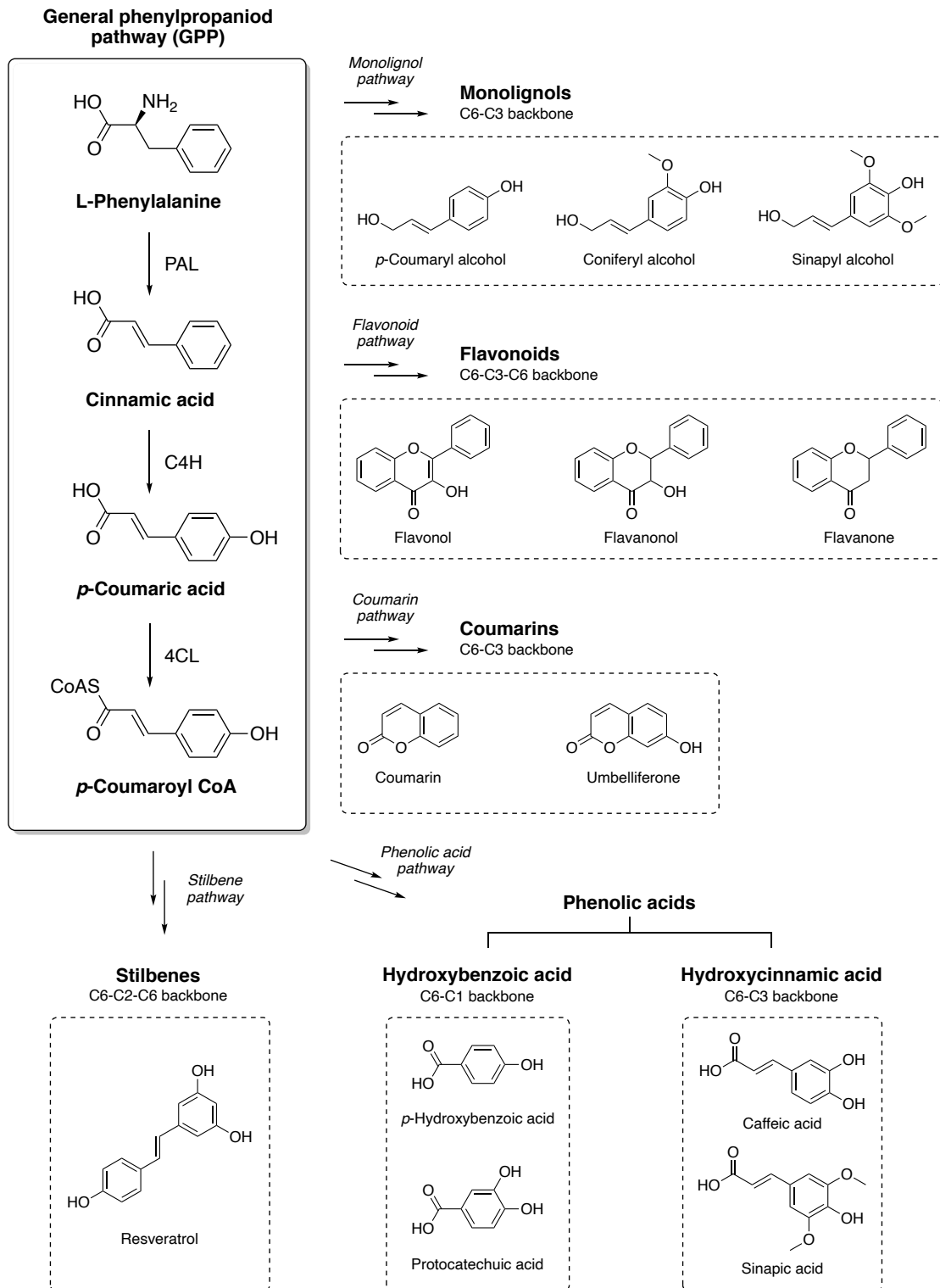


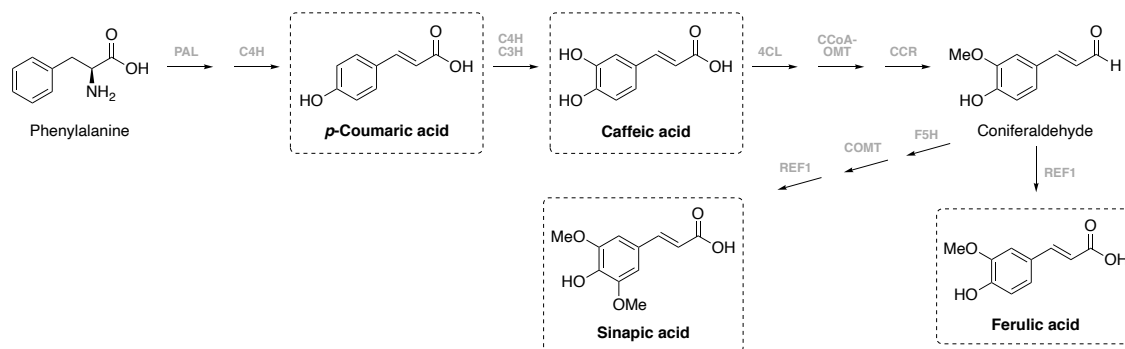
Figure 3. Schematic overview of phenylpropanoid biosynthetic pathways with representative compounds of each substance class (in dashed boxes). Initially, L-phenylalanine is converted to the central intermediate *p*-coumaroyl CoA in three steps. PAL: phenylalanine ammonia lyase, C4H: cinnamate 4-hydroxylase, 4CL: 4-coumaroyl CoA ligase. Figure adapted from Deng *et al.*⁵¹

The latter group can be further split into hydroxybenzoic acids (C6-C1 backbone) and hydroxycinnamic acids (C6-C3 backbone) (Figure 3). Compounds derived from the phenylpropanoid pathway are well recognized for their vital role in plant protection and survival and their beneficial health effects in humans.⁵¹

The core of phenylpropanoid biosynthesis is comprised of the general phenylpropanoid pathway (GPP), consisting of two to three steps, which branches into specific side pathways, depending on the compound class, thereby channeling carbon flux from the primary to the secondary metabolism. In most plants, the general phenylpropanoid pathway starts from phenylalanine, originating from the Shikimate pathway. Deamination of phenylalanine catalyzed by phenylalanine ammonia lyase (PAL) is considered the initial reaction step. The formed cinnamic acid is further hydroxylated by cinnamate 4-hydroxylase (C4H) and eventually transformed into the corresponding cinnamoyl-CoA ester by 4-coumaroyl CoA ligase (4CL), which represents a central intermediate for off-branching pathways towards diverse phenylpropanoids.⁵¹ In some plants (e.g. monocots) however, the action of tyrosine ammonia lyase (TAL) or bifunctional ammonia lyase (PTAL) can bypass cinnamic acid formation, by directly converting tyrosine into *p*-coumaric acid.^{51,52}

Despite the high abundance of phenolic acids in the plant kingdom, several steps of their biosynthetic pathways remain unclear or unknown to date, while transformation of intermediates shared with the monolignol pathway are much better understood.⁵¹ While the transformation of *p*-coumaric acid into caffeic acid, catalyzed by C4H/C3H membrane protein complex,⁵³ is relatively straightforward, formation of ferulic or sinapic acid is much more complex. Coniferaldehyde, a central intermediate of the monolignol pathway, is obtained from caffeic acid in three steps of CoA-ester formation (by 4-hydroxycinnamoyl-CoA ligase), methyl group transfer (by caffeoyl-CoA O-methyl transferase), and reduction (by cinnamoyl-CoA reductase).^{51,54} Ferulic and sinapic acid are probably oxidation products of the corresponding aldehyde precursors, coniferaldehyde and sinapaldehyde, the latter of which is obtained in three steps of hydroxylation (by ferulate 5-hydroxylase), methyl group transfer (by caffeic acid O-methyl transferase) and oxidation (by *Arabidopsis* REF1) from coniferaldehyde (Scheme 2).^{51,55,56}

Scheme 2. Simplified schematic overview on the biosynthesis of selected hydroxycinnamic acids (in dashed boxes). The underlying genes encoding enzymes responsible for the respective transformations are indicated in gray. PAL: phenylalanine ammonia lyase, C4H: cinnamate 4-hydroxylase, C3H: *p*-coumarate 3-hydroxylase, 4CL: 4-hydroxycinnamoyl-CoA ligase, CCoAOMT: caffeoyl-CoA O-methyl transferase, CCR: cinnamoyl-CoA reductase, F5H: ferulate 5-hydroxylase, COMT: caffeic acid O-methyl transferase, REF1: reduced epidermal fluorescence 1. Figure adapted from Deng *et al.*⁵¹



3.1.1.2 Role of phenolic acids in plants

Lignocellulose represents the major fraction of plant biomass and is largely responsible for their structural rigidity.^{57,58} It is mainly composed of cellulose, hemicellulose, a heteropolysaccharide which can be covalently linked to lignin, and lignin, which is an aromatic heteropolymer. Due to its complex architecture, containing various types of covalent carbon-carbon and carbon-oxygen cross-linkages, the susceptibility of lignin to hydrolytic degradation by invading microbes is relatively low, thus providing a persistent physical barrier against lignocellulose breakdown.⁵⁸ The exact composition of lignocellulose and its interconnectivity greatly varies amongst different plant species.^{57–59} In addition to the physical barrier provided by lignins, chemical modification of polysaccharides further reduces their enzymatic digestibility.⁶⁰ That is, substitution of the polysaccharide main chain or cross-linkage between polysaccharide chains mainly by acetylation or esterified phenolic acids, predominately ferulic and *p*-coumaric acid.^{60,61} Due to their bifunctional character (*i.e.* hydroxy and carboxyl function), phenolic acids are further able to interconnect the highly hydrophobic lignin and hemicelluloses, but also lignin itself.⁶² In particular, ferulate residues on xylan play a significant role as nucleation sites during lignin polymerization, thereby anchoring lignin to the polysaccharide fraction.^{59,63} Depending on the plant species, different phenolic acid modifications of lignocellulose are present, including esterification to arabinose residues of carbohydrates (*e.g.* 5-*O*-feruloyl

groups), ester- or etherification to lignin and coupling of two polysaccharides by dehydrodiferulic acid bridges, which can be formed from two ferulic acid units by peroxidase activity (Figure 4).^{59,62} It was demonstrated, that such ferulic acid dehydrodimer bridges drastically enhance persistence of polysaccharides against polysaccharidases, cellulases, xylanases, pectinases and laccases released by plant pathogens.⁵⁹

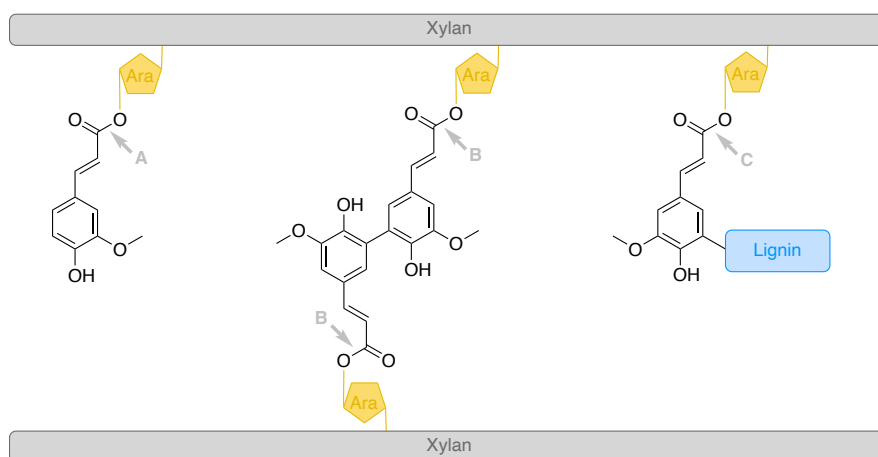


Figure 4. Schematic structure of arabinoxylan, a hemicellulose type as found in the primary cell wall of various plants. The polysaccharide xylan main-chain is decorated with α -arabinopyranose units (in yellow), which are involved in cross-linking *via* ferulic acid residues. Different types of interconnections are commonly observed: (A) 5-*O*-feruloyl group, (B) 5-*O*-diferuloyl group cross-linking two polysaccharide chains, (C) 5-*O*-feruloyl-lignin. Figure adapted from Williamson *et al.*⁶¹

Hence, hydrolytic enzymes able to release esterified phenolic acids from their association to biopolymers (*e.g.* oligosaccharides or lignin) play a crucial role in lignocellulose degradation/utilization and have been isolated from a wide range of microorganisms.⁶¹ It was also shown that such enzymes often act synergistically with hemicellulases (*e.g.* xylanases and pectinases), which facilitate access and action of side-chain specific hydrolases like acetyl xylan esterases or feruloyl esterases.⁶⁴

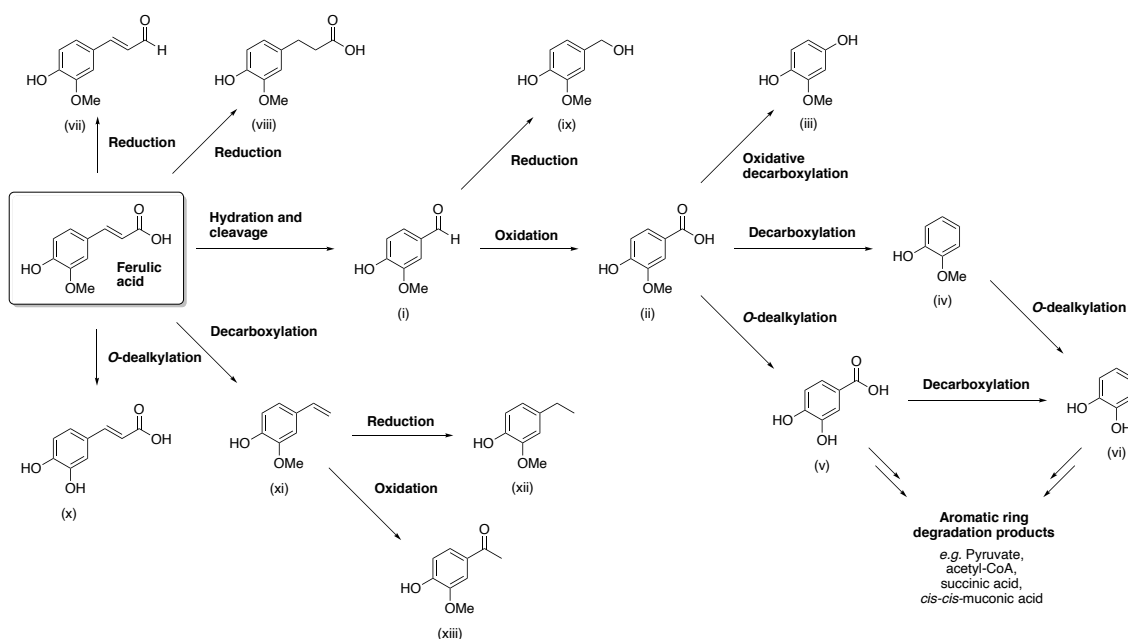
3.1.1.3 Role of phenolic acid decarboxylases

Due to the anti-microbial properties of phenolic acids,^{65–69} it is unsurprising that microorganisms often found in association to the habitat of rotten plants or woods developed strategies to detoxify this compound class.

Microorganisms have evolved a plethora of different pathways to metabolize phenolic acids. Metabolic routes, commonly found in bacteria, yeasts and fungi

during degradation of ferulic acid are summarized in Scheme 3. Transformation of ferulic acid into vanillin (Scheme 3, compound i) *via* cleavage of acetate plays a central role in the degradation cascade.⁷⁰ This conversion might also proceed *via* a CoA-dependent mechanism where ferulic acid is initially converted into feruloyl CoA by feruloyl CoA synthetase (FerA) and further into vanillin and acetyl CoA by feruloyl CoA hydratase/lyase (FerB).⁵⁸ Oxidation of vanillin yields vanillic acid (Scheme 3, compound ii), which in turn is a central node for further transformations into 2-methoxyhydroquinone (Scheme 3, compound iii) *via* oxidative decarboxylation, guaiacol (Scheme 3, compound iv) *via* decarboxylation or protocatechuic acid (Scheme 3, compound v) *via* O-demethylation.⁷⁰ Protocatechuic acid and catechol (Scheme 3, compound vi) are the main intermediates to undergo ultimate oxidative aromatic ring cleavage.^{70,71} Likewise, the degradation of *p*-coumaric acid and caffeic acid also involves formation of protocatechuic acid (Scheme 3, compound v) as a central metabolite.⁷¹

Scheme 3. Possible pathways for microbial ferulic acid degradation. Common intermediates include (i) vanillin, (ii) vanillic acid, (iii) 2-methoxyhydroquinone, (iv) guaiacol, (v) protocatechuic acid, (vi) catechol, (vii) coniferyl aldehyde, (viii) dihydroferulic acid, (ix) vanillyl alcohol, (x) caffeic acid, (xi) 4-vinylguaiacol, (xii) 4-ethylguaiacol, (xiii) 4-acetylguaiacol. Figure adapted from Rosazza *et al.*,⁷⁰ Bugg *et al.*⁵⁸ and Tinikul *et al.*⁷¹



Furthermore, reductive transformations of ferulic acid or metabolites are known, leading to saturated derivatives, aldehydes or alcohols (Scheme 3, compound

vii-ix), but also formation of caffeic acid (Scheme 3, compound x) *via* O-dealkylation is described. Finally, ferulic acid can also undergo non-oxidative decarboxylation to yield 4-vinylguaiacol (Scheme 3, compound xi), which was the first observed degradative route of phenolic acids.⁷⁰

Upon phenolic acid exposure, expression of phenolic acid decarboxylases is upregulated as part of a phenolic acid stress-response (PASR) mechanism in microorganisms.⁷² In *Bacillus subtilis*, transcription of the *padC* gene, encoding phenolic acid decarboxylase PadC, is negatively regulated by the repressor PadR, which is constitutively expressed at low levels. In presence of phenolic acids, the binding of PadR is inactivated and the transcription of *padC* enabled.^{72,73} This is in contrast to other organisms exhibiting comparable PAD activities, like *Lactobacillus plantarum*⁷⁴ or *Pediococcus pentosaceus*,⁷⁵ where PadR is subject of negative autoregulation, which initially leads to increased PadR production upon phenolic acid exposure, thus requiring higher levels of phenolic acids to effectively release the repressor.^{72,74,75} These regulatory differences allow *B. subtilis* to react faster when exposed to phenolic acids, thus rendering it more competitive within its commonly nutrient-poor habitat (*i.e.* soil and plant ecosystem).⁷²

While the toxicity of phenolic acids, possibly *via* their cell membrane permeabilizing ability due to their lipophilic character,⁷⁶ was well studied, knowledge on the anti-microbial properties of the formed vinyl phenol products was scarce. However, due to comparable composition of the aromatic moieties and even more hydrophobic character compared to phenolic acids, similar membrane damaging effects could be expected.⁷⁷ When the impact of both phenolic acids and vinyl phenols on growth of several gram-positive and gram-negative bacteria was studied, considerable differences in sensitivity towards both compound classes were observed. Several gram-positive bacteria, which had the ability to face phenolic acid toxicity by PAD activity, were insensitive towards vinyl phenols. In contrast, tested gram-negative strains were relatively insensitive towards phenolic acids, whereby they displayed significant inhibition by the corresponding vinyl phenols.⁷⁸ This contradicts previous assumptions, that vinyl phenols are less toxic than their phenolic acid counterparts.^{72,74,75} Particularly, wildtype *E. coli* and PAD expressing recombinant *E. coli* strains were significantly inhibited by 4-vinyl phenol. In

contrast, upon exposure to *p*-coumaric acid, transcription of *aaeB* and *marA* genes, both of which encode efflux pumps for antimicrobial substances, was upregulated, leading to the conclusion that the observed toxicity of vinyl phenols could be due to the absence of appropriate efflux systems for this compound class and resulting accumulation in the cytoplasm.⁷⁸

3.1.2 Cofactor-independent phenolic acid decarboxylases

While a plethora of phenolic acid decarboxylases exists in nature, the subclass of cofactor-independent decarboxylases is particularly fascinating. In contrast to other metal- or cofactor-dependent decarboxylases,⁷⁹ these enzymes only rely on the set of canonical amino acids to promote non-oxidative decarboxylation merely *via* acid-base catalysis. Yet, due to absence of a cofactor, the substrate itself has to serve as an electron sink during the reaction, thus severely limiting the substrate scope to *para*-hydroxylated cinnamic acid derivatives.

3.1.2.1 Structure and mechanism

Despite the role of phenolic acid decarboxylases in the biodegradation of plant biomass (see 3.1.1.3) and their crucial contribution to the organoleptic properties of fermented foods (*e.g.* wine or beer)^{80,81} being well recognized long ago, only little was known about their catalytic mechanism up until the beginning of the 21st century. Hashidoko and Tahara reported stereospecific protonation during decarboxylation of 4-hydroxycinnamic acids by a decarboxylase from *Klebsiella oxytoca*, exclusively yielding products with the newly incorporated proton in (*E*)-orientation. They further suggested a mechanism *via* a quinone methide intermediate, due to the mandatory presence of a *para*-hydroxy group.⁸²

In 2010, Mancheño and coworkers published a crystal structure of *Lactobacillus plantarum* PAD (*Lp*PAD) (PDB: 2W2A) together with mutational and computational analyses in order to further resolve the catalytic mechanism of cofactor-free PADs.⁸³ The core of the structure was composed of a β -sandwich consisting of two four-stranded anti-parallel β -sheets with an additional, sharply bent β 9 strand interacting with both sheets. Additionally, α -helices were found at both *C*- and *N*-terminal ends and between β 8 and β 9 (3 helices). Particularly, this flattened β -barrel core structure is highly conserved, both in sequence and

structure, amongst related PAD enzymes, whereas the C- and N-terminal regions, as well as helical and loop regions are more variable (Figure 5).⁸³

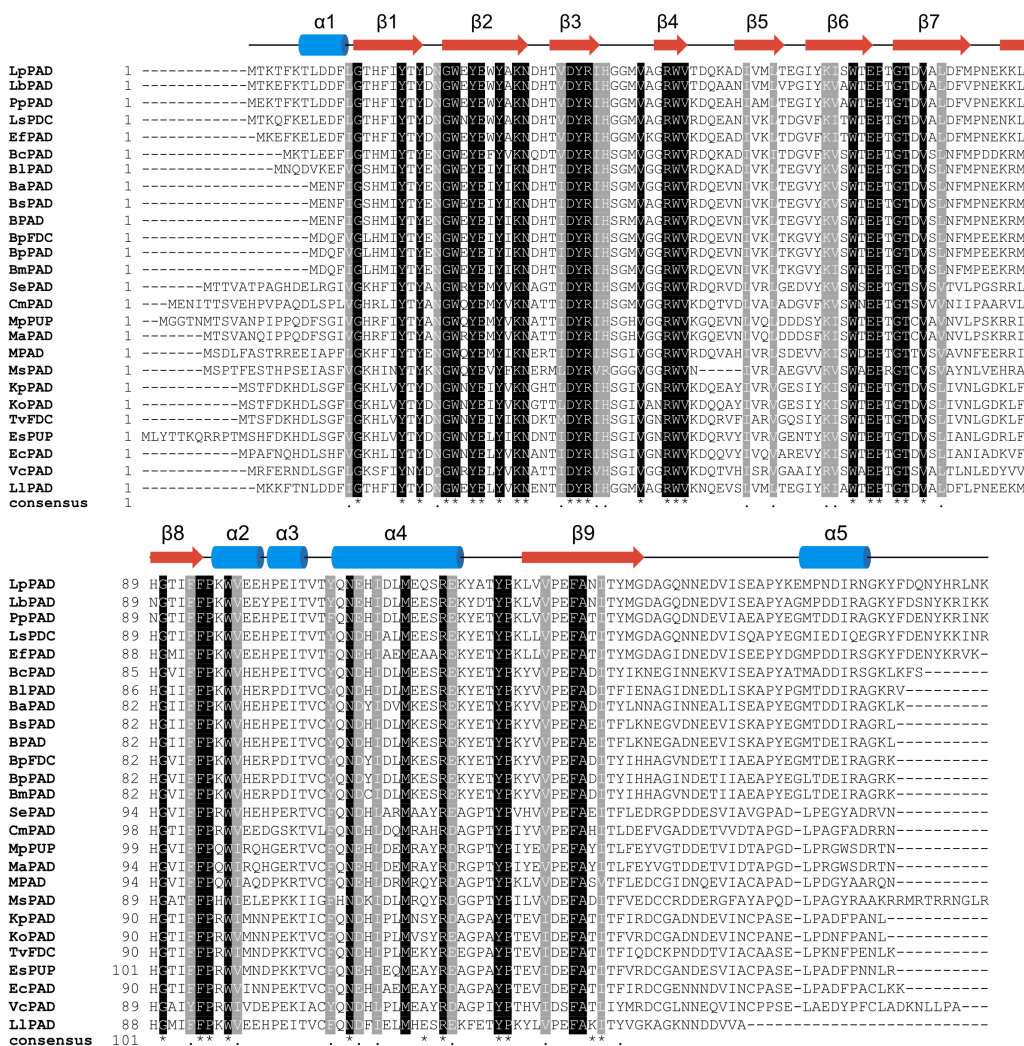
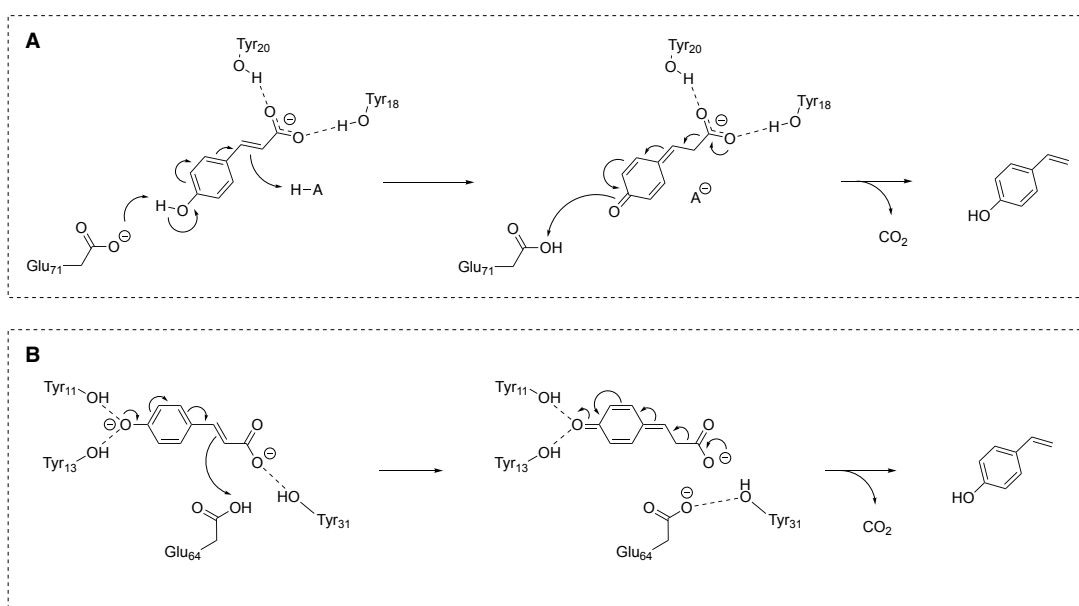


Figure 5. Multiple sequence alignment of *LpPAD* and other phenolic acid decarboxylases. Secondary structure elements of *LpPAD* are highlighted as blue barrels (α -helices) or red arrows (β -sheets). *LpPAD*: *Lactobacillus plantarum* PAD (F9ULL2), *LbPAD*: *Lactobacillus brevis* PAD (Q03TU3), *PpPAD*: *Pediococcus pentosaceus* PAD (Q9F3X2), *LsPDC*: *Lactobacillus sakei* *p*-coumaric acid decarboxylase (Q38UX6), *EfPAD*: *Enterococcus faecium* PAD (Q3Y2T7), *BcPAD*: *Bacillus coagulans* PAD (G2TIS4), *BIPAD*: *Bacillus licheniformis* PAD (Q65FC9), *BaPAD*: *Bacillus velezensis* (formerly *B. amyloliquefaciens*) PAD (A7Z928), *BsPAD*: *Bacillus subtilis* PAD (O07006), *BPAD*: *Bacillus* sp. PAD (Q8KNX7), *BpFDC*: *Bacillus pumilus* ferulate decarboxylase (Q45361), *BpPAD*: *Bacillus pumilus* PAD (A8FAY2), *BmPAD*: *Bacillus mesentericus* PAD (Q9EXR7), *SePAD*: *Saccharopolyspora erythraea* PAD (A4FMV0), *CmPAD*: *Clavibacter michiganensis* PAD (A5CN25), *MpPUP*: *Mycobacterium paratuberculosis* putative uncharacterized protein (Q743A0), *MaPAD*: *Mycobacterium avium* PAD (A0A6B9BIN7), *MPAD*: *Methylobacterium* sp. PAD (B0UN01), *MsPAD*: *Methylocella silvestris* PAD (B8EIN6), *KpPAD*: *Klebsiella pneumoniae* PAD (A6TFA1), *KoPAD*: *Klebsiella oxytoca* PAD (A6BMM9), *TvFDC*: *Trichomonas vaginalis* FDC (A2F476), *EsPUP*: *Cronobacter sakazakii* (formerly *Enterobacter sakazakii*) PUP (A7MLR8), *EcPAD*: *Pectobacterium carotovorum* (formerly *Erwinia carotovora*) PAD (Q6DB32), *VcPAD*: *Vibrio cholerae* PAD (Q9KPX2), *LIPAD*: *Lactococcus lactis* PAD (A2RN76). The alignment was created with Clustal Omega⁸⁴ and visualized with BoxShade. Identical residues are highlighted in black, similar residues in gray. Figure adapted from Rodríguez *et al.*⁸³

This lipocalin-like fold was also reported for PADs from *B. subtilis* and *B. pumilus*.^{85,86} Whereas the interior of the barrel structure was found to be mostly composed of hydrophobic amino acids, some of the rare polar residues turned out to be catalytically relevant (Arg48 and Glu71) in addition to the flexible entrance region (Tyr20). Mutational analysis of these residues revealed a drastic (Y20F and R48Q) or entire (E71S) loss of activity without impairing the overall enzyme structure. By modeling the substrate *p*-coumaric acid into the active site, contacts to the putative active site residues, including interactions of the *para*-hydroxy group with Arg48 and Glu71 and the carboxylate with Tyr18 and Tyr20 were observed. Additionally, the non-polar parts of the substrate were involved in hydrophobic interactions with aliphatic and aromatic residues (Tyr26, Tyr38, Ile40, Val45, Val77, Trp69, Phe94). Based on these results, a mechanism for *Lp*PAD was proposed (Scheme 4).⁸³

Scheme 4. Proposed mechanisms for PAD-like enzymes. (A) Catalytic mechanism proposed for PAD from *L. plantarum* (*Lp*PAD).⁸³ Tyr18 and Tyr20 anchor the substrate *via* the carboxylate while Glu71 deprotonates the *p*-hydroxy group, thereby initiating an electron rearrangement towards a quinone methide intermediate. After release of CO₂, the aromatic structure is restored and the final product 4-vinylphenol is formed. HA: unknown proton donor. (B) Suggested mechanism for *Bs*PAD based on DFT calculations by Sheng *et al.*⁸⁷ The phenolic hydroxy function is coordinated by Tyr11 and Tyr13. Upon electron rearrangement, a quinone methide intermediate is formed and the C=C bond is protonated by Gly64. After cleavage of CO₂, the aromatic system is restored by forming the product 4-vinylphenol.



It was suggested that the mechanistic cascade is initiated by abstraction of the phenolic proton by deprotonated Glu71 (stabilized by Arg48), followed by an

electron rearrangement along the conjugated system. The now nucleophilic alkene carbon accepts a proton from a proton donor to yield a quinone methide intermediate. As there is no potential proton donating residue in close vicinity, the most probable proton donor was suggested to be water. In the second step, the intermediate collapses, thereby releasing CO₂ by restoring the aromatic system.⁸³

After deposition of the initial *B. subtilis* PAD apoenzyme crystal structure (PDB: 2P8G) at 1.36 Å resolution by the Joint Center for Structural Genomics (JCSG) in 2007, Grogan and coworkers published another crystal structure of the same enzyme (Y19A variant) in complex with the substrate *p*-coumaric acid (PDB: 4ALB) (Figure 6).⁸⁶ Following from the position of the substrate in the active site, a reaction mechanism for *BsPAD* was suggested in accordance to the mechanism proposed by Mancheño and coworkers for PAD from *L. plantarum*.⁸³ As already suggested for *LpPAD* and now experimentally evidenced by Frank *et al.* for *BsPAD*, a structural shift of the β 1- β 2 loop is induced upon substrate binding, thereby closing the active site cavity (Figure 6).

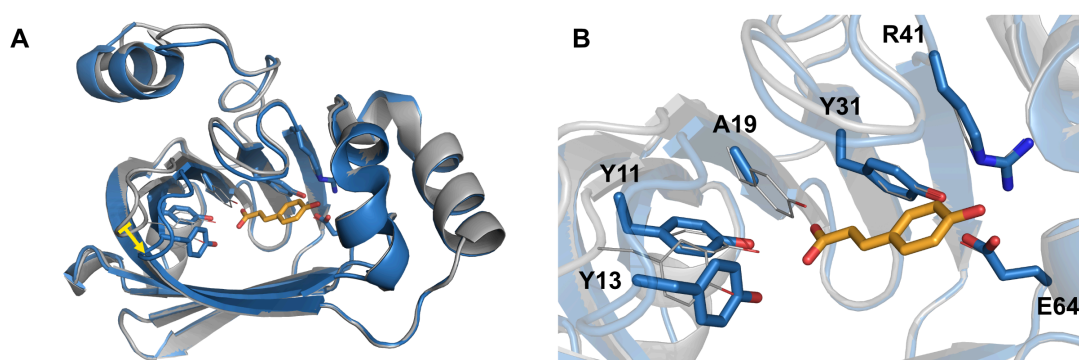


Figure 6. Crystal structures of *B. subtilis* PAD apoenzyme (in gray; PDB: 2P8G) and the Y19A mutant in complex with the substrate (in blue; PDB: 4ALB) *p*-coumaric acid (in orange). Relevant active site residues are highlighted as lines or sticks. (A) Alignment of both structures with the shift of the β 1- β 2 loop highlighted (yellow arrow). (B) Close-up of the active site residues. Structures were visualized with PyMOL.⁸⁸

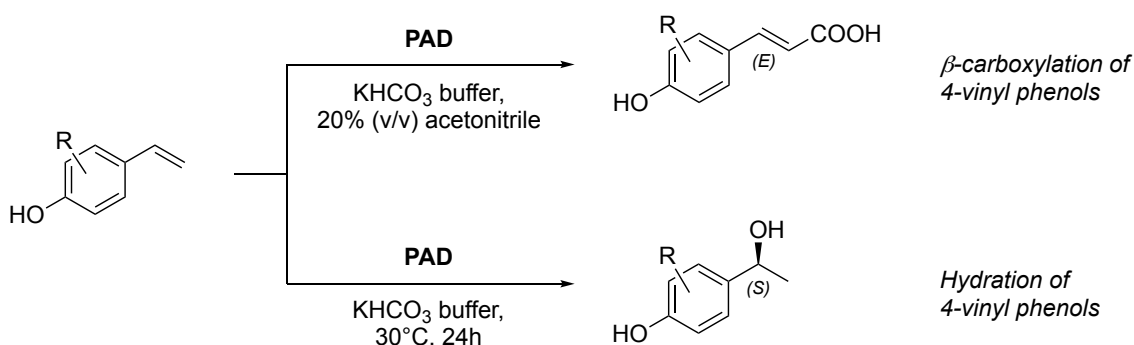
Though, this suggested mechanism was the subject of ongoing discussions and Himo and coworkers suggested a 180° flipped orientation compared to the crystal structure of *BsPAD* based on density functional theory (DFT) calculations. It was proposed, that binding of the phenolic hydroxy group by Tyr11 and Tyr13 instead of Glu64 is more likely from an energetic point of view. Furthermore, Glu64 would be appropriately positioned to act as a proton donor at the C=C bond (Scheme

4).⁸⁷ Such a binding pattern was also suggested for the β -carboxylation and hydration activity of PAD^{89–91} and would be in agreement with other enzymes with related *para*-hydroxylated aromatic substrates, such as vanillyl alcohol oxidase (VAO)⁹² or hydroxycinnamoyl-CoA hydratase lyase (HCHL),⁹³ where the hydroxy function is engaged by two conserved tyrosine residues.

3.1.2.2 Alternative reactivities of PAD

While the ability of benzoic acid (de)carboxylases to directly carboxylate aromatic systems was rather well-studied, a related activity of phenolic acid decarboxylases was unknown. Wuensch *et al.* were the first to report the reverse carboxylation activity of these enzymes exclusively occurring at C- β of the alkene group of 4-vinyl phenols, yielding the corresponding (*E*)-configured *p*-hydroxycinnamic acids (Scheme 5).⁹⁴

Scheme 5. Hydration and β -carboxylation of 4-vinyl phenols catalyzed by PADs. Scheme adapted from Kourist *et al.*⁹⁵



Initially, this reverse reactivity was confirmed for phenolic acid decarboxylases from *Bacillus amyloliquefaciens* (PAD_Ba) and *Lactobacillus plantarum* (PAD_Lp), which were able to catalyze formation of the carboxylated products in rather low conversion (2–30 %). In order to revert the energetically favored and quasi-irreversible decarboxylation reaction, use of a highly concentrated bicarbonate buffer (3 M) was essential to provide an excess of CO_2 . Quite remarkably, carboxylation proceeded regioselectively at C- β of the alkene moiety leaving all aromatic substituents unchanged, which renders this reactivity unique without a known chemical complement.⁹⁴ In a later work, Wuensch *et al.* provided an in-depth study of the reaction scope and mechanism. They identified related decarboxylases from *Mycobacterium colombiense* (PAD_Mc), *Lactococcus lactis*

(PAD_LI), *Methylobacterium* sp. (PAD_Ms), *Pantoea* sp. (PAD_Ps), and *Enterobacter* sp. (FDC_Es) by similarity search (40-80 %) using PAD_Lp as a template. While most of the studied enzymes tolerated a broad variety of substituents on the aromatic ring (e.g. alkyl-, alkoxy or halide groups), substitutions on the alkene proved impeding. Furthermore, presence of a *para*-hydroxy group as a mechanistic prerequisite was also confirmed for the carboxylating activity of PAD.⁸⁹

Another interesting promiscuous activity of phenolic acid decarboxylases, the asymmetric hydration of 4-vinyl phenols, was also discovered by Wuensch *et al.* as a side-reactivity observed during the study of β -carboxylation (Scheme 5).⁹⁰ Interestingly, the extent of hydration surpassed carboxylation in most of the studied cases. It was proposed that the formed quinone methide intermediate acts as a Michael acceptor during nucleophilic 1,6-addition of water. Initially, as during β -carboxylation, a quinone methide oxyanion intermediate is formed by nucleophilic addition of bicarbonate. Yet, this intermediate collapses upon front-side nucleophilic attack by a highly coordinated water molecule, thus explaining the observed (S)-selectivity.⁹⁰

3.1.3 Recent applications

3.1.3.1 Fermentative production of *p*-hydroxystyrenes from glucose

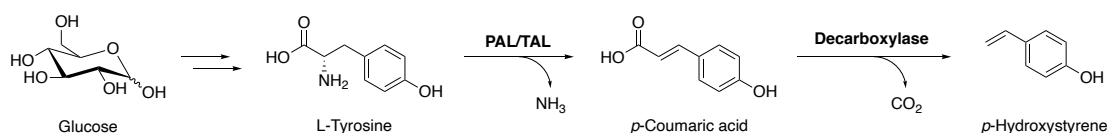
The ability of higher plants to produce a great variety of aromatic compounds by essentially starting from glucose is appealing. By taking advantage of the shikimate pathway, which is fed by intermediates of glycolysis and the pentose phosphate pathway, plants are capable of producing aromatic amino acids, which are further utilized in the phenylpropanoid pathway to form diverse aromatic compounds.⁹⁶

In this context, formation of the industrially relevant polymer precursor *p*-hydroxystyrene from glucose was envisioned by utilizing fermentatively produced L-tyrosine together with phenylalanine/tyrosine ammonia lyase (PAL/TAL) catalyzed deamination and decarboxylation of *p*-coumaric acid in a single host (Table 2).

In an initial approach, Sariaslani and coworkers used *E. coli* strain NST 74, which was designed for phenylalanine production, to heterologously co-express fungal

PAL/TAL from *Rhodotorula glutinis* and different bacterial PDCs from *B. subtilis*, *P. fluorescens* and *L. plantarum*.⁹⁷ Under optimized conditions, 0.4 g L⁻¹ of *para*-hydroxystyrene were obtained from glucose under phosphate limiting conditions, which is close to the upper limit causing cell death in *E. coli* (0.6 g L⁻¹) (Table 2, entry 1). It has to be noted, that due to the fact that the *E. coli* strain NST 74 was meant for production of phenylalanine, considerable amounts (0.5 g L⁻¹) accumulated during fermentation and were further converted into cinnamic acid by PAL/TAL, which accumulated up to 1.2 g L⁻¹. The predominant production of L-phenylalanine over L-tyrosine, together with the considerable toxicity of acetate formed during the fermentation and of the final product *p*-hydroxystyrene, were the major bottlenecks of the process. In this regard it was argued, that *in-situ* product removal strategies could improve the product titer.⁹⁷ This was confirmed by using the same recombinant *E. coli* strain as optimized before in an extractive fermentation setup using 2-undecanone as second phase. In that approach, the product titer could be increased almost 5-fold to 1.8 g L⁻¹ (Table 2, entry 2).⁹⁸

Table 2. Reported systems for the fermentative production of *p*-hydroxystyrene from glucose via deamination of L-tyrosine and decarboxylation of *p*-coumaric acid. PAL/TAL: phenylalanine/tyrosine ammonia lyase, PDC: phenolic acid decarboxylase.



Entry	Host	PAL/TAL	Decarboxylase	Product titer	Ref.
1	<i>E. coli</i> NST 74	PAL/TAL (<i>R. glutinis</i>)	PDC (<i>L. plantarum</i>)	0.4 g L ⁻¹ (H ₂ O)	97
2	<i>E. coli</i> NST 74	PAL/TAL (<i>R. glutinis</i>)	PDC (<i>L. plantarum</i>)	1.8 g L ⁻¹ (2-undecanone)	98
3	<i>P. putida</i> S12	PAL/TAL (<i>R. toruloides</i>)	PDC (<i>L. plantarum</i>)	17.6 g L ⁻¹ (1-decanol)	99

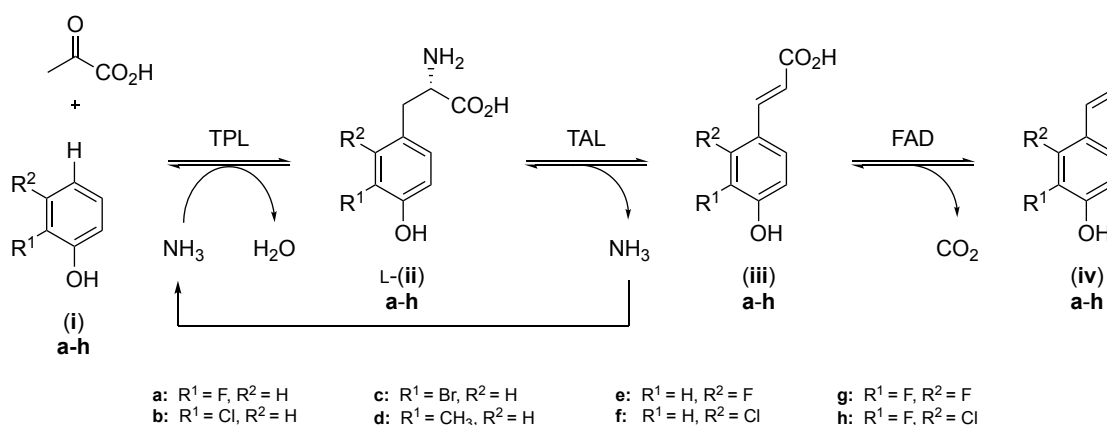
To avert several problems met when *E. coli* was used as production host, Verhoef *et al.* implemented a comparable setup using solvent tolerant *Pseudomonas putida* S12 strains.⁹⁹ Both genes encoding PAL/TAL from *Rhodospiridium toruloides* and PDC from *L. plantarum* were introduced and heterologously

expressed in the host. As the studied strains were originally optimized for production of *p*-coumarate and phenol, both strains exhibited elevated metabolic flux towards L-tyrosine. By additionally knocking out a *p*-coumarate degradative pathway by inactivating the *fcs* gene encoding feruloyl-CoA synthetase, product yield was further optimized. By using a biphasic fed-batch fermentation, the product yield could be increased to 17.6 g L⁻¹ (Table 2, entry 3). It was argued, that due to the inherent solvent tolerance of *P. putida* S12, 1-decanol could be used for *in-situ* product removal, which is more efficient for extraction of *p*-hydroxystyrene and thus the product titer could be increased. Further it was suggested, that genomic integration of the expression construct would result in a more stable expression host and process.⁹⁹

3.1.3.2 Enzymatic and chemo-enzymatic cascades

Similarly to previously reported fermentative approaches,^{97–99} Busto *et al.* designed a three-step *in-vitro* enzymatic cascade, which enabled the formation of 4-vinyl phenols from unprotected phenols *via* C-C bond formation, deamination and decarboxylation (Scheme 6).¹⁰⁰

Scheme 6. Three-step enzymatic cascade for the formation of 4-vinyl phenols (**iv**) from phenols (**i**) *via* C-C bond formation, deamination and decarboxylation. TPL: tyrosine phenol lyase, TAL: tyrosine ammonia lyase, FAD: ferulic acid decarboxylase. Scheme adapted from Busto *et al.*¹⁰⁰

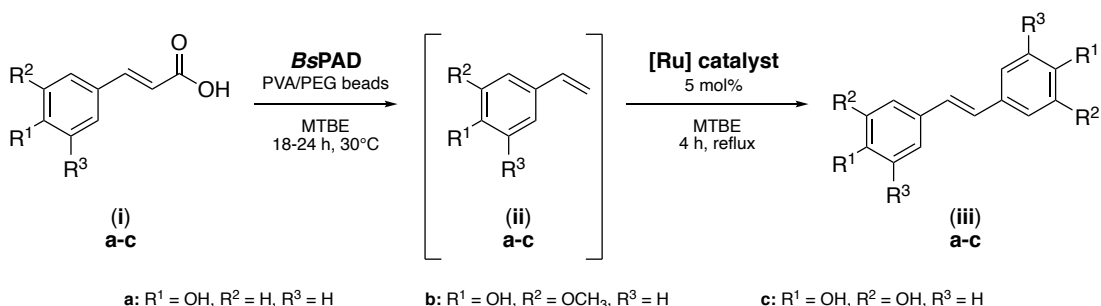


Tyrosine phenol lyase (TPL) from *Citrobacter freundii* catalyzed coupling of the respective phenol derivative (Scheme 6, compound **i**) to pyruvate and the resulting tyrosine derivative (Scheme 6, compound **ii**) was consequently deaminated by tyrosine ammonia lyase (TAL) from *Rhodobacter sphaeroides*. Finally, the formed 4-hydroxycinnamic acids (Scheme 6, compound **iii**) were

decarboxylated by ferulic acid decarboxylase (FAD) from *Enterobacter* sp. to yield the respective *para*-vinylated phenols (Scheme 6, compound **iv**). Remarkably, the whole cascade only required pyruvate as stoichiometric reagent, as ammonia released during deamination was directly recycled and used for the TPL-catalyzed reaction. The applicability was demonstrated by transforming several 2- (**a-d**) and 3- (**e-f**) monosubstituted and 2,3-disubstituted (**g-h**) phenols (Scheme 6). Over 99 % conversion was reached within 8-24 h and the desired *para*-vinylated phenols were isolated in 65-83 % yield. Undesirable *ortho*-vinylation or polymerization of the reactive vinyl phenols was not observed under the employed reaction conditions.¹⁰⁰

Previously, Gómez Baraibar *et al.* combined PAD-catalyzed decarboxylation and a Ru-catalyzed metathesis which gave rise to bio-based resveratrol analogs, recognized for their antioxidant properties, from *p*-hydroxycinnamic acids (Scheme 7). Encapsulation of the biocatalyst in a polyvinyl alcohol/polyethylene glycol cryogel provided an aqueous microenvironment for enzymatic decarboxylation, while the metal-catalyzed reaction step could proceed in pure organic solvent. This spatial separation allowed for simple separation of the enzyme, followed by a drying step (MgSO₄) before metathesis, thus avoiding intermediary work-up and isolation of the highly reactive 4-vinyl phenol intermediates (Scheme 7, compound **ii**).¹⁰¹

Scheme 7. Two-step one-pot chemo-enzymatic cascade for the synthesis of resveratrol derivatives (**iii**) *via* enzymatic decarboxylation of 4-hydroxycinnamic acids (**i**) and [Ru]-catalyzed metathesis of 4-vinyl phenols (**ii**). BsPAD: phenolic acid decarboxylase from *Bacillus subtilis*, PVA/PEG: polyvinyl alcohol/polyethylene glycol cryogel, MTBE: methyl *tert*-butyl ether, [Ru]: Grubbs II or Grubbs-Hoveyda II catalyst. Scheme adapted from Kourist *et al.*⁹⁵

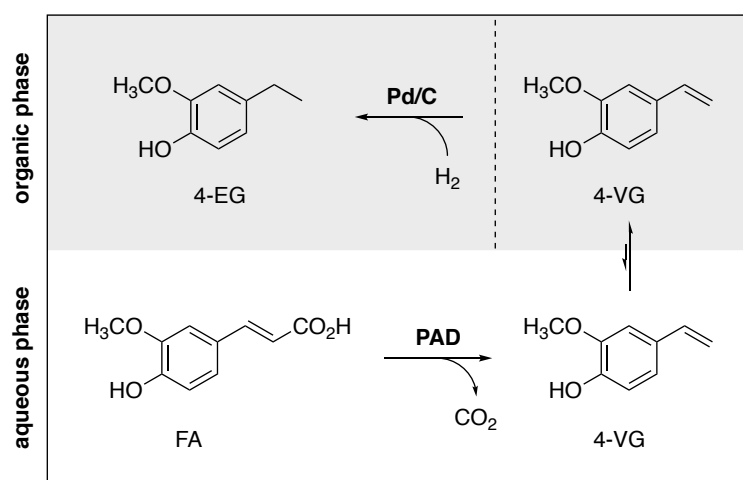


The synthetic usability of this one-pot cascade was demonstrated on a 100 mg scale (3 g L⁻¹ substrate) and 36-90 % of the desired products (compounds **iii a-**

c) could be isolated.¹⁰¹ In this vein, and inspired by the satisfactory activity of BsPAD in deep eutectic solvents (DESs),¹⁰² Ríos-Lombardía *et al.* aimed at the implementation of a similar one-pot cascade in non-conventional solvents. Yet, the promising activity of Grubbs II catalyst in DES (ChChI:Gly 1:2) when using styrene as a model substrate was unmet with all tested PAD-derived vinyl phenols.¹⁰³

Pesci *et al.* developed a chemo-enzymatic cascade process for the synthesis of the aroma compound 4-ethylguaiacol. This was accomplished by combining decarboxylation of ferulic acid by PAD from *Mycobacterium colombiense* (McPAD) and chemical hydrogenation of the resulting alkene *via* palladium on charcoal (Pd/C). Implementation of a two-liquid phase system for *in-situ* product removal turned out to be essential to both overcome inhibition and enzyme deactivation caused by the decarboxylation product and provide a suitable environment for chemical hydrogenation (Scheme 8).¹⁰⁴

Scheme 8. Two-phase system for sequential PAD-catalyzed decarboxylation of ferulic acid, followed by hydrogenation *via* Pd/C in organic solvent. The dashed line represents transfer of the organic phase into a second reactor for hydrogenation. 4-EG: 4-ethylguaiacol, 4-VG: 4-vinylguaiacol, FA: ferulic acid, PAD: phenolic acid decarboxylase. Scheme adapted from Pesci *et al.*¹⁰⁴

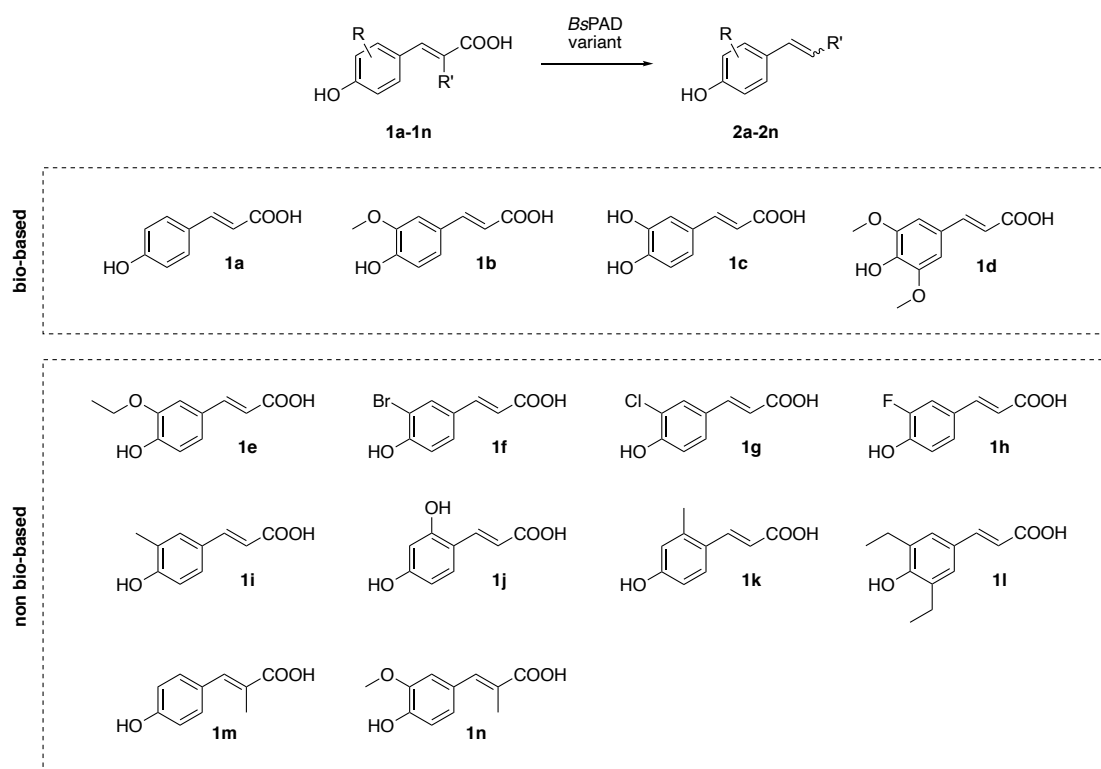


The overall process was further optimized by conducting the hydrogenation step in a separate reactor, which allowed removal and recycling of the organic solvent *via* distillation and led to accumulation of the final product 4-ethylguaiacol.¹⁰⁴

3.2 Aim of the project

Within this work, phenolic acid decarboxylase from *Bacillus subtilis* (BsPAD) and a series of enzyme variants thereof (I85A, I85V, I85L) were studied. The main interest was to investigate the underlying non-mechanistic determinants for substrate acceptance and how use of sterically and electronically differing substrates is reflected in the resulting reaction rate. For this purpose, and besides the well-known bio-based substrates of PAD (**1a-d**), a series of “non-natural” substrates (**1e-n**) was employed to elucidate the impact of subtle alterations of the substrate structure onto enzymatic conversion (Scheme 9). All synthetic substrates were kindly provided by Dr. Christoph Winkler and Prof. Wolfgang Kroutil (Institute of Chemistry, Organic and Bioorganic Chemistry, University of Graz).

Scheme 9. Overview on different substrates studied in BsPAD-catalyzed decarboxylation. Besides bio-based substrates (**1a-d**), various synthetic substrates (**1e-n**) were investigated.



To further benefit from the robustness and high activity of BsPAD, common limitations of enzymatic processes had yet to be resolved. The applicability of PAD-catalyzed decarboxylation was thus investigated within the framework of

different collaborations. In cooperation with Dr. Javier González-Sabín (Entrechem, S.L., Oviedo, Spain), use of non-conventional solvents, more specifically deep eutectic solvents (DESs), was studied in order to alleviate the limited substrate solubility in aqueous reaction systems. Following from that, implementation of a chemo-enzymatic tandem reaction in continuous flow was envisioned in collaboration with Assoc. Prof. Heidrun Gruber-Wölfler (Institute of Process and Particle Engineering, TU Graz).

3.3 Results and discussion

The substrate scope of wildtype PAD is relatively narrow not only due to the obligatory presence of a hydroxy moiety in *para*-position to the acid group as a mechanistic prerequisite, but also due to the spatial limitation within the active site. For instance, while ferulic acid **1b** is a typically accepted substrate, the slightly bulkier sinapic acid **1d** is not converted.^{101,105} By iterative saturation mutagenesis of active site residues, Morley *et al.* discovered that the substitution I85A enables PAD from *Bacillus pumilus* (*Bp*PAD) to accept sinapic acid **1d** as a substrate.¹⁰⁵ This was explained by reduction of steric bulk within the active site by exchanging isoleucine to alanine, which would in turn allow accommodation of the slightly larger substrate.

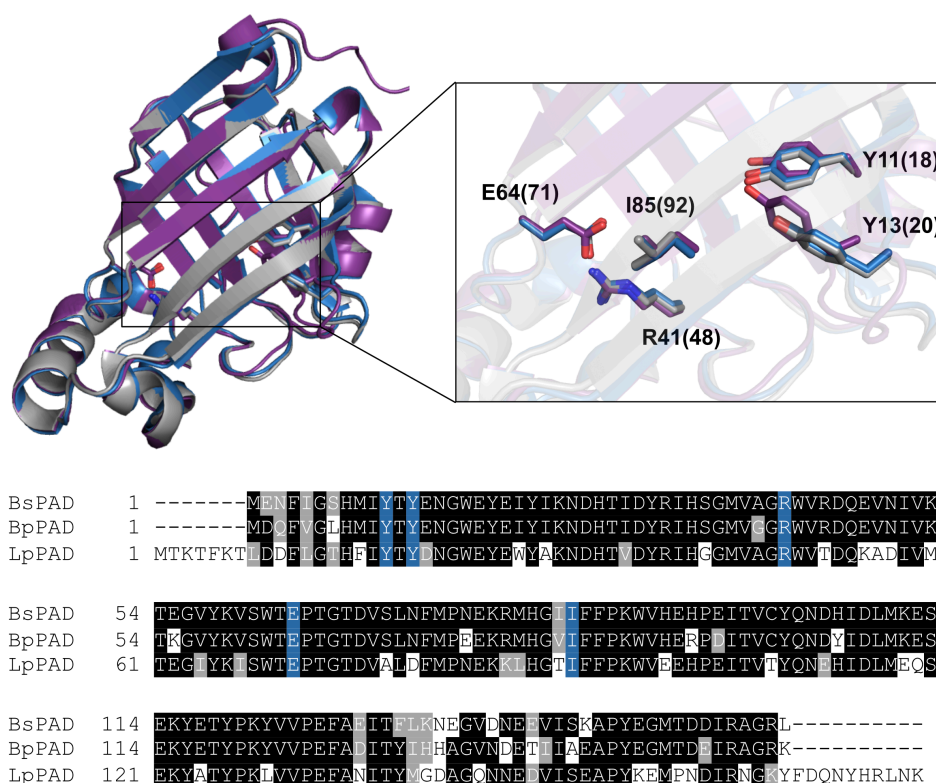


Figure 7. Top: Aligned crystal structures of PAD from *B. subtilis* (*Bs*PAD, PDB: 2P8G) in blue, *B. pumilus* (*Bp*PAD, PDB: 3NAD) in gray and *L. plantarum* (*Lp*PAD, PDB: 2GC9) in purple with relevant active site residues highlighted as sticks. Residue numbers of *Bs*PAD and *Bp*PAD are equivalent and the corresponding residues of *Lp*PAD are given in parentheses. Bottom: Multiple sequence alignment of PAD from *B. subtilis* (*Bs*PAD, UniProtKB accession number O07006), *B. pumilus* (*Bp*PAD, UniProtKB accession number Q45361) and *L. plantarum* (*Lp*PAD, UniProtKB accession number F9ULL2) created with Clustal Omega⁸⁴ and visualized with BoxShade. Identical residues are highlighted in black, similar residues in gray and relevant active site residues in blue. Structures were visualized with PyMOL.⁸⁸

Given the high sequence identity (84 % overall pairwise identity and 100 % pairwise identity of relevant active site residues; Figure 7) and overall structural similarity (Figure 7) of *Bp*PAD and *Bs*PAD, we were curious, if the analogous substitution would also lead to an expanded substrate scope in *Bs*PAD.

In order to prove this hypothesis and to further elucidate the influence of steric bulk within the active site onto substrate acceptance of *Bs*PAD, we characterized *Bs*PAD variants with three different residues at this crucial position (I85A, I85V, I85L) and compared the obtained results to the wildtype enzyme.

3.3.1 Generation of *Bs*PAD I85V and I85L variants

Variants I85V and I85L were created by site-directed mutagenesis and the respective plasmid encoding variant I85A was thankfully received from Álvaro Gómez-Baraibar (Ruhr University Bochum, Bochum, Germany). Primers for QuikChange were designed to introduce the desired point mutations in the *Bs*PAD gene (I85V and I85L) using pET28a[*Bs*PAD_I85A] as a template. All primers had 3' overhangs and a short overlapping region of 18 bp (Appendix C, Table 56). Due to the relatively high GC content in the targeted region, a gradient was chosen for primer annealing (57.0-65.1 °C). After confirming the success of the QuikChange reaction by agarose gel analysis (Figure 8) and followed by *Dpn*I digest, all samples of the respective variant were pooled and used for transformation of *E. coli* TOP10.

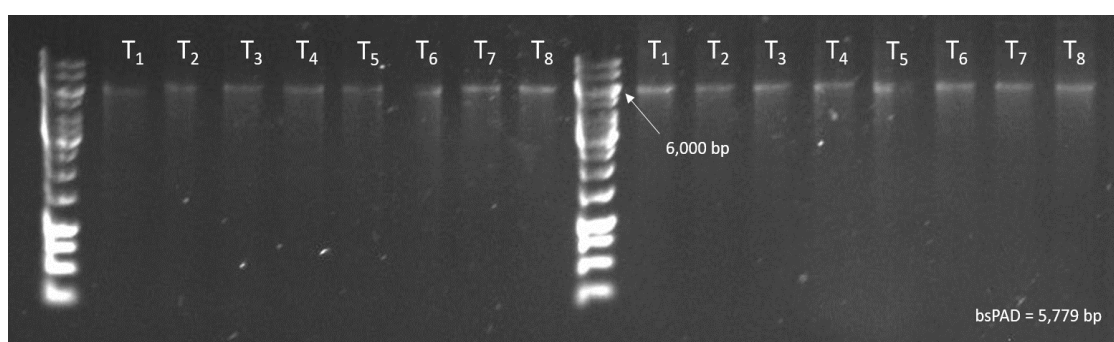


Figure 8. Agarose gel analysis of the QuikChange reaction with temperature gradient ($T_1 = 57\text{ }^{\circ}\text{C}$, $T_2 = 57.7\text{ }^{\circ}\text{C}$, $T_3 = 58.6\text{ }^{\circ}\text{C}$, $T_4 = 59.8\text{ }^{\circ}\text{C}$, $T_5 = 61.1\text{ }^{\circ}\text{C}$, $T_6 = 62.5\text{ }^{\circ}\text{C}$, $T_7 = 63.8\text{ }^{\circ}\text{C}$, $T_8 = 65.1\text{ }^{\circ}\text{C}$); *Bs*PAD I85L (left) and *Bs*PAD I85V (right). GeneRuler 1 kb DNA ladder (Thermo Scientific) was used as standard.

After incorporation of the desired mutations was confirmed by Sanger sequencing, plasmids were used for transformation of *E. coli* BL21(DE3) for means of heterologous enzyme production.

3.3.2 Heterologous production and enzyme purification

The use of pET vectors in combination with *E. coli* BL21(DE3) is a classical yet not flawless approach. The T7 promoter, as used in the pET vector series, is one of the strongest promoters known and widely used for recombinant production of proteins.¹⁰⁶ By using T7-controlled expression constructs, high levels of transcription can be reached in *E. coli* strains, which contain the λ -DE3 lysogen encoding the gene for phage T7 RNA polymerase under the control of the *lacUV5* promoter.^{106,107} Yet, the high potency of IPTG to act as inducer even at low concentrations renders fine-tuning of recombinant protein levels impossible.¹⁰⁶ The burden imposed on *E. coli* by the addition of IPTG might thus result in increased levels of non-functionally expressed proteins (e.g. in inclusion bodies).^{108,109}

Unlike the previously reported conditions,¹⁰¹ heterologous production of BsPAD wildtype and variants was attempted without the use of autoinduction medium. Cultures were incubated at 37°C until cell density (OD₆₀₀) reached 0.5-0.8. Upon induction with IPTG (0.1 mM final concentration), cultivation temperature was reduced to 20°C and cultures were incubated overnight (approx. 18-20 h). By using terrific broth (TB) medium, final cell densities of up to 25 (OD₆₀₀) could be obtained. After cell lysis and separation of cellular debris, high amounts of solubly produced PAD WT and variants could be detected in the respective cleared lysates (see for instance Figure 9, CFE). Furthermore, PAD-containing cell-free extracts were sufficiently stable at 4°C (for at least one week) or in a lyophilized form, when stored at -20°C (at least several months). The availability of lyophilized cell-free extract turned out to be very useful for the preparative-scale isolation of decarboxylation products (see 3.3.4), as due to the highly specific reactivity of PAD no undesirable side-reactivities, as often caused by native *E. coli* enzymes, were observed (judged by HPLC and GC-MS analysis).

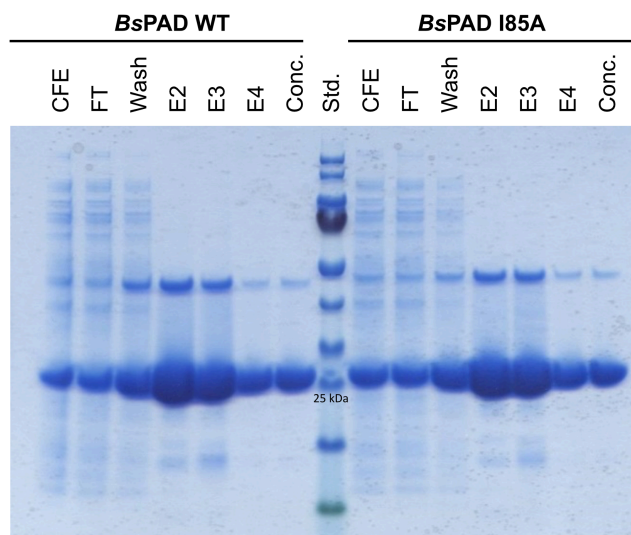


Figure 9. SDS-PAGE analysis of different fractions obtained from Ni-affinity purification of *BsPAD* wildtype (left) and *BsPAD* I85A (right) (MW = 19.1 kDa). CFE: cell-free extract, FT: flow through, E: elution fraction, Conc.: desalted and concentrated combined elution fractions (E1-E7). PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

For determination of specific enzyme activity, PAD variants were purified from the cleared lysates *via* an *N*-terminal His₆-tag by Ni-affinity chromatography. By using standard protocols, such as those provided by the supplier of the column material, satisfactory amounts of purified PAD could be obtained. Interestingly, the final purified protein contained an additional protein (approx. 40-50 kDa in size), however in negligible amounts compared to PAD (Figure 9, conc.). This band could also represent the (insufficiently denaturated) *BsPAD* homodimer usually present in solution.¹¹⁰ Determination of purification yield was hence omitted due to exceeding binding capacity of the column material. For the purification of *BsPAD* I85L and I85V variants, the purification strategy was slightly adapted: The amount of CFE used for purification was adapted to the amount of column material in order to prevent overloading and the column was washed with 20 column volumes (CV) of binding buffer (instead of 10 CV) after applying the sample. In this setup, 30.0 mg (I85L) and 36.7 mg (I85V) purified PAD were obtained from 5.6 mL (14.4 mg mL⁻¹ total protein) and 5.4 mL (14.8 mg mL⁻¹ total protein) CFE, respectively, which accounts for 38 and 46 % of total protein. This further corresponds to a purification yield of 850-945 mg PAD per liter cultivation medium (Table 3). The corresponding SDS gels are shown in Appendix F (Figure S 10-Figure S 11).

Table 3. Yields of BsPAD I85L and I85V obtained by purification via Ni-affinity chromatography from the respective cell free extracts. Purification yield (mg L^{-1} medium) was calculated based on 35 mL (I85L) and 25 mL (I85V) CFE obtained from a 200 mL culture, respectively.

Variant	Purified protein (mg mL^{-1} CFE)	Total protein (mg mL^{-1} CFE)	PAD per total protein	Purification yield (mg L^{-1} medium)
I85L	5.4	14.4	38 %	945
I85V	6.8	14.8	46 %	850

As the achievable percentage of target protein overproduction is mostly reported to be in the range of 50 % when using T7-promoter-based expression constructs,^{106,111} the herein obtained shares of around 40 % are very close to this reference value. This indicates that satisfactory amounts of PAD are produced in a soluble form and that inclusion body formation seems to be insignificant in that case.

The purified proteins could be stored in 20 % glycerol at -20°C (in small aliquots) for over one year without significant loss of specific activity, thus once more underlining the robustness of PAD and the considerably facilitated handling connected to that.

3.3.3 Determination of specific activity of BsPAD and variants

Previously, wildtype PAD from *Bacillus subtilis* was reported to readily convert bio-based hydroxy cinnamic acids like *p*-coumaric acid **1a** ($312 \pm 40 \text{ U mg}^{-1}$), ferulic acid **1b** ($188 \pm 2 \text{ U mg}^{-1}$) and caffeic acid **1c** ($129 \pm 4 \text{ U mg}^{-1}$), but showed only little activity towards the related sinapic acid **1d** ($1.0 \pm 0.3 \text{ U mg}^{-1}$).¹⁰¹ A similar substrate preference was observed for the related enzyme from *Bacillus pumilus*.^{112,113} In search of an enzyme variant able to efficiently convert sinapic acid **1d**, Morley *et al.* saturated several positions within the BpPAD active site and identified the variant I85A. By only introducing a single point mutation, the substrate scope could be expanded towards more sterically demanding substrates.¹⁰⁵

To accomplish an in-depth study of steric and electronic contributions to substrate acceptance, we investigated conversion of a series of bio-based (**1a-d**) and non-biobased (**1e-n**) substrates (Scheme 9) by wildtype BsPAD and variants I85A, I85V and I85L.

At first, activity of *BsPAD* WT towards *p*-coumaric acid **1a** (40 U mg^{-1}), ferulic acid **1b** (80 U mg^{-1}) and caffeic acid **1c** (18 U mg^{-1}) was confirmed (Table 4, entry 1-3 and Figure 10). Yet, absolute values and the clear preference of the wildtype enzyme for ferulic acid **1b** differed from previous reports.¹⁰¹ Likewise, a low activity of 0.36 U mg^{-1} towards sinapic acid **1d** was detected (Table 4, entry 4 and Figure 10). Generally, PAD variants I85V and I85A showed a reduced activity towards the smaller substrates **1a-c**, indicating less successful positioning in the active site caused by increased space thus leading to inefficient conversion (Table 4, entry 1-3 and Figure 10). Contrarily, more sterically demanding substrates such as sinapic acid **1d** or **1e**, a homolog of ferulic acid **1b**, were only scarcely converted by wildtype PAD. Yet, by alternatively using enzyme variants with expanded active site, activity could be increased between eight- (I85V) and 13-fold (I85A) for **1d** or six- (I85V) and 26-fold (I85A) for **1e** (Table 4, entry 4-5 and Figure 10). While the overall activity of variant I85L was low, it was still evident, that larger substrates were not accepted at all, thus reinforcing the tendency observed over all tested variants (Table 4, entry 1-5 and Figure 10).

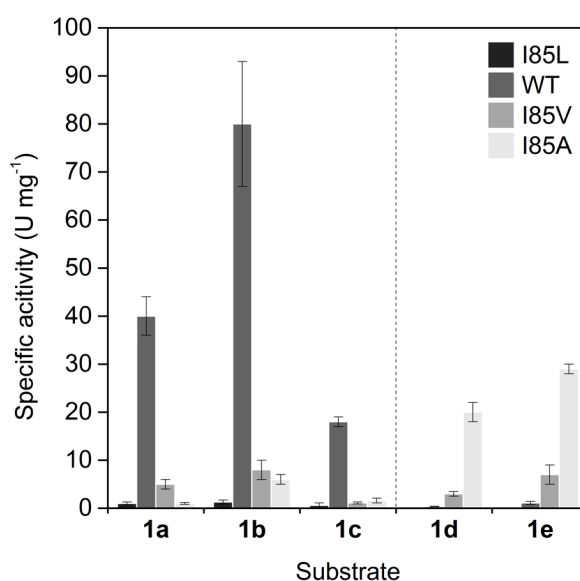


Figure 10. Specific activity (U mg^{-1}) of *BsPAD* wildtype and variants towards phenolic acids **1a-1e** in buffer. Reaction conditions: 10 mM substrate, 5 % (v/v) DMSO, KPi buffer (50 mM, pH 6), 5 μg (WT) or 50 μg (I85A/V/L) purified PAD, 30°C, 600 rpm. Figure taken from Schweiger *et al.*¹⁰²

By introducing halide substituents on the aromatic ring, pronounced activity-enhancing effects were observed. Specific activities of wildtype *BsPAD* were increased by three- to six-fold compared to *p*-coumaric acid **1a** when employing

brominated substrate **1f** (124 U mg^{-1} ; Table 4, entry 6) or the corresponding chlorinated compound **1g** (238 U mg^{-1} ; Table 4, entry 7). Similar effects were observed for PAD I85A, yet with an overall reduced activity due to reasons discussed above (Table 4, entry 6-7). Unexpectedly, when proceeding to the related fluorinated compound **1h**, the trend of activity enhancement was discontinued within this series of substrates for both the wildtype enzyme and I85A variant (Table 4, entry 8). Quite unexpectedly, conversion of the fluorinated substrate **1h** never exceeded 50 % when employing the wildtype enzyme, while such a behavior was not observed with PAD I85A. Also, when using differing substrate concentrations, conversion halted at approximately 4-5 μmol , which corresponds to 40-50 % conversion at 10 mM substrate loading (Figure 11). While this appears like an inhibitory effect caused by the reaction product **2h**, it remains puzzling why the I85A variant is unaffected thereof.

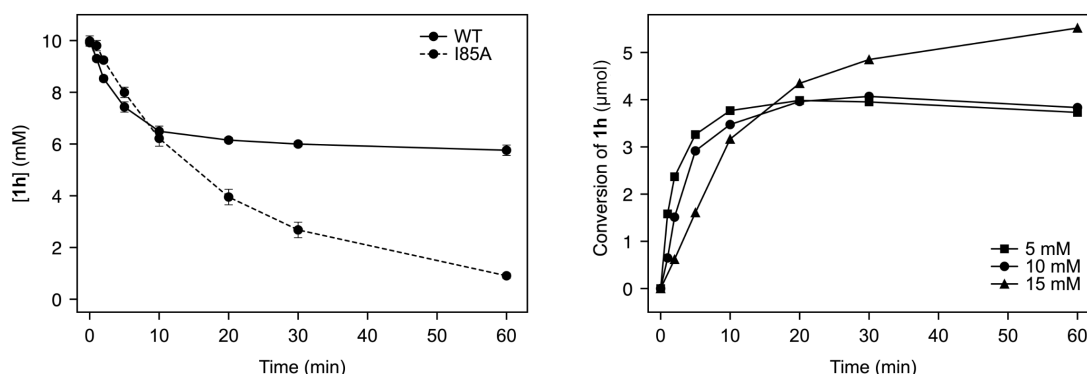
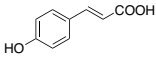
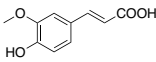
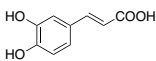
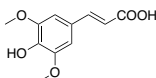
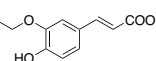
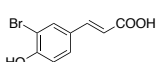
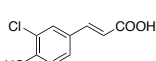
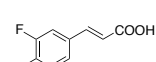
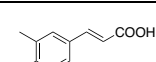
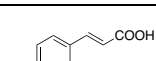
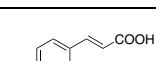
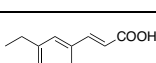
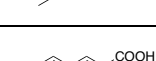
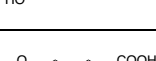


Figure 11. Conversion of substrate **1h** by BsPAD WT and I85A. Left: Decrease of concentration of **1h** (in mM) during conversion by BsPAD WT (5 μg enzyme, solid line) and I85A (50 μg enzyme, dashed line). Right: Conversion of **1h** (in μmol) by BsPAD WT (5 μg enzyme) at 5 mM (square), 10 mM (circle) and 15 mM (triangle) substrate concentration.

Besides halides, a methyl group in *meta* position to the acrylic acid moiety (substrate **1i**) also proved beneficial in terms of PAD activity and led to an enhanced activity of 177 U mg^{-1} and 5 U mg^{-1} for PAD WT and I85A, respectively (Table 4, entry 9). This influence on activity is remarkable, as it cannot be explained *via* electronic contributions, especially when compared to the rate enhancement observed for the strongly electron withdrawing halide substituents (substrate **1f-h**).

Table 4. Specific activity (U mg^{-1}) of BsPAD wildtype and variants I85A, I85V and I85L towards phenolic acids **1a-1n**.^[a]

Entry	Substrate	Specific activity (U mg^{-1})			
		PAD I85L	PAD WT	PAD I85V	PAD I85A
1	 1a	1.0 ± 0.3	40 ± 4	5 ± 1	1.0 ± 0.2
2	 1b	1.3 ± 0.4	80 ± 13	8 ± 2	6 ± 1
3	 1c	0.6 ± 0.5	18 ± 1	1.1 ± 0.2	1.6 ± 0.5
4	 1d	n.c.	0.36 ± 0.03 ^[b]	3.0 ± 0.5	10 ± 2
5	 1e	n.c.	1.1 ± 0.3	7 ± 2	29 ± 1
6	 1f	n.d.	124 ± 16	n.d.	8 ± 1
7	 1g	n.d.	238 ± 70	n.d.	9 ± 2
8	 1h	n.d.	145 ± 20	n.d.	7 ± 1
9	 1i	n.d.	177 ± 20	n.d.	5 ± 1
10	 1j	n.d.	n.c.	n.d.	7 ± 2
11	 1k	n.d.	n.c.	n.d.	n.c.
12	 1l	n.d.	n.a. ^[c]	n.d.	n.a. ^[c,d]
13	 1m	n.d.	0.09 ± 0.03 ^[e]	n.d.	n.c.
14	 1n	n.d.	n.c.	n.d.	n.c.

[a] Reaction conditions: 10 mM substrate, 5 % (v/v) DMSO, KPi buffer (50 mM, pH 6), 5 μg (WT) or 50 μg (I85A/V/L) purified PAD (unless stated otherwise), 30°C, 600 rpm. [b] 500 μg purified PAD WT was used (100-fold amount). [c] Quantification was prevented by substrate insolubility. [d] Substrate conversion was observed visually through decrease in precipitate. [e] 50 μg purified PAD WT was used (10-fold amount). n.a.: not applicable, n.c.: not converted, n.d.: not determined.

When we studied substrates with substituents *ortho* to the acrylic acid group (**1j-k**), it was interesting to see that a hydroxy group at this position (substrate **1j**) was exclusively accepted by BsPAD I85A (7 U mg⁻¹; Table 4, entry 10) and even converted at a higher rate than the corresponding *meta* substituted compound caffeic acid **1c**. Contrarily, substrate **1k**, carrying a methyl group at the same position, was not accepted at all (Table 4, entry 11), while both PAD WT and I85A showed high activity towards the corresponding *ortho* substituted compound **1i**. In view of the enhanced activity of BsPAD I85A towards larger substrates, we envisaged conversion of substrate **1l**, similar to sinapic acid **1d**, by this enzyme variant. Yet this compound proved highly insoluble in aqueous medium which prevented quantification of the activity of PAD I85A towards this substrate. However, when comparing a suspension of **1l** after addition of either PAD WT or PAD I85A, a clear reduction of turbidity was observed in the latter case, thus indicating conversion of **1l** by PAD I85A (Table 4, entry 12). The wildtype enzyme did not show any activity towards substrate **1l**, which is in accordance with previous results. Quite intriguingly, also compound **1m** with a methyl group in α -position to the carboxylic acid was accepted as substrate, albeit only by wildtype BsPAD and at a low rate of 0.09 U mg⁻¹ (Table 4, entry 13). Based on the fact, that ferulic acid **1b** is preferred over *p*-coumaric acid **1a** by BsPAD WT, we assumed, that installation of an additional methoxy group could have beneficial effects on the ability of wildtype PAD to convert this compound (substrate **1n**). Yet, substrate **1n** was neither accepted by wildtype BsPAD nor the I85A variant (Table 4, entry 14). Though, the low tolerance of PAD towards substrates substituted at α -position is in accordance with previous studies of Glueck and coworkers, where substitutions at this position were described as “prohibitive”.⁷⁹ While rates observed in the series of halide substituents might indicate a rate enhancement with increased electronegativity of the substituent (Cl > Br), the trend was discontinued by the corresponding fluorinated substrate **1h**. This could either be the effect of substrate or product toxicity or a less stable accommodation of this substrate in the active site, due to the reduced atomic radius of fluorine compared to the other halides studied. This is also reflected in the fact, that ferulic acid (**1b**) was converted faster than the unsubstituted coumaric acid (**1a**) by all variants studied. From the results obtained in this study, a general comparison of different types of substituents seems infeasible. That is, while halides are

considered electron-withdrawing, alkyl groups rather have an opposite impact and alkoxy groups mainly contribute by mesomeric stabilization, yet all of them can lead to enhanced activity. These contradictory effects of rate acceleration by both electron withdrawing and donating groups suggest, that electron density on the aromatic ring rather plays a minor role and stable accommodation of the substrate, which is in turn determined by substituent size, is predominant. Certainly, several parameters like the electronic structure of the substrate or solvent effects additionally control the activity towards a given compound, yet in the present case of BsPAD it seems that steric effects are the main determinant for substrate acceptance and clearly outweigh electronic effects. This general trend is in accordance with what was previously described by Faber and coworkers.^{89,90,94}

3.3.4 Preparative-scale production of *p*-vinyl phenol derivatives

In order to provide sufficient amounts of 4-vinyl phenol derivatives for investigations of follow-up reactions (*i.e.* for implementation of chemo-enzymatic cascade reactions, see 3.3.6), preparative-scale decarboxylation reactions were executed.

Initially, we considered to use higher substrate concentrations as this would intensify the overall process. However, even when only 20 mM of **1a** were used, the enzymatic reaction was highly inefficient and did not proceed to full conversion, unlike previously reported (Table 5, entry 1).¹⁰¹ This could either be a result of insufficient buffer capacity, *i.e.* unfavorable changes in pH, or toxic effects of the formed vinyl phenol products. While the latter could not be prevented in a monophasic approach, the first issue could be tackled by (continuous) re-adjustment of pH. Yet, this in turn would require elaborate reaction setups or specialized equipment. Alternatively, a biphasic reaction setup for *in-situ* product removal could be used. Gómez Baraibar reported the successful use of isooctane,¹⁰¹ we however aimed to provide the produced vinyl phenols in a low-boiling solvent, which can later on be more carefully evaporated at lower temperatures to prevent polymerization. Such solvents, like methyl *tert*-butyl ether (MTBE), are yet unsuitable for a biphasic bioconversion at 30°C, as they would fully evaporate over time. Considering the high efficiency of BsPAD under the conditions used for characterization (see 3.3.3), implementation of

similar conditions for preparative-scale conversions was an obvious choice. Further, the great specificity of PAD and the high percentage of heterologously produced enzyme allowed for direct use of (lyophilized) cell-free extracts, while still obtaining pure reaction products in acceptable time periods.

In an initial attempt, purified *BsPAD* wildtype was used to decarboxylate *p*-coumaric acid (**1a**, 4.16 mmol) at 20 mM substrate concentration in buffer (Table 5, entry 1). Unlike in analytical reactions (see 3.3.3), addition of DMSO as co-solvent was omitted, and substrates were simply suspended in buffer. It has to be noted that purified enzyme was merely used due to availability but offers no particular advantage compared to crude cell-extract for reasons discussed above. Despite the enzyme amount was scaled according to previous experiments ($5 \mu\text{g mL}^{-1}$), enzyme activity was impaired, most probably by the elevated substrate concentration and thus, additional enzyme had to be added after 3 h to obtain full conversion.

Table 5. Summarized conditions studied in preparative-scale decarboxylation reactions catalyzed by *BsPAD*. The respective substrate was mixed with 50 mM KPi buffer (pH 6) in an Erlenmeyer flask and enzyme was added. *BsPAD* wildtype was used for converting substrate **1a-c** and *BsPAD* I85A for **1d**. The reactions were incubated at 30°C and 120 rpm.

Entry	Substrate	Amount (mmol)	Concentration (mM)	Enzyme formulation	Enzyme amount	Reaction time
1	1a	4.16	20	purified	2x 2 mg ^[a]	4 h
2	1a	4	10	CFE	2 mL	1.5 h
3	1b	2.5	10	CFE	1 mL	o/n
4	1c	2.5	10	CFE	2 mL	o/n
5	1d	2.5	10	CFE	6 mL	o/n
6	1a	5	10	lyophilized	100 mg	o/n

[a] Initially, 2 mg purified *BsPAD* WT were added (according to analytical-scale reactions). As incomplete conversion was detected after 3 h (TLC), additional 2 mg of enzyme were added.

Next, cell-free extracts containing *BsPAD* wildtype (2 mL) were used to decarboxylate **1a** at 10 mM substrate concentration. Samples were taken for HPLC analysis, and quantitative conversion was already detected after 1 h (Table 5, entry 2). Likewise, ferulic acid **1b** and caffeic acid **1c** were successfully

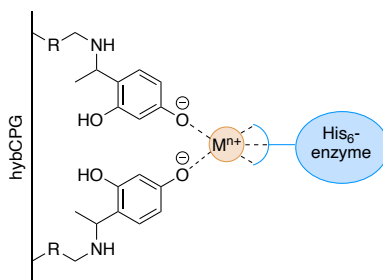
converted by using comparable amounts of CFE (Table 5, entry 3-4). For sinapic acid **1d** (Table 5, entry 5), BsPAD variant I85A was used at higher ratios compared to the reactions described before due to the lower specific activity towards this substrate (see 3.3.3). It has to be noted, that we did not aim at further optimization of CFE amounts or reaction time, as overnight incubation of the reactions and simple TLC analysis proved convenient in our hands. In most cases however, full conversion can most probably be achieved in a much shorter time if desired (*c.f.* Table 5, entry 2). Similarly, enzyme amounts were non-optimized when lyophilized CFE was used, but rather estimated based on the yield of lyophilized CFE from liquid (Table 5, entry 6). The latter approach was clearly preferred, due to the ease of preparation and good storage stability of lyophilized CFE at -20°C, which further avoided on-demand preparation of fresh cell-free extracts for each reaction.

The formed products could be easily extracted from the aqueous reaction mixture with MTBE and the solvent was partially evaporated to obtain a 25-50 mM solution of the vinyl phenol derivatives. This solution could be stored at 4°C for at least several weeks without undesired polymerization. The required amount of products was isolated shortly before further use by evaporation of MTBE under reduced pressure and cooling with an ice bath.¹¹⁴

3.3.5 Immobilization of BsPAD

In order to provide PAD in a heterogeneous format for the implementation of a chemo-enzymatic cascade (see 3.3.6), immobilization on EziG™ particles was envisaged. EziG™ carrier material is composed of controlled porosity glass (CPG), which might be coated with different polymer films in order to tailor the surface polarity (hybrid CPG or hybCPG). Due to the inert, non-swelling and incompressible nature of glass and its interconnecting pore structure, this carrier offers superb flow-through properties and is suitable for batch and continuous applications in aqueous as well as organic medium.^{115,116} Carriers are prepared by coupling 2,4-dihydroxyacetophenone to long-chain amino alkyl moieties on the CPG surface, followed by reduction and complexation of metal ions (commonly Fe³⁺ or Co²⁺). Thus, affinity of the target enzyme's His₆-tag to the complexed metal ions can be exploited for simultaneous purification and immobilization (Scheme 10).¹¹⁶

Scheme 10. Schematic constitution of EziG™ carrier materials. This matrix can be used to bind His₆-tagged enzymes. hybCPG: Hybrid controlled porosity glass (hybCPG). Figure adapted from Engelmark Cassimjee *et al.*¹¹⁶



Based on previous investigations (Gómez Baraibar, unpublished), EziG™ 2 emerged as the most suitable carrier for PAD. This modification of EziG™ possesses a hydrophobic surface conferred by a polymer layer consisting of vinylbenzyl chloride units.^{115,116} Immobilization of BsPAD was done according to the manufacturer's instructions, yet, due to absence of a quick activity assay, loading of BsPAD was analyzed by BCA assay and SDS-PAGE (see 7.3.4 and 7.3.5). Therefore, a defined amount of EziG™ 2 was added to a defined volume of cell-free extract with known total protein content and incubated according to the instructions. Then, the carrier was separated by centrifugation and washed with buffer. The protein content of the recovered supernatant after binding was again determined and the amount of bound protein calculated (Table 6). Additionally, cell-free extract and supernatant were analyzed *via* SDS-PAGE. In an initial attempt, a relatively large amount of PAD-containing CFE (20 mL, 9.4 mg mL⁻¹ total protein) was mixed with 50 mg of EziG™ 2. A total of 15.6 mg of protein was determined to be bound to the carrier, thus resulting in a loading of 31 % (w/w) (Table 6, entry 1). Yet, a decrease of BsPAD in the recovered supernatant was not observed after SDS-PAGE analysis (Appendix F, Figure S 12).

Following from this initial experiment, the strategy was slightly adapted. In order to be able to observe a decrease of BsPAD in the cell lysate on the SDS-Gel, the ratio of carrier to lysate was enhanced and the immobilization procedure was repeated with the resulting supernatant and fresh carrier (Table 6, entry 2-4). In the first immobilization step (5 mL CFE and 100 mg carrier), a loading of 16 % (w/w) was calculated based on the determined protein concentration (Table 6, entry 2).

Table 6. Immobilization of *BsPAD* on EziG™ 2 carrier material. The protein content of the respective solutions used for immobilization was determined before and after binding to the carrier and the bound protein and resulting carrier loading was calculated thereof. SN: Supernatant recovered after separation of loaded carrier.

Entry	Sample	Volume (mL)	Protein concentration (mg mL ⁻¹)	Bound protein (mg)	EziG™ 2 (mg)	Loading
1	CFE	20	9.4	15.6	50.0	31 % (w/w)
	SN		8.6			
2	CFE	5	9.4	16.5	100.6	16% (w/w)
	SN1		6.1			
3	SN1	4	6.1	0.8	101.1	0.8 % (w/w)
	SN2		5.9			
4	SN2	3	5.9	0.6	100.2	0.6 % (w/w)
	SN3		5.7			

While an intensification of the overall immobilization procedure was intended by adapting the carrier to lysate ratio, it seemed, that the reduced excess of *BsPAD* and less efficient mixing rather impaired efficient carrier loading. Quite unexpectedly, during following iterated immobilization steps, only insignificant amounts of protein were bound to the carrier, which resulted in negligible carrier loadings (Table 6, entry 3-4). As high amounts of *BsPAD* were still detectable in all supernatant fractions (Figure 12), reasons for this behavior remained unclear.

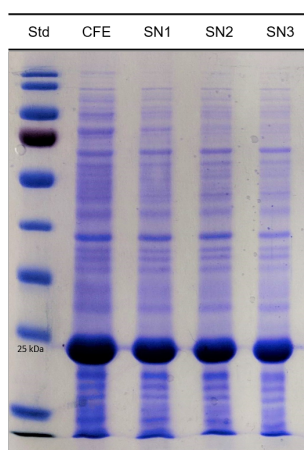


Figure 12. SDS-PAGE analysis of *BsPAD*-containing cell-free extract (CFE) and supernatant fractions (SN1-SN3) obtained after binding to EziG™ 2 carrier material. Results correspond to Table 6, entry 2-4. *BsPAD* (19.1 kDa) appears slightly below the 25 kDa band of the standard (PageRuler Prestained Protein Ladder, Thermo Scientific).

Despite activity of EziG-BsPAD was confirmed under standard conditions (see 7.4.1.1), long-term operational stability of this biocatalyst preparation, especially at a pH < 7 and fluctuations of pH due to decarboxylation, remained questionable.^{115,117}

Due to the aforementioned issues and the unveiled poor predictability and reproducibility of carrier loading, we refocused on entrapment of BsPAD in alginate beads (see 3.3.6). This approach proved similarly straightforward and allowed direct entrapment of defined amounts of PAD-containing lyophilized CFE.¹¹⁴ Prior enrichment or purification steps were not required due to reasons discussed elsewhere (see 3.3.2 or 3.3.6).

3.3.6 Applicability of BsPAD in deep eutectic solvents (DESs)

Despite its robustness and high activity, larger-scale applications of PAD are mostly restricted by the exceptionally low solubility of coumaric acids and their derivatives of only around 10 mM in aqueous media. We opted for the use of deep eutectic solvents (DESs), as they are often considered an environmental benign alternative to ionic liquids. Thereby common limitations of aqueous systems, like substrate solubility and product stability could be tackled, probably due to the extensive hydrogen bonding network present in this type of solvent.¹¹⁸ Despite the rapidly increasing number of successful examples for biocatalysis in DESs, the actual usability for a particular system cannot be taken for granted as enzyme tolerance towards DESs is difficult to predict *a priori*.¹¹⁹

Three different choline chloride (ChCl) based DES (ChCl:glycerol 1:2, ChCl:sorbitol 1:1, ChCl:urea 1:2) were tested in the PAD-catalyzed decarboxylation of caffeic acid **1c** in a 1:1 mixture with water and PAD tolerated all but ChCl:urea. Due to the high viscosity of ChCl:sorbitol however, it was omitted and ChCl:glycerol was used in all further experiments. In order to increase volumetric productivity, substrate loading was increased from 10 to up to 300 mM, where full conversion was reached after 4 h (Figure 13).¹⁰² It has to be noted that despite the apparently high catalyst loadings of 33 % (w/w) relative to the substrate, PAD was employed as freeze-dried cell-free extract, thereby resulting in a considerably lower effective catalyst amount. The relatively high quantities of catalyst needed were thus offset by the ease of preparation and handling.

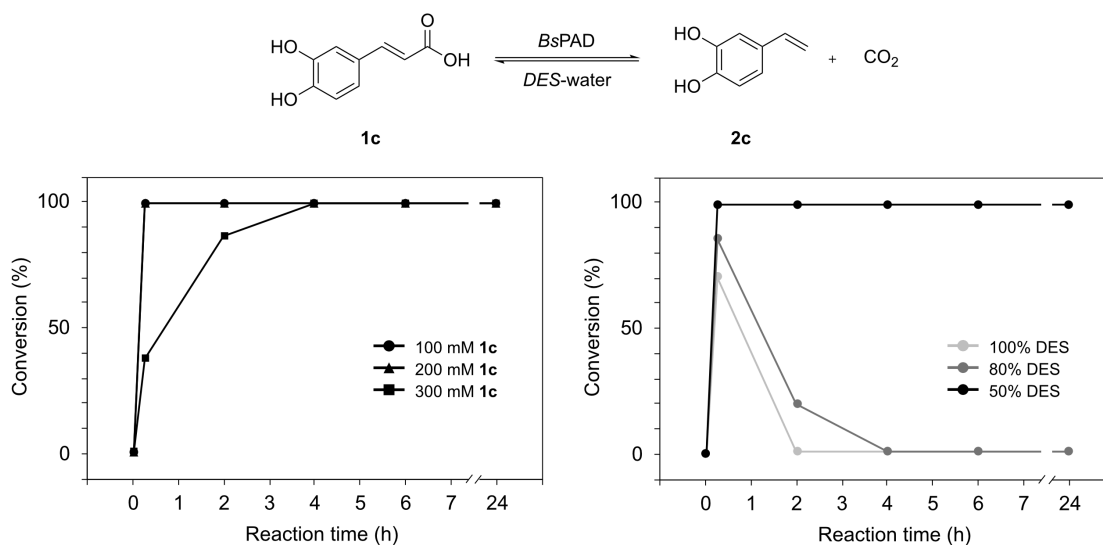
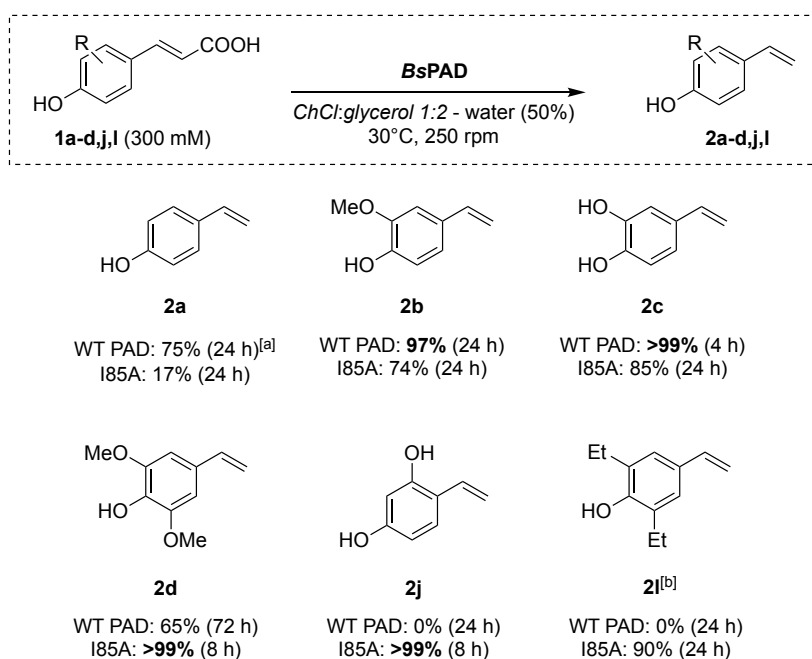


Figure 13. Time course of BsPAD-catalyzed decarboxylation of caffeic acid **1c** (3:1 (w/w) **1c**:PAD_{CFE}) at 30°C and 250 rpm. Decarboxylation of 100-300 mM **1c** in 50 % DES (left) and decarboxylation of 100 mM **1c** in different DES-water mixtures (right).

Performance of PAD was further tested at higher percentages of DES (80 to 100 %) and similar initial activities were observed as for a DES-water ratio of 1:1. Yet, after elongated incubation, reversal of the formed product towards the substrate was monitored with HPLC at DES-contents higher than 50 % (Figure 13).¹⁰² This observation was striking, as PAD was already reported previously to catalyze the reverse carboxylation reaction as well, yet only in presence of high bicarbonate concentrations as external CO₂ source.⁸⁹ Indeed, deep eutectic solvents were reported in the literature to efficiently capture carbon dioxide.¹²⁰ Water contents of at least 50 % proved to be the only approach to effectively release the emerging CO₂ from the reaction mixture, unlike application of heat or vacuum. At this point however, it was yet unclear if the specific characteristics of DES were still retained at a water content of 50 %, as Edler and coworkers previously reported the loss of DES nanostructures of ChCl:urea at over 42 % water content.¹²¹ To verify this for our reaction system, reactivities in 50 % DES-water mixtures, as used in this study, were compared to aqueous solutions of DES components without prior DES formation. Despite PAD still showing some activity in the latter, results obtained in 50 % DES-water mixtures were beyond that, thus clearly demonstrating that – at least in parts – beneficial properties of this neoteric solvent were still retained at 50 % hydration.¹⁰²

The scope of the established reaction setup was tested with selected compounds **1a-d,j,l** at 300 mM substrate concentration and both PAD wildtype and I85A (Scheme 11).

Scheme 11. Substrate scope of BsPAD-catalyzed decarboxylation of **1a-d, j, l** in 50 % aqueous mixtures of DES (ChCl:glycerol 1:2). [a] Conversion of >99 % at 200 mM in 2 h, [b] Performed at 10 mM with 1 % (v/v) DMSO, WT: wildtype.



Most of the tested compounds were efficiently converted by either the wildtype or I85A variant of PAD at the applied substrate concentration, generally following the trend of substrate acceptance as observed in aqueous medium. Only **2l** remained highly insoluble under the applied conditions, thus preventing substrate concentrations higher than 10 mM (in presence of 1 % (v/v) DMSO).

Quite interestingly, while conversion of caffeic acid **1c** by PAD is clearly preferred in DES, the related constitutional isomer **2j** was converted at comparable relative activities in both DES and conventional aqueous medium. Such remarkable differences in activity as observed for **1c** might be the result of differentiated interactions between solvent and substrate. This in turn can directly influence substrate availability for the enzyme and thus result in an altered effective Michaelis constant.^{122–124}

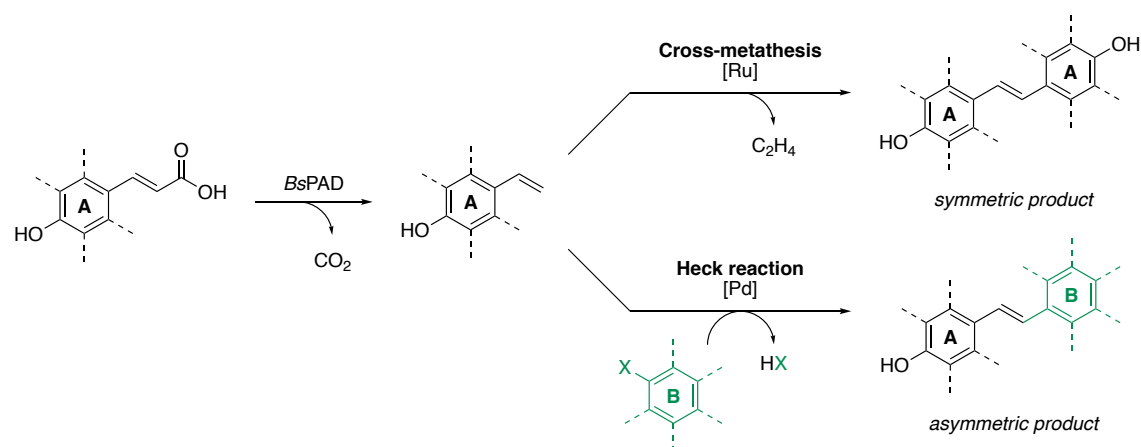
In view of the promising activity of PAD in non-conventional solvents, and previous reports of successful transition-metal-catalyzed reactions in DES,^{118,125} implementation of a chemo-enzymatic cascade was envisioned. Finding a

compatible reaction environment for both chemo- and biocatalyst is a commonly encountered hurdle during implementation of cascade processes.¹²⁶ Often, these issues can be solved by introducing a biphasic system or compartmentalization. In this vein, Gómez Baraibar *et al.* previously reported entrapment of BsPAD in a cryogel matrix. This approach provided a suitable aqueous micro-environment for the enzyme, while the Ru-catalyzed reaction step could be carried out in pure organic solvent.¹⁰¹ Yet such biphasic reaction mixtures can be complex and suffer from difficulties in handling or upscaling. Additionally, the use of, mostly flammable or toxic, organic solvents is often undesired.

Following this previous work, we conceived the synthesis of stilbene derivatives starting from bio-based coumaric acids.¹⁰¹ Due to their well-known antioxidant and health-beneficial properties, stilbenes pose a highly interesting substance class.^{127,128} Yet, the previously reported approach was limited to symmetrically substituted products, inherently determined by the nature of the Ru-catalyzed metathesis reaction. Contrastingly, similar yet asymmetric products can be obtained by using differently substituted aryl halides in a Pd-mediated Heck reaction (Scheme 12).

In that sense and attending to our ultimate goal of implementing a chemo-enzymatic cascade in continuous flow, a suitable heterogeneous Pd⁰ catalyst and immobilized preparation of BsPAD for application in deep eutectic solvent were yet to be identified.

Scheme 12. Conceptual overview of the different coupling products obtained from Ru-catalyzed cross-metathesis and Pd-catalyzed Heck reaction starting from biocatalytically produced 4-vinyl phenols. Dashed bonds represent potential substituents. X: halide.



Initially, both reaction steps were optimized separately in batch and in flow, before they were eventually combined to a reaction cascade. Regarding the biocatalytic decarboxylation, a suitable heterogeneous enzyme formulation needed to be identified first. Originally, it was intended to immobilize BsPAD on EziG™ particles *via* affinity of the enzyme's His-tag to metal ions, which would enable enrichment of the target enzyme from the cell-free lysate concurrently to immobilization (see 3.3.5). This approach was however dismissed due to unclear long-term operational stability of such a preparation,¹¹⁷ especially during decarboxylation (*i.e.* changes in pH) as in the present case. Further considering the high percentage of heterologously produced BsPAD in the cell-free lysates (see 3.3.2), its largely retained activity after lyophilization and purity of obtained reaction products, even when using crude enzyme preparations, renders prior purification steps or enrichment during immobilization dispensable. As an equally straightforward and easy alternative approach, PAD was instead immobilized in 2 % alginate beads (average diameter: 2 mm), by directly using lyophilized PAD-containing cell-free extract.

After confirming activity of immobilized PAD in batch (Table 7, entry 1), heterogeneity of the biocatalyst preparation was investigated by means of a “hot”-filtration test. For this purpose, the alginate beads were filtered off after 45 min, which stopped the reaction. This indicated, that – at least no active – enzyme was leaching from the beads under the applied reaction conditions.¹¹⁴

Next, we tested if the chosen heterogeneous Pd-substituted cerium-tin-oxide catalyst ($\text{Sn}_{0.79}\text{Ce}_{0.2}\text{Pd}_{0.01}\text{O}_{2.5}$) was suitable to catalyze Heck coupling in the applied solvent system. In addition to the DES-buffer mixture as used in the enzymatic reaction step, a 1:1 (v/v) mixture of water and ethanol containing 1.5 eq of each iodobenzene (**3**) and K_2CO_3 was required for this step, thus rendering the resulting mixture a salt solution rather than a deep eutectic solvent.¹²¹ While full conversion was reached within 30 min (Table 7, entry 2), the isolated yield was as low as 35 % due to several potential side reactivities in batch like polymerization of **2a**, or formation of homo-coupling products. Importantly, when both steps were combined in a sequential one-pot reaction, neither remaining **1a** nor components of the CFE negatively influenced the Heck reaction, thus proving compatibility of the individually optimized reactions for continuous flow.¹¹⁴

Table 7. Summarized reaction conditions for the individually optimized reaction steps in batch and flow, namely BsPAD-catalyzed decarboxylation of *p*-coumaric acid **1a** and Pd-catalyzed cross-coupling of the resulting 4-vinyl phenol **2a** with iodobenzene (**3**). In all cases, ChCl:glycerol 1:2 was used as DES and Sn_{0.79}Ce_{0.2}Pd_{0.01}O_{2.5} as heterogeneous Pd catalyst.

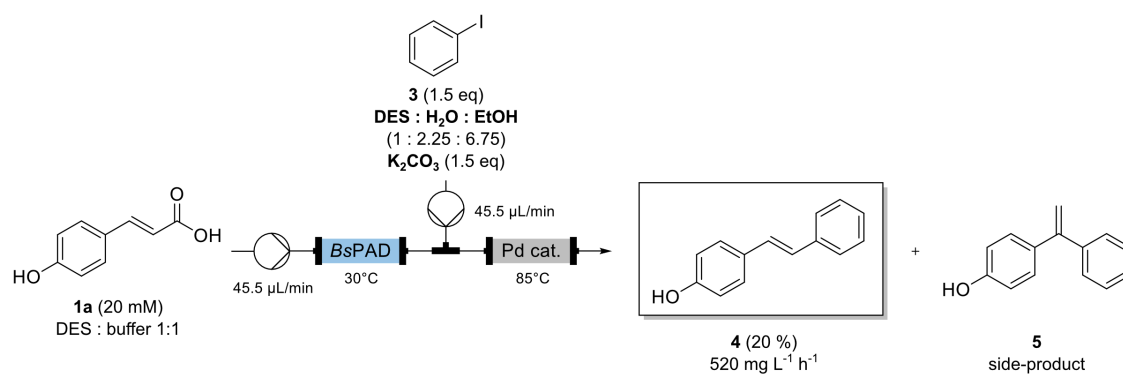
Entry	Reaction step	Setup	Reaction conditions	Catalyst amount	Conversion ^[a]
1	Enzymatic	Batch	0.7 mmol 1a in 10 mL DES:buffer 1:1 (v/v), 30°C	38.3 mg BsPAD _{CFE} ^[b]	100 % (6h)
2	Heck coupling	Batch	0.7 mmol 2a , 1.5 eq 3 and K ₂ CO ₃ in 20 mL DES:buffer:EtOH:H ₂ O 1:1:1:1 (v/v/v/v), 85°C	0.05 mol%	100 % (0.5 h)
3	Enzymatic	Flow	20 mM 1a in DES:buffer 1:1 (v/v), 30°C, 45.5 μL min ⁻¹	160 mg BsPAD _{CFE} ^[b] (2x 40 mm × 8 mm i.d., 4 mL total)	100 % (steady-state)
4	Heck coupling	Flow	10 mM 2a , 1.5 eq 3 and K ₂ CO ₃ in DES:buffer:EtOH:H ₂ O 1:1:1:1 (v/v/v/v), 85°C, 91 μL min ⁻¹	2 g Pd-cat. (40 mm × 8 mm i.d.)	90 % (steady-state)

[a] Conversion determined with HPLC. [b] Lyophilized CFE was dissolved at 38.3 mg mL⁻¹ in 50 mM KP_i buffer (pH 6) containing 2 % (w/v) Na-alginate and the mixture was cross-linked in a 2 % (w/v) BaCl₂ solution.

In order to implement the established reactions in flow, both heterogeneous catalysts were packed into stainless steel columns (Table 7, entry 3-4), which could then be individually heated to the optimal reaction temperature in a water bath. After testing each step individually (Table 7, entry 3-4), slight adaptations were made in order to optimize the Heck reaction. To enhance substrate solubility, the solvent mixture of the feed for the second step was slightly adapted to DES:EtOH:H₂O 1:6.75:2.25, thus resulting in a final composition of 30 vol% DES, 25 vol% buffer, 34 vol% EtOH and 11 vol% H₂O after combining the feed with the product stream of the first step. Further, the Pd-catalyst amount was increased to 6 g (120 mm × 8 mm i.d.). In the combined continuous setup, an initial concentration of 20 mM *p*-coumaric acid (**1a**) was chosen to ensure complete substrate dissolution and both pumps were operated at 45.5 μL min⁻¹ (in total: 91 μL min⁻¹) (Scheme 13). The residence time was determined to be

30 min and 45 min for decarboxylation and Heck coupling, respectively. Thorough flushing of the system before operation turned out to be crucial to prevent clogging by denatured, loosely bound enzyme. In this regard, an additional filter was installed before entrance to the second reactor. In this way, the process could be operated over 16 h in a steady-state with full conversion of both **1a** and **2a** and a space-time yield of $4.8 \text{ g L}^{-1} \text{ h}^{-1}$ and $0.52 \text{ g L}^{-1} \text{ h}^{-1}$ for decarboxylation and Heck coupling, respectively (Scheme 13).

Scheme 13. Schematic representation of the established chemo-enzymatic tandem reaction combining BsPAD-catalyzed decarboxylation of *p*-coumaric acid **1a** and Pd-catalyzed cross-coupling of the resulting 4-vinyl phenol **2a** (not shown) with iodobenzene (**3**) to yield (*E*)-4-hydroxy stilbene **4** and the side product 4-(1-phenylvinyl)phenol (**5**) in aqueous DES mixtures.



The obtained yield of the desired (*E*)-4-hydroxy stilbene **4** was however limited to only 20 %. While formation of the side product 4-(1-phenylvinyl)phenol (**5**), due to similar reactivity of both sp^2 vinylic carbon atoms of **2a**, was more apparent in flow than in batch, emergence of biphenyl as a homo-coupling product of iodobenzene (**3**) could be diminished. Additionally, polymerization of **2a** could result in a low yield by formation of undetectable compounds. While polymerization of the formed 4-vinyl phenol **2a** was not observed during sole enzymatic conversions in DES-water mixtures,¹⁰² this might not hold true for the following Heck coupling step at elevated temperatures (85°C) and altered solvent composition, where the characteristic nanostructural network of DES and thus its stabilizing properties are absent due to the high degree of dilution.¹²¹

3.4 Summary

In this work, phenolic acid decarboxylase from *Bacillus subtilis* (BsPAD) and a series of newly created enzyme variants – I85A, I85V and I85L – were studied in order to shed light on the determinants of substrate scope and influences on reaction rate of this enzyme. This was accomplished by investigating the activity of BsPAD and its variants towards several, both bio-based (**1a-d**) and artificial (**1e-n**), substrates (Figure 14).

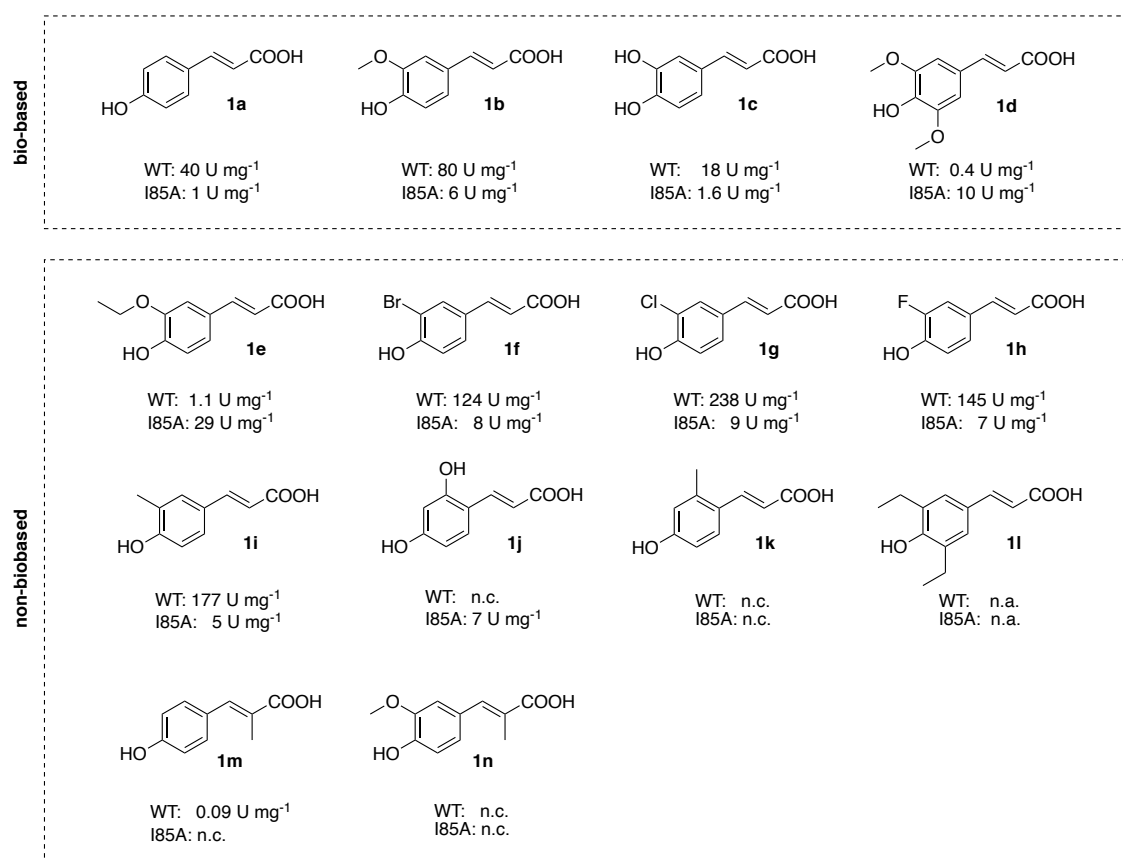


Figure 14. Summary of bio-based (**1a-d**) and non-biobased (**1e-n**) substrates tested in BsPAD-catalyzed decarboxylation. Specific activities (U mg⁻¹) determined for each substrate and BsPAD WT or I85A variant are stated. Experimental details are given in Table 4. n.c.: not converted, n.a.: not applicable.

While BsPAD wildtype largely tolerated small substituents *meta* to the acrylic acid moiety (*i.e.* methoxy (**1b**), hydroxy (**1c**), halides (**1f-h**), methyl (**1i**)), *ortho* substituents (**1j-k**) and multi-substituted substrates (**1d**, **1l**, **1m-n**) were not accepted. In contrast, variant I85A could also accommodate bulkier substrates (**1d-e**, **1l**) and also tolerated *ortho* substituents (**1j**) to some extent. Alkene substitutions were generally not accepted, and only substrate **1m** was converted

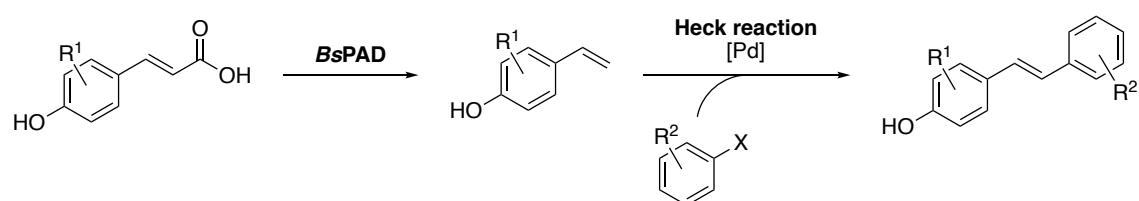
by BsPAD WT, albeit at a very low rate. Overall, steric effects were predominant for the substrates studied, and enzyme activity seemed to be closely connected to the ability of accommodating a given substrate in a productive binding mode. To further exploit the exceptional robustness and high activity of BsPAD, the extremely low solubility of the *para*-hydroxycinnamic acid substrates was tackled by implementing an alternative solvent system for biocatalysis. By using deep eutectic solvents (DES), more specifically choline chloride (ChCl) and glycerol in a molar ratio of 1:2, substrate solubility could be significantly enhanced while maintaining activity of BsPAD. Dilution of pure DES with 50 % water proved crucial to enable efficient release of CO₂ from the reaction mixture, thereby preventing reversal of the desired decarboxylation reaction. In this optimized setup, up to 300 mM of substrate were successfully converted within 24 h. Additionally, we observed an interesting switch of substrate preference depending on the solvent used, as caffeic acid **1c** was clearly preferred in DES. These results highlight the importance of solvent engineering to enable efficient biocatalytic processes but also to influence substrate preference.

Prompted by the promising performance of BsPAD in deep eutectic solvents, we envisaged implementation of a chemo-enzymatic cascade in DES which should combine PAD catalyzed decarboxylation and Pd catalyzed Heck coupling to enable synthesis of asymmetrically substituted stilbene derivatives. By using encapsulated enzyme and a heterogeneous Pd catalyst, a fully integrated two-step continuous flow process could be established. The use of deep eutectic solvents enabled higher substrate concentrations for enzymatic decarboxylation (20 mM) and provided compatible reaction conditions for both reaction steps. Furthermore, isolation of the reactive 4-vinyl phenol intermediates was prevented in this setup. The implemented process was operated over 16 h in a steady-state with full conversion of both substrate **1a** and intermediate **2a**, and a space-time yield of 4.8 g L⁻¹ h⁻¹ and 0.52 g L⁻¹ h⁻¹ for decarboxylation and Heck coupling, respectively. While the overall yield of (*E*)-4-hydroxy stilbene **4** was limited to 20 % due to byproduct formation, the established process nevertheless highlighted the usability of non-conventional solvents, like DES, to overcome incompatibility issues faced in tandem-catalytic processes.

3.5 Outlook

We could recently demonstrate the usability of deep eutectic solvents for the implementation of a chemo-enzymatic tandem reaction in continuous flow (Scheme 14). Yet, the presented reaction is only exemplary and numerous variations of the substrate and aryl halide scaffold are conceivable, thus, in theory, enabling the preparation of numerous different stilbene derivatives. In this respect, recent efforts aimed at optimization of the overall cascade setup.¹²⁹

Scheme 14. General scheme of *BsPAD* catalyzed decarboxylation combined with Pd catalyzed Heck coupling to yield asymmetric stilbene derivatives.



As substrates of *BsPAD* also present bio-based aromatic compounds, the formed 4-vinyl phenols are interesting targets for the production of functionalized bio-based polystyrenes (Scheme 15). In this vein, studies are currently conducted aiming at the efficient preparative decarboxylation of caffeic acid and implementation of a suitable protecting group strategy as preparation for polymerization and characterization of the resulting materials.

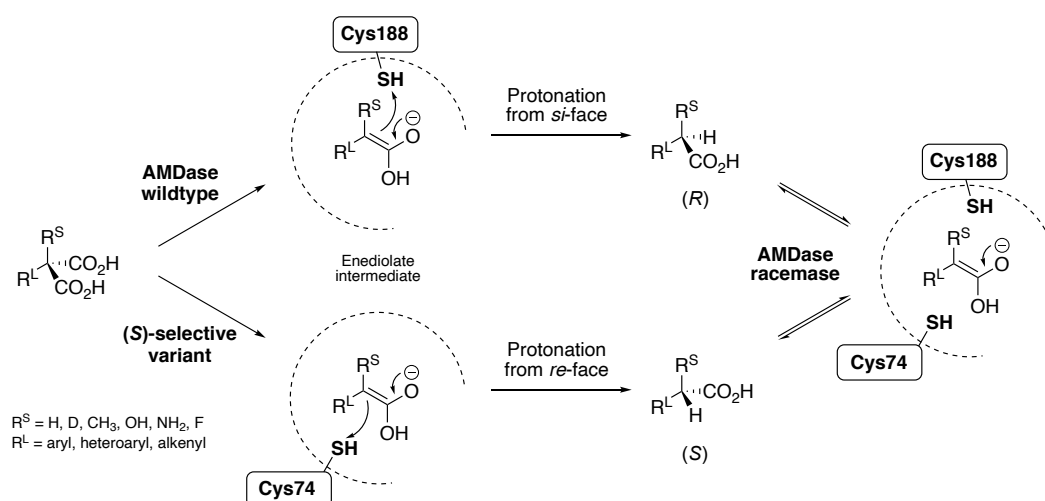
Scheme 15. Preparation of bio-based hydroxy styrene derivatives by PAD catalyzed decarboxylation, followed by protection and polymerization to yield functionalized polystyrenes. PG: protecting group.



4 Arylmalonate decarboxylase from *Bordetella bronchiseptica* (AMDase)

Arylmalonate decarboxylase (AMDase, EC 4.1.1.76), originally isolated from the soil bacterium *Bordetella bronchiseptica*, catalyzes the decarboxylation of aromatic- or alkenylmalonic acids to yield the corresponding optically pure monoacid without the aid of a cofactor. While the wildtype enzyme exhibits strict (*R*)-selectivity, enzyme engineering afforded efficient (*S*)-selective or racemizing enzyme variants (Scheme 16). The native function or the natural substrate of AMDase however remain unknown to date. The unique reactivity and broad substrate tolerance allows for the production of diverse aryl- or alkenylaliphatic carboxylic acids in outstanding optical purity, amongst them several α -arylpropionates with non-steroidal anti-inflammatory activity, the so-called profens.⁴³

Scheme 16. Overview on AMDase catalyzed asymmetric decarboxylation by wildtype (Cys188) and (*S*)-selective variants (Cys74) and AMDase catalyzed racemization (Cys188/Cys74).

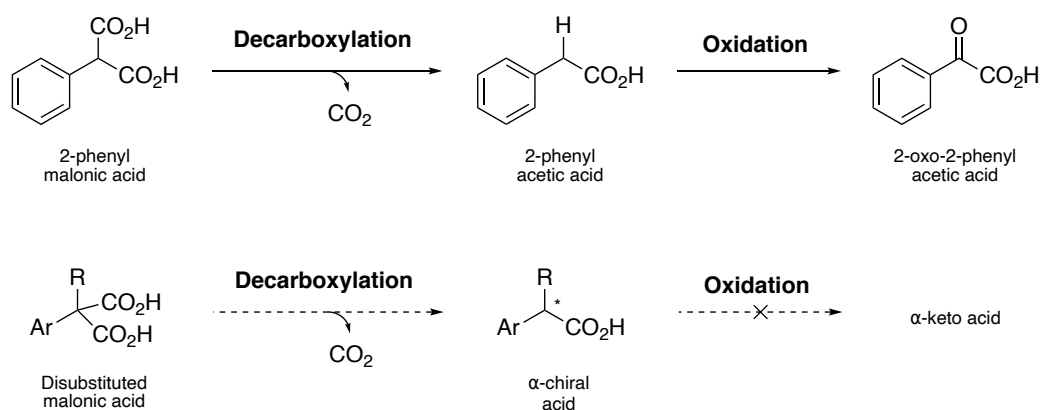


4.1 Theoretical background^a

4.1.1 Discovery of AMDase

Miyamoto *et al.* discovered arylmalonate decarboxylase (AMDase) in the 1990's in a screening to identify enzymes for the generation of chiral molecules from prochiral malonates by enzymatic decarboxylation.¹³⁰ As the decarboxylation of malonyl-ACP is a key step in metabolism, it appeared likely that the bacterial catabolism might possess malonate decarboxylases. The screening followed the assumption that decarboxylation of 2-phenylmalonic acid could represent the initial step in the metabolism, followed by oxidation to yield 2-oxo-2-phenylacetic acid. Accordingly, it was proposed, that analogous decarboxylation of disubstituted malonic acids would lead to α -chiral acids, as further oxidation is not possible (Scheme 17).

Scheme 17. Assumed reaction pathway of 2-phenylmalonic acid converting microorganisms (top) and proposed route towards α -chiral acids *via* the same decarboxylative activity (bottom).



An assay for microorganisms with the ability to grow on phenylmalonic acid as the sole carbon source led to the identification of the soil bacterium *Alcaligenes bronchisepticus* KU 1201 (now: *Bordetella bronchiseptica*). It could be shown that whole cells of this bacterium also converted 2-methyl-2-phenylmalonic acid and analogous substrates with various aromatic residues.^{130,131}

The decarboxylase was purified from the microorganism and named arylmalonate decarboxylase (AMDase), due to its preference towards

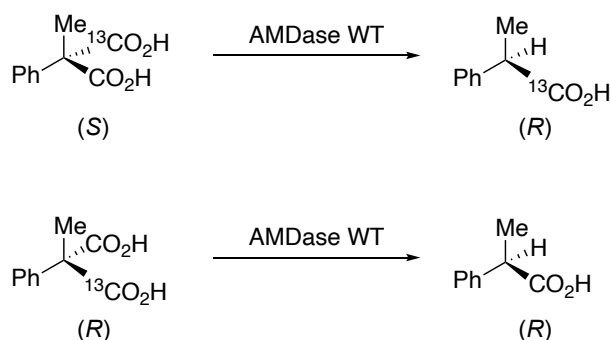
^a A review containing parts of this chapter is currently in preparation.

α -arylmalonates. Characterization of the purified enzyme revealed that it acts without a cofactor and is biotin-independent. Yet, the native role and substrate of AMDase remained unclear, as aryl malonates are not naturally abundant.¹³¹

4.1.1.1 Early suggestions on enzyme mechanism and selectivity

Shortly after the enzyme was discovered, Miyamoto *et al.* aimed at revealing the enzyme stereoselectivity by means of isotope labelling studies.¹³² For that, both enantiomers of ¹³C-labelled 2-methyl-2-phenylmalonic acid were synthesized. Products of the enzymatic reaction were analyzed *via* ¹³C NMR spectroscopy, which disclosed, that if the (*S*)-substrate was used, the (*R*)-configured product still contained the ¹³C-labelled carboxylate, whereas in case of the (*R*)-substrate, enrichment of ¹³C was not detected in the remaining carboxylate of the (*R*)-configured product (Scheme 18). These results indicated, that, in both cases, exclusively the pro-(*R*) carboxylate is cleaved, yielding the final product *via* inversion of configuration. Later on, this finding was confirmed by Okrasa *et al.* using ¹⁸O labelling of the enantiotopic carboxylate groups.¹³³

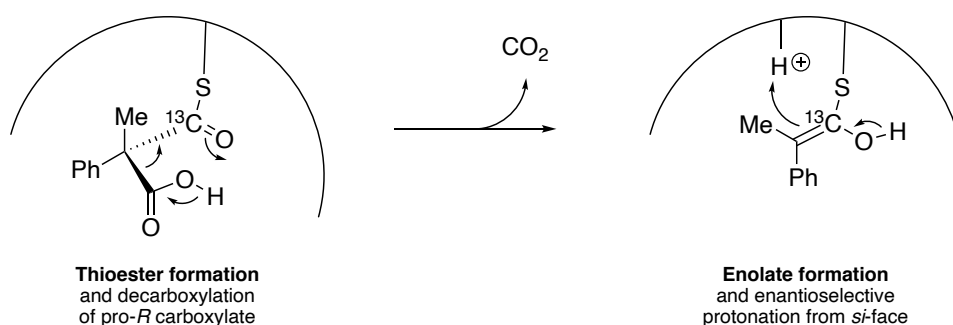
Scheme 18. Detected reaction products after decarboxylation of ¹³C-labelled chiral malonate substrates by wildtype AMDase indicates, that the pro-(*R*) carboxylate is eliminated during the reaction in any case.



The observation that AMDase is inhibited by sulfhydryl reagents indicated that AMDase is likely to be a thiol decarboxylase.^{131,132} Thus, the initial step of the reaction was believed to be the formation of a thioester between the pro-(*S*) carboxylate of the substrate and a cysteine residue in the active site (*cf.* activation by coenzyme A in fatty acid biosynthesis¹³⁴). The observed inversion of configuration, however, was in discrepancy to other known decarboxylases, where configuration was strictly retained.¹³⁴ This prompted the authors to suggest

either a S_E2 -type concerted mechanism with complete inversion, or the formation of an intermediary enolate, which further gets protonated in an enantioselective fashion from the *si*-face (Scheme 19).¹³² Formation of a charged intermediate was also supported by the finding that substrates bearing electron-withdrawing groups (EWGs) were converted at a higher rate, due to their ability to stabilize the supposedly formed carbanion intermediate.¹³¹

Scheme 19. First postulated mechanism based on the assumption, that AMDase acts as a thiol decarboxylase. After thioester formation between the substrate and a cysteine in the active site, the pro-(*R*) carboxylate is cleaved and the resulting enolate is enantioselectively protonated from the *si*-face.



Identification of the AMDase gene and sequence analysis thereof disclosed that the enzyme contains four cysteine residues (C101, C148, C171, C188).¹³⁵ All were present in a reduced state and only mutation of C188 to serine (C188S) proved detrimental to enzyme activity (k_{cat}), indicating its critical role in the reaction mechanism.¹³⁶

Soon thereafter, however, it became apparent that thioester formation does not play a role in the mechanism of AMDase. Instead, it was suggested that Cys188 rather acts as a proton donor. This became obvious after the wildtype enzyme, having the more acidic cysteine residue in the active center, was inactivated at a $pH > 9$, and the corresponding C188S variant was not. Interesting evidence was also found among quite remotely related (30 % homology) enzymes from the racemase and isomerase family.¹³⁷ Those enzymes, which were well-studied with regards to their mechanism, also exhibit a highly conserved cysteine residue in this region, which functions as a proton donating residue.¹³⁸

4.1.2 Inversion of enantioselectivity and introduction of racemase activity

Some enzymes with about 30 % homology to AMDase were identified *via* PSI-BLAST (position-specific iterative basic local alignment search tool), all belonging to the class of isomerases (EC 5). Glutamate racemase (*Lactobacillus fermenti*), aspartate racemase (*Streptococcus thermophilus*), hydantoine racemase (*Pseudomonas* sp. strain NS671) and maleate isomerase (*Alcaligenes faecalis*) all shared the conserved cysteine at position 188 with AMDase, while the enzymes from the isomerase class had an additional cysteine at around residue 74 (Figure 15).¹³⁹ As the reaction mechanism and crystal structure of glutamate racemase were already well-studied, two cysteine residues on both sides of the substrate were proposed to be essential for the racemizing activity, by either abstracting or donating protons from opposite sites in a so-called two-base mechanism.¹³⁸

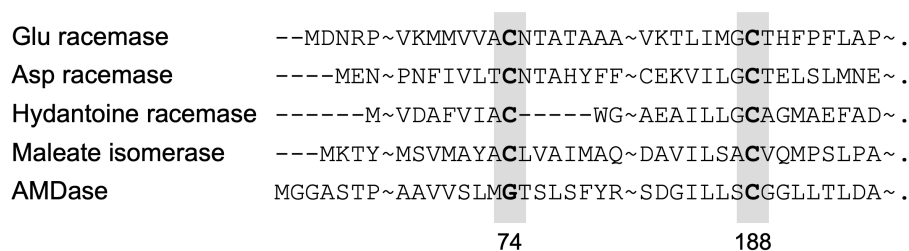


Figure 15. Amino acid sequence homology between AMDase and selected racemases identified *via* PSI-BLAST. Glutamate racemase from *Lactobacillus fermenti*, aspartate racemase from *Streptococcus thermophilus*, hydantoine racemase from *Pseudomonas* sp. strain NS671, maleate isomerase from *Alcaligenes faecalis* and AMDase from *Bordetella bronchiseptica*. Figure adapted from Ijima *et al.*¹³⁹

Inspired by these findings, Ijima *et al.* decided to introduce a cysteine at position 74 instead of the glycine residue present in AMDase.¹³⁹ Additionally, previous studies have shown that the mutation G188S led to a drastic decrease of native AMDase activity.¹³⁶ While the proton-donating ability of serine is already low compared to cysteine, the authors proposed, that an amino acid without any acidic proton would be beneficial for the enantiomeric excess of the obtained product. Surprisingly, the mutant G188A proved to be completely inactive. Hence, the double-variant G74C/C188S was prepared, and was found to produce the opposite enantiomers in 94-96 %ee as compared to wildtype AMDase. Yet,

activity of the G74C/C188S variant even fell below the tremendously reduced activity of C188S.¹³⁹

Interestingly, also for the (*S*)-selective variant S36N/G74C/C188S,¹⁴⁰ preference for cleavage of the pro-(*R*) carboxylate was observed by isotope labeling studies, thus confirming that enantioselective reprotonation of the enolate intermediate is decisive for the final configuration of the product.¹⁴¹ As already observed by Iijima *et al.*, the single amino acid exchange G74C rendered arylmalonate decarboxylase a racemase.¹⁴² Notably, the decarboxylase activity of AMDase was preserved in the G74C variant, yielding racemic arylpropionates from arylmalonates. Kinetic analysis disclosed, that in terms of catalytic efficiency ($k_{\text{cat}}/K_{\text{M}}$), decarboxylation ($0.96 \text{ s}^{-1}\text{mM}^{-1}$) exceeded racemization ($0.56 \text{ s}^{-1}\text{mM}^{-1}$), which was also reflected by the racemic product already detected at an early stage of the reaction.¹⁴²

4.1.3 Crystal structure and mechanism

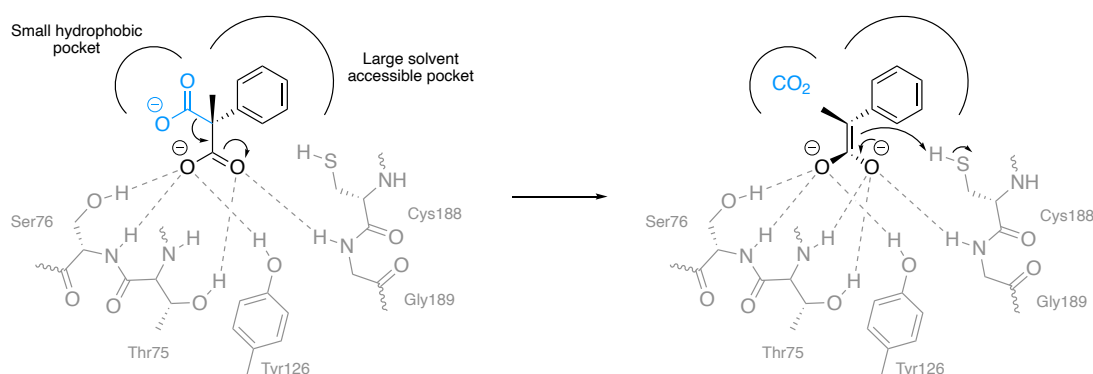
Despite the crystal structure of AMDase remained unsolved until 2008, considerable insights were already acquired by that time, including suggestions on the enzyme mechanism,^{132,137,141} identification of key residues,^{137,143} switch in enantioselectivity^{139,140} and introduction of a racemase activity.¹⁴² Yet, the exact mechanism for stabilization of the highly unstable enediolate intermediate remained elusive. While other enzymes with related enolate intermediates use Mg^{2+} for stabilization,¹⁴⁴ activity of AMDase is not dependent on metal ions or other cofactors.¹³¹ Further, delocalization of electron density into the aromatic ring system might partly account for stabilization of the enediolate but would be insufficient to explain the observed highly efficient enzymatic decarboxylation.¹³³ By solving the first crystal structure of wildtype AMDase (PDB: 3DG9), Okrasa *et al.* were able to unravel key features of the enzyme activity and selectivity in more detail.

Interestingly, in the core structure consisting of two four-stranded parallel β -sheets surrounded by several α -helices, a tightly bound phosphate ion was found near the active site cysteine 188. A total of six hydrogen bonds were established by side-chain and backbone interactions to Thr75, Ser76, Tyr126, Cys188 and Gly189, belonging to two adjacent oxyanion holes.¹³³ The authors suggested that the phosphate might resemble the position of the enediolate

intermediate. In retrospective, this structural motif termed “dioxyanion hole” was also found in related enzymes of the isomerase family and mutational studies on Thr75 and Ser76 or Tyr126 of AMDase have already demonstrated the essential role of this region for enzyme activity.^{133,143}

After reconfirming loss of the pro-(*R*) carboxylate by a ¹⁸O-labelling strategy, the enediolate intermediate resulting from decarboxylation of 2-methyl-2-phenylmalonate was placed in the active site. It became obvious, that the phenyl moiety occupies a large solvent accessible pocket, whereas the small methyl substituent is left at an orientation, where only little space is available. Considering these restrictions and the fixed position of the pro-(*S*) carboxylate entrapped in the dioxyanion hole, it was suggested that the pro-(*R*) carboxylate must point towards a small hydrophobic pocket (Leu40, Val43, Val156, Tyr48), which could act as a driving force for decarboxylation by hydrophobic destabilization (Scheme 20). Thus, the key for cofactor-free decarboxylation by AMDase was proposed to be the extensive stabilization of one carboxylate, while the other one is destabilized by unfavorable electrostatic interactions.¹³³

Scheme 20. Proposed binding mode and interaction network of 2-methyl-2-phenylmalonate (left) and the resulting enediolate intermediate (right) in the active site of AMDase WT. The pro-(*R*) carboxylate is shown in blue and the dioxyanion hole in gray. Scheme adapted from Okrasa *et al.*¹³³



This proposed binding mode was confirmed when AMDase was co-crystallized with the mechanism-based inhibitor benzylphosphonate ($K_i = 5.2$ mM) (PDB: 3IP8).¹⁴⁵ Similarly, the phosphonate dianion was tightly engaged by six hydrogen bonds to the residues of the dioxyanion hole and the phenyl residue was positioned in the large pocket within the Gly189-Gly190 amide bond and Pro14 through van der Waals interactions.

Obata *et al.* solved the crystal structure of the G74C/C188S variant containing the ligand 2-phenyl acetate (PDB: 3IXL) and of the G74C variant (PDB: 3IXM) and WT enzyme (PDB: 3DTV) in a sulfate associated form.¹⁴⁶ While the overall structure of AMDase G74C/C188S was mostly preserved when compared to wildtype structures either containing BnzPO (PDB: 3IP8) or PO_4^{3-} (PDB: 3DG9) in the active site (RMSD of C_α atoms approx. 0.25 Å), considerable conformational differences became obvious by comparing to the SO_4^{2-} -associated structures of the wildtype and G74C variant (PDB: 3DTV and 3IXM).

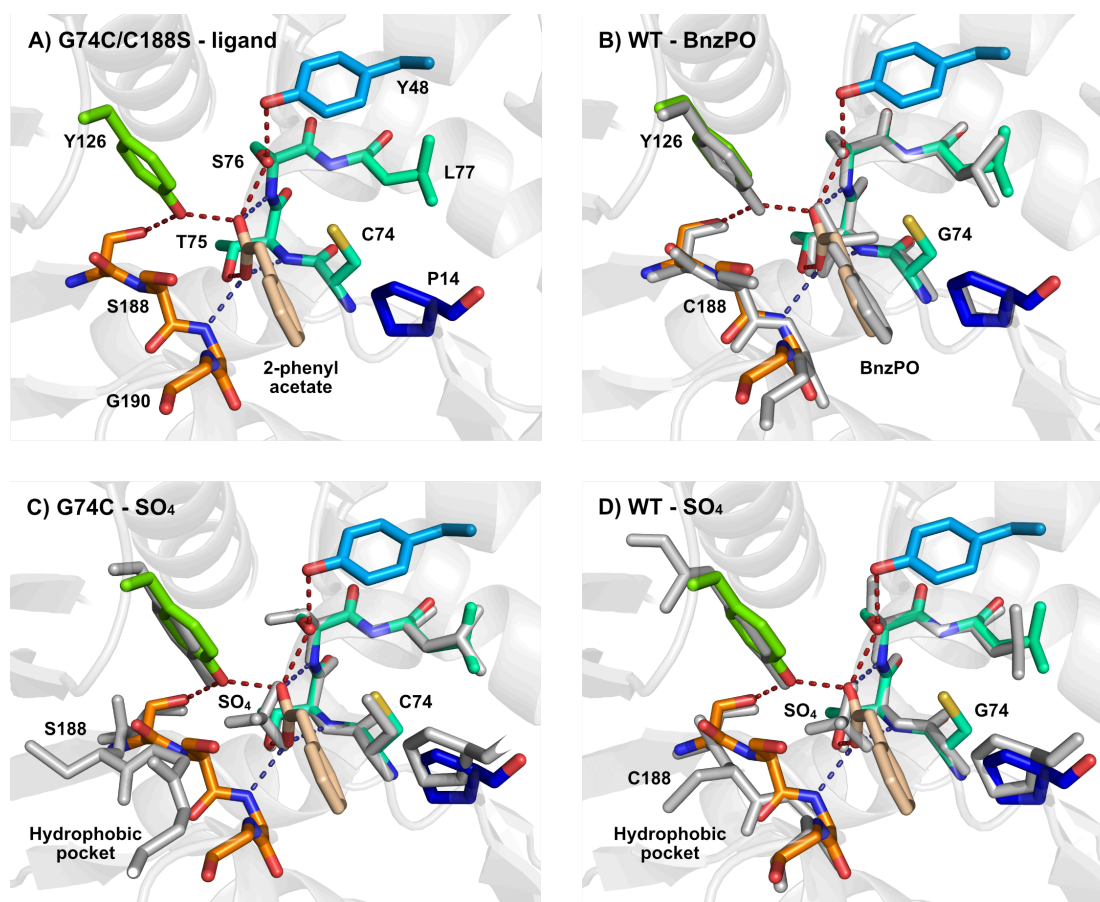


Figure 16. Representation of the active site of (A) AMDase G74C/C188S in association with 2-phenyl acetate (PDB: 3IXL) and superimposed with (B) AMDase WT in complex with benzylphosphonate (BnzPO) (PDB: 3IP8), (C) AMDase G74C in complex with SO_4^{2-} (PDB: 3IXM) and (D) AMDase WT in complex with SO_4^{2-} (PDB: 3DTV). Active site residues of AMDase G74C/C188S are shown in color and the respective superimposed structure in gray. Labelling of residues refers to the respective superimposed structure. Dotted lines represent H-bonds formed between the carboxylate and residue side-chains (red) or backbone NH groups (blue). Figure created with PyMOL⁸⁸ according to Obata *et al.*¹⁴⁶

It was suggested that a positional shift of Cys188-Gly189-Gly190 towards Gly74-Thr75 was induced upon ligand binding, thereby triggering the formation of a

hydrophobic network covering the active site (Val43, Thr154, Val156). In the unliganded (sulphate-associated) crystal structures, such cluster was not observed, as the corresponding loop regions were not contacting each other. Additionally, in this conformation, the key residue 188 is rotated away from the ligand towards the hydrophobic pocket in a rather “non-productive” orientation. These observations indicated the presence of either an “open” or “closed” conformation, regulated by ligand binding (Figure 16).¹⁴⁶ A similar difference between structures of the empty and the ligand-bound enzyme was found for the related glutamate racemase.¹⁴⁷

Most importantly, however, was the observation, that the position of Cys74 of the G74C/C188S variant was indeed in mirror symmetry to C188 of AMDase WT with regards to the C_α atom of the enediolate intermediate, which gets protonated during the reaction, thereby confirming the early proposed rationale for inversion of enantioselectivity.^{139,143} In this binding mode, configuration of the final product is only dependent on the enantioface-selective protonation, which occurs from *si*-face in case of Cys188 or from *re*-face in case of Cys74, resulting in (*R*)- or (*S*)-2-phenylpropionat, respectively in accordance to experimental data.¹⁴⁸

4.1.4 Computational modelling on the enzyme mechanism

The reaction mechanism of AMDase was also subject of computational studies. Lind and Himo studied the mechanism by means of density functional theory (DFT) calculations.¹⁴⁹ By employing the quantum chemical cluster approach, the mechanism, stationary points thereof and enantioselectivity were investigated with the aid of two different truncated active site models. The smaller model contained only partly the residues of the dioxyanion hole and the catalytic Cys188 (81 atoms), whereas the second model additionally included residues from the small and large binding pockets (223 atoms). Results obtained from using the small model indicated that a two-step mechanism proceeding *via* a planar enediolate intermediate is plausible, with the decarboxylation step being rate-limiting. However, binding modes leading either to cleavage of the pro-(*R*) or pro-(*S*) carboxylate were very similar in energy (1.9 kcal mol⁻¹ for initial binding pose; 1.5 kcal mol⁻¹ for transition state) probably due to the absence of the binding pockets determining the orientation of the remaining substituents, therefore being unable to explain the observed enantioselectivity with this small

model (>99 %ee corresponds to at least 3 kcal mol⁻¹). By using the larger model, the calculated energies for binding and decarboxylation were 14.1 and 18.3 kcal mol⁻¹ higher for the formation of the (S)-enantiomer than for the (R)-product, mostly caused by unfavorable steric clashes of the large aryl substituent and the residues of the small binding pocket. When the much smaller methylvinylmalonate was studied, ambiguous results were obtained concerning absolute transition state energies, probably being a result of the too small and rigid models of the binding pockets. Both binding modes and transition states, however, were much closer in energy as compared to methylphenylmalonate. This would indicate, that stereoselectivity is exclusively determined by substrate binding in case of bulky aromatic compounds, whereas in case of smaller substrates, the decarboxylation transition state could also contribute to enantiodiscrimination.¹⁴⁹ Just recently, Dasgupta *et al.* emphasized that geometric constraints such as introduced in quantum-chemical studies can lead to artifacts like imaginary vibrational frequencies and impaired efficiency of the overall optimization process.¹⁵⁰ The authors thus introduced soft harmonic confining potentials to the terminal atoms of the model, thereby avoiding the artificial strain and rigidity of fixed-atom truncated active site models. By employing this system, they were able to reproduce the results from previous studies¹⁴⁹ and further claimed, that this methodology is easy to implement and can dramatically reduce optimization efforts.¹⁵⁰

Busch *et al.* used semiempirical QM/MM calculations to study the mechanism of the AMDase racemase variant G74C.¹⁵¹ MD simulations revealed that both catalytic cysteines need to be in their deprotonated state for successful catalysis. In contrast to glutamate racemase, where co-catalytic residues activate the cysteines,¹⁵² water is suggested to deprotonate the corresponding residues of AMDase G74C prior to substrate binding, which are then further stabilized by thiolate holes. This might also explain the pronounced pH-dependency of the reaction. Further semiempirical calculations indicate a stepwise mechanism similar to decarboxylation, where a delocalized π -electron system within the substrate is necessary to stabilize the enediolate intermediate.¹⁵¹ Likewise to previous studies,¹⁴⁹ there was no indication for a concerted mechanism as described for Glu racemase.¹⁴⁷

Another study conducted by Karmakar *et al.* aimed at the simulation of CO₂ and product release from the active site.¹⁵³ By performing (steered) MD simulations it was demonstrated that release of the decarboxylated product is energetically unfavorable and might even surpass the decarboxylation step, thus potentially rendering product release the true rate-limiting step (WT: ~14 kcal mol⁻¹ for decarboxylation vs. ~23 kcal mol⁻¹ for product release). The authors claimed that this might also partly account for the reduced activity of the G188S/G74C variant (20 000-fold),¹⁴⁸ as the calculated energy barrier for product release was even higher in this case (~20 kcal mol⁻¹ for decarboxylation vs. ~37 kcal mol⁻¹ for product release).¹⁵³

4.1.5 Substrate scope

Already early after AMDase discovery, Miyamoto and Ohta thoroughly characterized the enzyme and its substrate scope.¹³¹ These fundamental findings, obtained in absence of any structural and mechanistic knowledge on the enzyme, provided the basis for studies up to the present, by demonstrating the essential limitations of AMDase substrates. (1) Steric effects play a crucial role for the small substituent, as mono-substituted malonates were converted faster than the corresponding methylated substrates. Further, the ethyl analogue was not converted. (2) The aromatic substituent is essential for activity, as substrates with a benzyl-, phenoxy- or phenylthio- instead of an aryl-group were inert. This was also true for malonate (R^L = H) or methylmalonate (R^L = CH₃). (3) Only free malonic acids can act as a substrate, as the corresponding mono- and diesters were not accepted. (4) Electron-withdrawing substituents on the aromatic moiety enhance the reaction rate by better stabilizing the proposed carbanion intermediate.

Table 8 summarizes substrates studied in AMDase catalyzed decarboxylation and racemization. The large substituent R^L offers major variability within the substrate scaffold. Phenyl-,^{130,131,156–161,133,136,142,143,145,148,154,155} 2-naphthyl-^{131,139,140,142,143,148} and thienyl-^{130,131,139,140,142,143,159} malonic acids are typical substrates of AMDase and omnipresent in literature. Over time, the substrate scope was extended to a vast number of variously substituted phenyl derivatives (*p*-Me,^{131,157} *p*-*iso*-butyl,¹⁵⁶ *p*-NO₂,¹⁵⁷ OMe,^{130,131,157} F,^{131,158} Cl,^{130,131,154} CF₃,¹⁵⁸ *p*-Ph-*m*-F,^{42,157} *m*-COPh¹⁵⁷), naphthyl derivatives (1-naphthyl,¹³¹

6-methoxynaphthalen-2-yl^{142,130,148,156,157}) heterocyclic (pyridinyl,^{157,159} furanyl and bicyclic systems¹⁵⁹) and also alkenyl^{145,157} substrates.

Table 8. Overview on the AMDase substrate scope.

Substituent	Generally accepted	Accepted by racemase	Not accepted
R^L	 R = H, <i>p</i> -Me, <i>p</i> - <i>iso</i> -Bu, <i>p</i> -NO ₂ , <i>p/m/o</i> -OMe, <i>p/m/o</i> -F, <i>p/m/o</i> -Cl, <i>p/m/o</i> -CF ₃ , <i>p</i> -Ph- <i>m</i> -F, <i>m</i> -COPh	 X = S; R = H X = O; R = Me, OMe, 2 Me	 X = CH ₂ , O, S
	 R = H, OMe	 R ¹ = H; R ² = H, Me, Et, Ph R ¹ = Me; R ² = Me	 Me H
R^S	H, D, Me, OH, F, NH ₂	Et	<i>i</i> Pr <i>n</i> Pr
EWG	CO ₂ H	NO ₂	COMe COH CO ₂ Me CO ₂ Et CONH ₂ CH ₂ OH COSEt CN

With regards to the small substituent R^S , it was soon recognized, that, besides H and CH₃, other small groups like D,¹⁶¹ F,¹⁵⁸ OH or NH₂¹⁵⁵ were generally tolerated, while larger alkyl groups (ethyl¹³¹ or propyl^{142,157}) were not converted or only at a very low rate.^{142,157} This was later on explained by structural analysis, as only limited space is available at the position left for this group.^{133,146}

As already exposed by Miyamoto *et al.* in 1992, presence of the free di-acid is indispensable for decarboxylation to occur.¹³¹ This was also proven valid for the racemization activity of AMDase, as carboxylate derivatives such as alcohols, amides, nitriles,¹⁴² esters,^{142,157} thioesters, ketones and aldehydes¹⁵⁷ were completely inactive. Thus, stabilization of the free carboxylate *via* the complex network within the dioxyanion hole seems crucial, regardless of decarboxylation or racemization. The only exceptional case known so far is (nitromethyl)benzene, which was accepted as a substrate by AMDase racemase G74C and G74C/V43A.¹⁵⁷ It was generally concluded, that the substrate structure seems more decisive for successful conversion, than α -proton acidity of the α -aryl propionates.¹⁵⁷

In terms of spatial arrangement, Miyamoto *et al.* proposed that a co-planar alignment of the phenyl group and α -substituent would allow optimal overlap of the π -orbitals in order to stabilize the formed charge on the enolate intermediate.¹⁵⁴ In case of *ortho*-substituted substrates like *o*-chlorophenyl malonic acid, two such conformations come into consideration, namely a *syn*- and *anti*-periplanar arrangement (Figure 17).

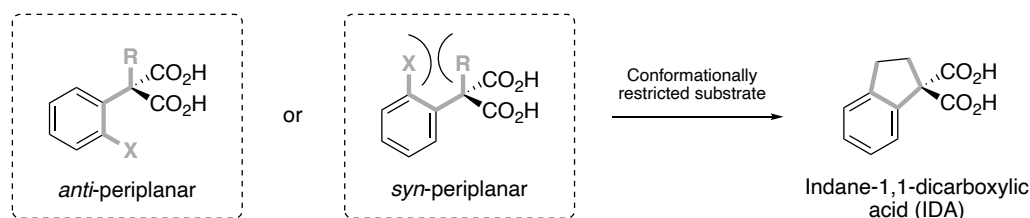


Figure 17. Coplanar arrangement of aromatic ring and α -substituent R in a *syn*- or *anti*-periplanar conformation with respect to the *ortho*-substituent X (bold gray bonds). Indane-1,1-dicarboxylic acid (IDA) serves as conformationally restricted model substrate to study entropic contributions for substrate acceptance.

Kinetic studies revealed that, while the activity of AMDase towards *o*-chlorophenylmalonic acid even surpassed phenylmalonic acid, the corresponding α -methyl-substituted compound was not accepted at all. This led to the assumption, that a *syn*-periplanar arrangement is necessary for AMDase-mediated decarboxylation, which might in turn disfavor substrates with an additional α -substituent, due to steric clashes with the *ortho*-substituent.¹⁵⁴

Ab initio calculations were thus performed to assess the energy barriers encountered during rotation along the C_{α} - C_{Ar} bond. Results obtained for

o-chlorophenylmalonic acid indicated two stable conformations corresponding to a *syn*- or *anti*-periplanar orientation (approx. 0° or 360° and 180° respectively). The small energy difference between both structures (< 1 kcal mol⁻¹) implies, that steric repulsion between *o*-Cl and α -H is negligible. Repulsive contributions were mainly caused by interaction of the chlorine residue with the carboxylates (Figure 18). In case of α -(*o*-chlorophenyl)- α -methylmalonic acid, an *anti*-periplanar or perpendicular arrangement of substituents was energetically favored, whereas *syn*-periplanar-like conformations were found to be unstable (Figure 18).¹⁵⁴

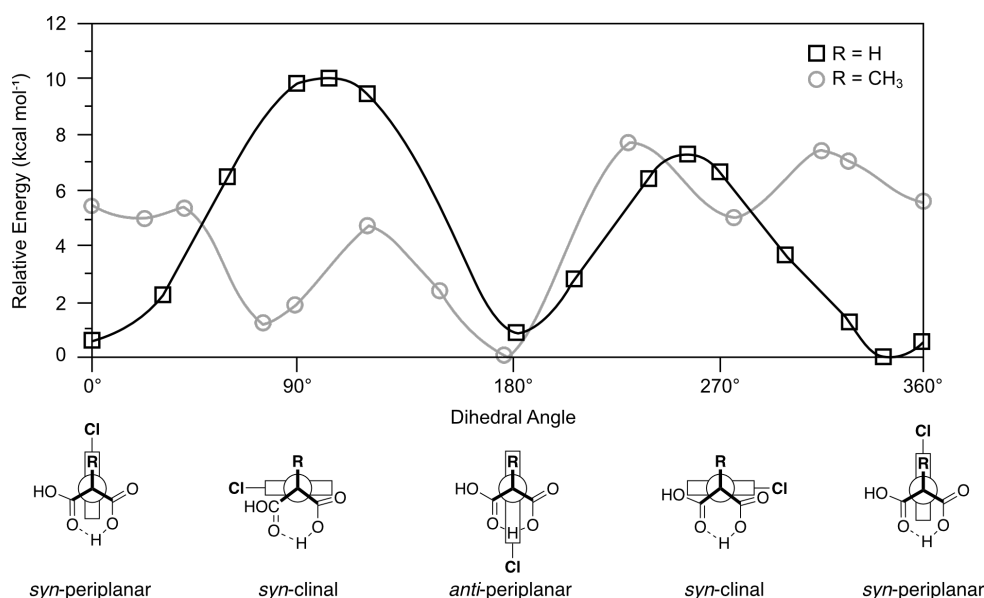


Figure 18. Potential energy diagram for the C-C bond rotation of *o*-chlorophenylmalonic acid (black squares) and α -(*o*-chlorophenyl)- α -methylmalonic acid (gray circles) calculated with HF/3-21G*. The phenyl ring is represented as rectangle in the Newman projection of the respective rotamers. Figure adapted from Miyamoto *et al.*¹⁵⁴

It was further proposed that, by using the conformationally restricted indane-1,1-dicarboxylic acid (IDA), a *syn*-periplanar conformation would be mimicked, thereby lowering the activation entropy ΔS^\ddagger (Figure 17). Indeed, the model substrate IDA was accepted by AMDase in contrast to the corresponding α -methyl- α -(*o*-tolyl)malonic acid, thus underlining the inability of the latter to overcome rotational energy barriers towards a suitable *syn*-periplanar arrangement for conversion.¹⁵⁴ The remarkably low K_M value observed for IDA in comparison to other studied substrates further affirmed the suitability of such spatial arrangement for AMDase-catalyzed decarboxylation.^{154,162} In this context it was also proposed that a hydrophobic pocket within the active site might be

responsible for correct positioning of unrestricted substrates *via* CH- π interactions,¹⁶³ thereby reducing the activation entropy in a similar manner.¹⁶²

4.1.6 Enzyme engineering

In view of the dramatically reduced efficiency of the initially created variants with inversed enantioselectivity¹³⁹ or racemizing activity¹⁴² in comparison to the native enzyme, considerable efforts to recover or even to increase the activity by enzyme engineering have been made until today.

4.1.6.1 Enzyme engineering of C188X/G74C-based variants

At the time, when Ijima *et al.* first described the AMDase G74C/C188S variant, the obtained inversion of enantioselectivity by merely employing the yet uncovered spatial mirror symmetry of residues 188 and 74 was intriguing. Yet, the variant suffered from an incisive loss of activity due to diminished k_{cat} (Table 9).¹³⁹ Early attempts on recovering the activity by random mutagenesis led to a variant (S36N/G74C/C188S) with increased (10-fold) activity among a total of 50 000 screened clones.¹⁴⁰ This reflected the difficulty of enzyme evolution purely relying on randomization in the absence of rational guidance.

After structural data of AMDase became available and, inspired by previous findings, that variations of the hydrophobic pocket residues can strongly enhance enzyme activity,¹⁴⁵ Miyauchi *et al.* also attempted directed evolution of the G74C/C188S variant. This time, structural information provided the basis for three rounds of directed evolution *via* iterative saturation mutagenesis (ISM)³¹ of Ser188 in the first place, and residues of the hydrophobic pocket (Leu40, Val43, Tyr48, Leu77, Val156, Met159) in the following rounds (NNK codon for introduction of all 20 amino acids) (Figure 19).¹⁴⁸ The critical role of the Cys188-substituting residue on enzyme activity was demonstrated earlier by the G74C/C188A variant being completely inactive.¹³⁹ From the first round of screening, variant G74C/C188G was identified to exhibit 5.6-fold increased activity. By using this variant as a template for the second generation, a triple mutant (G74C/M159L/C188G) with 210-fold activity evolved. Interestingly, the mutant selected from the last screening round (Y48F/G74C/M159L/C188G) with 920-fold increased activity carried the mutation Y48F, which was already found beneficial in the second-generation mutants. All identified beneficial mutations

within the hydrophobic pocket were due to hydrophobic substitutions, thus underlining the high importance of this destabilizing environment for decarboxylation.¹⁴⁸

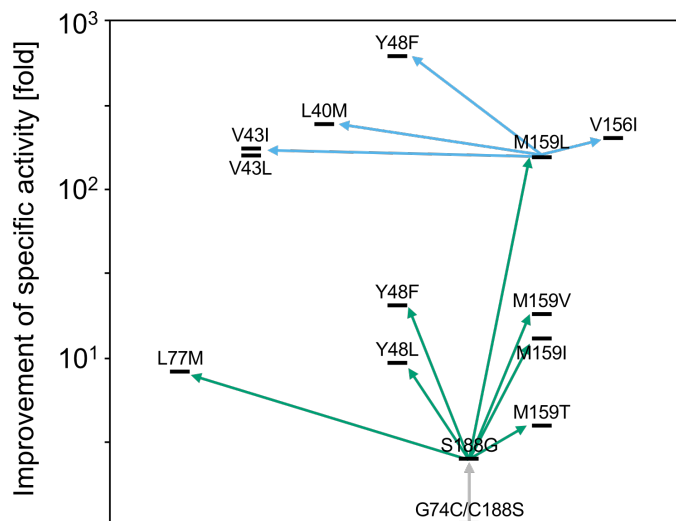


Figure 19. Enzyme variants evolved from AMDase G74C/C188S in three rounds of directed evolution *via* iterative saturation mutagenesis (ISM) and their improvement of specific activity towards phenylmalonate. First (gray), second (green) and third (blue) screening are shown as arrows. Figure adapted from Miyauchi *et al.*¹⁴⁸

As selection was based on conversion of phenylmalonate, activity towards 2-aryl-2-methylmalonates differed considerably. While the G74C/M159L/C188G variant showed a comparably increased activity towards production of Naproxen (220-fold higher specific activity), the quadruple variant Y48F/G74C/M159L/C188G, superior for conversion of phenylmalonate, exhibited only reduced activity. Additionally, the slightly impaired enantioselectivity of the G74C/C188S variant was completely abolished in the C188G-based variants (>99 %ee), indicating potential disadvantageous interactions of the serine moiety with the enolate intermediate.¹⁴⁸

Due to previously created variants, synergistic effects through multiple amino acid exchanges within the hydrophobic pocket were strongly indicated. Yoshida *et al.* conducted site-saturation mutagenesis (SSM) on the previously evolved G74C/M159L/C188G (CLG) variant,¹⁴⁸ which allowed for simultaneous variation of multiple sites.¹⁵⁶ Further, the authors argued, that evolution towards 2-methyl substituted arylmalonates would be advantageous compared to simple phenylmalonate, as the corresponding optical pure products are of major interest.¹⁶⁴ Again, residues from the hydrophobic pocket were chosen as

mutagenesis targets, due to their vicinity to the aryl residue (Leu40, Val156) or the proton-donating Cys74 (Met73). AMDase CLG was used as the starting variant, due to its superior activity in the formation of Naproxen¹⁴⁸ and a limited set of amino acids was introduced at the selected sites *via* the VTK codon (Leu, Ile, Val and Met). Surprisingly, only variants with mutations at V156 were selected from the first-generation library, thus indicating its crucial role in the decarboxylation mechanism. The best variant CLG-L (G74C/V156L/M159L/C188G) was chosen for the second round of SSM (Table 9), where the region around the important residues Val156 and Met159 was targeted (Val43 and Ala125). Interestingly, from the three selected variants with improved activity, two included the mutation A125P, which seemed surprising, as A125 was part of an α -helix. The most active variant identified in this screening was CLGL-IP (V43I/G74C/A125P/V156L/M159L/C188G), where the catalytic efficiency (k_{cat}/K_M) was improved 2.1-fold compared to CLG-L and over 9 000-fold compared to the initial G74C/C188S variant (Table 9).¹⁵⁶

Table 9. Kinetic parameters of AMDase WT and (S)-selective variants towards 2-methyl-2-phenylmalonic acid according to Yoshida *et al.*¹⁵⁶ Relative activity was calculated by dividing the catalytic efficiency (k_{cat}/K_M) of each variant by the catalytic efficiency of G74C/C188S variant.

AMDase variant	k_{cat}/K_M (s ⁻¹ mM ⁻¹)	Relative activity	Ref.
WT	279/26.9	28 090	133
G74C/C188S (CS)	0.0048/13	1	156
G74C/M159L/C188G (CLG)	1.1/1.8	1 655	156
G74C/V156L/M159L/C188G (CLGL)	1.7/1	4 604	156
V43I/ G74C/A125P/V156L/M159L/C188G (CLGIPL)	3.8/1.1	9 356	156

The CLGIPL variant turned out to be highly (S)-selective and most active amongst the (S)-selective enzyme variants towards all tested substrates. Notably, the precursor malonate of Ibuprofen was converted by AMDase CLGIPL with 1.3 U mg⁻¹ (not converted by G74C/C188S) which was also more than twofold faster than the wildtype enzyme (0.53 U mg⁻¹). Also Naproxen was formed by CLGIPL (17 U mg⁻¹) with high activity, which was yet surpassed by the (R)-selective wildtype (88 U mg⁻¹).¹⁵⁶

The AMDase CLGIPL variant also turned out highly efficient in the conversion of the malonate precursor of Flurbiprofen (55 U mg^{-1}) when compared to the wildtype (33.1 U mg^{-1}) or the (S)-selective CLG variant (15.8 U mg^{-1}).⁴² Interestingly, AMDase wildtype possesses higher activity towards the formation of Naproxen compared to Flurbiprofen, while AMDase CLGIPL shows an inversed preference. This emphasizes once more that the inherent reactivity of a substrate is less decisive for AMDase activity compared to the ability of productive binding in the active site and the sterical and electronical interactions resulting thereof.

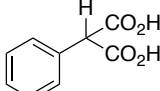
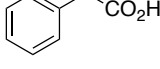
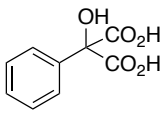
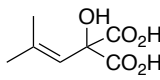
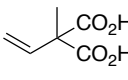
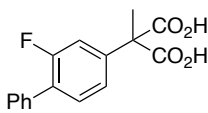
4.1.6.2 Enzyme engineering of WT-based variants

By solving the AMDase WT crystal structure in complex with the mechanism-based inhibitor benzyl phosphonate, Okrasa *et al.* were able to further rationalize interactions of the enzyme active site residues and the substrate scaffold.¹⁴⁵ They further proposed, that alkenyl moieties might also afford suitable delocalization of electron density during the reaction, similarly to the previously studied substrates with aryl residues. Indeed, the tested alkenyl substrates proved to be suitable for AMDase-catalyzed decarboxylation, yet at lower efficiencies as their aromatic counterparts. Despite their relatively low turnover numbers (k_{cat}), most of the alkenyl substrates often showed a higher affinity (K_{M}) for AMDase than the respective phenyl derivatives. To prove the hypothesized interaction patterns in the active center, two sets of residues (Pro14, Pro15 and Gly190 from the large binding pocket and Val43 and Met159 from the hydrophobic pocket) were selected for three rounds of iterative saturation mutagenesis (NNK coding for all 20 amino acids). Due to the established setup, mutants were screened for phenylmalonate, but were later on also tested towards the newly introduced alkenyl substrates.¹⁴⁵

By introducing the single point mutation M159V, the relative activity towards phenylmalonate was raised 51-fold, thus implying at the same time, that it is not the native substrate of AMDase (Table 10, entry 1). The double mutant P14V/P15G exhibited enhanced activity towards several different substrates (Table 10, entry 2-4), which could be attributed to a generally increased flexibility within the larger binding pocket. The applicability of the M159V and P14V/P15G

variants was later on also exemplified with a series of heterocyclic prochiral malonates.¹⁵⁹

Table 10. Kinetic parameters of AMDase wildtype-derived variants towards different substrates. Relative activities were taken from literature and were either calculated by dividing the catalytic efficiency (k_{cat}/K_M) of each variant by the catalytic efficiency of wildtype AMDase (entry 1-5) or the corresponding specific activities (entry 6). IPLL: AMDase V43I/A125P/V156L/M159L; n.d.: not determined.

Entry	Substrate	AMDase variant	k_{cat}/K_M ($\text{s}^{-1} \text{mM}^{-1}$)	Relative activity	Ref.
1		M159V	450/0.3	51	145
2		P14V/P15G	1143/3.5	11	145
3		P14V/P15G	99.8/15.3	1.9	145
4		P14V/P15G	34.4/8.0	1.5	145
5		G190A	20.8/0.8	4.6	145
6		IPLL	n.d.	6	42

The importance of substrate-fit was also demonstrated by the G190A variant, which converted 2-vinyl-2-methylmalonic acid, the smallest of all tested substrates, with 4.6-fold increased activity (Table 10, entry 5). The authors proposed that the slightly decreased space within the pocket might be beneficial for binding of the small vinyl group, which was also reflected in the lowered K_M value (0.8 mM) as compared to the wildtype (7.8 mM).¹⁴⁵

Gaßmeyer *et al.* showed that by transferring the set of amino acid substitutions previously found beneficial in the (*S*)-selective CLGIPL variant (V43I/A125P/V156L/M159L excluding C188G/G74C), a potent (*R*)-selective variant (IPLL) for production of (*R*)-Flurbiprofen is created. The quadruple variant converted the respective malonate with 209 U mg^{-1} to yield (*R*)-Flurbiprofen in 98 %*ee*, which represents a six-fold increase of specific activity compared to the native AMDase (Table 10, entry 6).⁴²

4.1.6.3 Enzyme engineering of G74C-based variants

The limited activity of the AMDase G74C racemase towards bulkier and industrially relevant α -arylpropionic acid derivatives (also referred to as profens) prompted Kourist *et al.* to create more suitable racemase variants by rational enzyme engineering.¹⁵⁷ While attempts to introduce co-catalytic residues (Asp, Glu, His) as present in the related glutamate racemase GluR¹⁵² at the corresponding positions of AMDase (Val13, Pro14, Gly190) only led to inactive mutants, variation of the hydrophobic pocket residues proved more successful. Destabilization within the hydrophobic pocket is crucial for decarboxylation, however, only plays a minor role in racemization. Yet, most of its constituting residues are in close proximity of the reaction center, thereby indirectly affecting the activity of the racemase. In case of AMDase G74C, MD simulations revealed that Val43 and Met159 are potent engineering targets, due to their vicinity to the small substituent. Two of the designed variants, AMDase G74C/V43A and G74C/M159L, were found to exhibit enhanced racemization activity, while efficiency towards decarboxylation was decreased at the same time, thus underlining the different determinants for both promiscuous activities. In total, activity of the G74C/V43A variant was 20-fold shifted from decarboxylation towards racemization. Remarkably, activity towards racemization of ketoprofen was also enhanced 30-fold.¹⁵⁷

In a later study, structure-guided protein engineering of the racemase was performed based on STD-NMR (saturation-transfer-difference NMR) analysis, which eventually led to the quadruple variant V43A/G74C/A125P/V156L with generally increased activity towards all tested profen derivatives and a maximum of 40-fold activity-increase towards Naproxen.¹⁶⁵

4.1.7 Recent applications

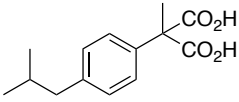
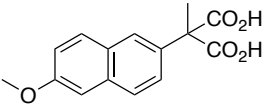
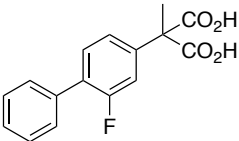
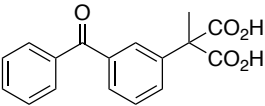
4.1.7.1 Production of non-steroidal anti-inflammatory drugs (NSAID)

Chiral α -aryl propionates are members of the group of non-steroidal anti-inflammatory drugs (NSAIDs) and possess anti-inflammatory as well as analgesic properties.¹⁶⁶ Members of this drug class belong to top-selling over-the-counter drugs and are used for the treatment of rheumatic disease (musculoskeletal pain) and acute or chronic pain.¹⁶⁷ NSAIDs inhibit prostaglandin biosynthesis by acting

on the cyclo-oxygenase (COX) enzyme system, which at the same time causes common adverse effects on gastrointestinal functions by non-specific inhibition of a different COX isoform.^{166,167}

Amongst the commonly known profens, only Naproxen is marketed as pure (S)-enantiomer, whereas Ibuprofen, Ketoprofen or Flurbiprofen are administered in a mixture together with their inactive (R)-enantiomers.¹⁶⁷ However, *in-vivo* unidirectional interconversion of the inactive (R)- to the pharmacologically active (S)-enantiomer can occur to a varying extent (*via* epimerization of the respective CoA-esters).^{167,168} Even though Flurbiprofen is mostly sold as a racemate, the (S)-enantiomer primarily accounts for the anti-inflammatory effect.¹⁶⁹ Yet, other biological effects of (R)-Flurbiprofen had been described,^{170–172} thus making the accessibility of both enantiomers a desirable goal.

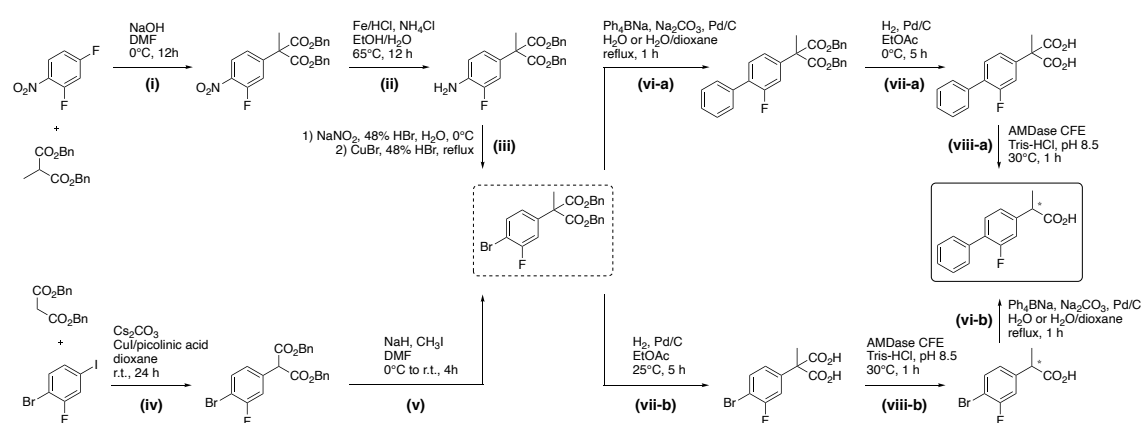
Table 11. Overview on AMDase substrates with relevance for research and industry. AMDase CLGIPL: V43I/G74C/A125P/V156L/M159L/C188G and IPLL: V43I/A125P/V156L/M159L. NSAID: non-steroidal anti-inflammatory drug.

AMDase substrate	Research interests	Ref.
 <p>Precursor malonate of Ibuprofen</p>	<ul style="list-style-type: none"> – Ibuprofen is sold as racemate – (S)-enantiomer acts as NSAID – Best variant: WT (R) or CLGIPL (S) 	156,157
 <p>Precursor malonate of Naproxen</p>	<ul style="list-style-type: none"> – Naproxen is sold as pure (S)-enantiomer – (S)-enantiomer acts as NSAID – Best variant: WT (R) or CLGIPL (S) 	42,148,156,157
 <p>Precursor malonate of Flurbiprofen</p>	<ul style="list-style-type: none"> – Flurbiprofen is sold as racemate – (S)-enantiomer acts as NSAID – (R)-enantiomer shows other activities^{170–172} – Best variant: IPLL (R) or CLGIPL (S) 	42,157
 <p>Precursor malonate of Ketoprofen</p>	<ul style="list-style-type: none"> – Ketoprofen is sold as racemate – (S)-enantiomer acts as NSAID – Not well accepted by AMDase to date 	157

The ability of AMDase to convert prochiral malonates into optically pure arylpropionic acids was early recognized in the context of profen synthesis, yet hampered by the availability of efficient enzyme variants or limited stability of the malonate precursors.¹⁷³ Table 11 should give an overview on NSAID precursors studied in AMDase catalyzed decarboxylation and the corresponding best-performing enzyme variants to date. Interestingly, several industrially employed routes for profen synthesis employ chemical malonate decarboxylation.^{174,175} Due to the inherent instability of the intermediately formed malonic acid, ester hydrolysis and decarboxylation are often carried out simultaneously, thus yielding racemic product mixtures.^{174–176} Yet, efforts to obtain pure enantiomers, like recrystallisation of diastereomeric mixtures or enzymatic kinetic resolution, are often tedious and suffer from low yields or limited enantiomeric excess.¹⁷³ In this respect, also enzymatic racemization under mild conditions could contribute to higher overall yields by recycling the undesired enantiomer for resolution processes.¹⁵⁷

In view of the aforementioned limitations of malonate-based routes towards Flurbiprofen, Enoki *et al.* attempted the implementation of an integrated chemo-enzymatic pathway towards Flurbiprofen including AMDase-mediated decarboxylation and the previously introduced protecting group strategy *via* benzyl esters.^{42,44} The authors envisioned two potential chemo-enzymatic routes similarly to previously described chemical synthesis strategies, yet adapted to the corresponding dibenzyl malonates (Scheme 21).^{173,176}

Scheme 21. Chemo-enzymatic synthesis routes towards optically pure Flurbiprofen. The upper route was inspired and adapted from an industrial chemical synthesis and the lower route was adapted to the aspired conditions of the chemo-enzymatic process.



Interestingly, the AMDase-catalyzed reaction step (Scheme 21, step (viii)) was planned at two different stages of the process. Either after Suzuki coupling, thereby using the established Flurbiprofen malonate as substrate (Scheme 21, step (viii-a)), or before Suzuki coupling, which would require that AMDase exhibits activity towards 2-(4-bromo-3-fluorophenyl)-2-methylmalonic acid (Scheme 21, step (viii-b)).⁴⁴ After necessary adaptations were established for the initial chemical steps towards the central intermediate dibenzyl 2-(4-bromo-3-fluorophenyl)-2-methylmalonate (Scheme 21, steps (i)-(v)), the authors focused on the second half of the cascade, including Pd-catalyzed Suzuki coupling (step (vi)), deprotection of the dibenzyl-protected malonates (step (vii)) and eventually AMDase catalyzed decarboxylation (step (viii)). It was suggested that palladium on charcoal (Pd/C) should be a suitable catalyst for both Suzuki coupling and cleavage of the benzyl esters. Notably, during hydrogenolysis, the central intermediate was less reactive (step (vii-b)) compared to the dibenzyl malonate precursor of Flurbiprofen (step (vii-a)), yet more stable towards spontaneous decarboxylation, which proved advantageous for obtaining a high final ee of the product.⁴⁴ While Suzuki coupling under mild conditions in water, using Pd/C as a catalyst and Ph₄BNa as phenyl group donor, was reported (step (vi-b)),¹⁷⁶ those particular conditions proved impractical for the corresponding reaction of the alternative route (step (vi-a)), due to solvent incompatibilities of the reagents. Alternative Suzuki coupling strategies in organic solvents were not considered at this point, by reasons of the aspired environmentally benign reaction conditions. Eventually, AMDase variants also proved active towards the yet uncharacterized substrate in step (viii-b) with 114.3 U mg⁻¹ (WT), 480.3 U mg⁻¹ (IPLL), 41.7 U mg⁻¹ (CLG) or 40.7 U mg⁻¹ (CLGIPL), which was comparable to the activities towards the precursor malonate of Flurbiprofen. Both enantiomers of Flurbiprofen could thus be produced in high yields and >99 %ee.⁴⁴

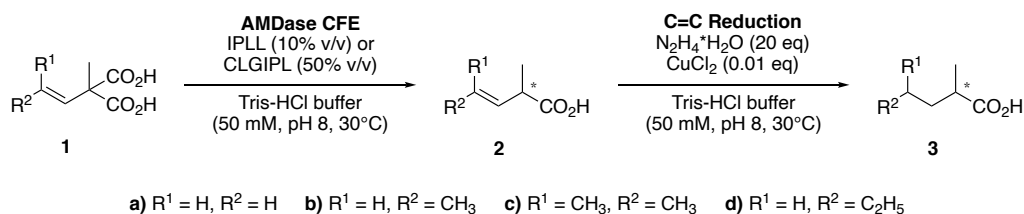
Using immobilized AMDase, Aßmann *et al.* achieved an optimized upscaled process for (S)-Naproxen synthesis, by careful analysis of kinetic parameters and in-line reaction monitoring *via* Raman spectroscopy. Importantly, AMDase CLGIPL was inhibited by the product, but only to an extent, where *in-situ* product removal was not necessary. A productivity of 140 kg_{product} kg_{enzyme}⁻¹ was calculated for the implemented process, which exceeds the minimum specified value for pharmaceutical processes of 50-100 kg_{product} kg_{enzyme}⁻¹. Further, product

isolation *via* precipitation was optimized and (S)-Naproxen was isolated in 92 % yield and 99 % ee.¹⁷⁷

4.1.7.2 Chemo-enzymatic cascade synthesis of optically pure alkanolic acids

AMDase requires a substituent with a π -electron system on the substrate and does not convert aliphatic malonic acids. Nevertheless, that enantiopure alkanolic acids could be accessed in two steps by combining AMDase-mediated decarboxylation and reduction of the non-activated double bond (Table 12). With other catalytic methods, such molecules are not easily accessible, particularly due to the difficulty of chemical catalysts to distinguish the structurally similar substituents.⁴⁶ While the activity of wildtype AMDase towards alkenyl malonic acids was already reported,¹⁴⁵ reactivity of other (*R*)- and (*S*)-selective enzyme variants was not yet studied with this class of substrates.

Table 12. Sequential chemo-enzymatic cascade of AMDase catalyzed decarboxylation of alkenyl malonates followed by C=C bond reduction with *in-situ* generated diimide.⁴⁶ After completion of the biocatalytic reaction step, hydrazine (20 eq) and CuCl₂ (0.01 eq) were added directly to the reaction mixture. The stated conversion refers to the final reaction step and was detected after 23 h.



Entry	Substrate	AMDase variant	Conversion (%)	ee of 3 (%)
1	1a	IPLL	>99	98 (<i>R</i>)
2	1a	CLGIPL	>99	66 (<i>S</i>)
3	1b	IPLL	80	>99 (<i>R</i>)
4	1b	CLGIPL	78	>99 (<i>S</i>)
5	1c	IPLL	20	>99 (<i>R</i>)
6	1c	CLGIPL	11	>99 (<i>S</i>)
7	1d	IPLL	86	>99 (<i>R</i>)
8	1d	CLGIPL	89	>99 (<i>S</i>)

As already demonstrated with other substrate types, AMDase IPLL and CLGIPL outperformed the wildtype enzyme and other (*S*)-selective variants, yet with a 19-170 and 190-2 200 times reduced activity, respectively, as compared to aromatic substrates. This reflected the lower capacity of an alkene to stabilize the enediolate intermediate.⁴⁶ Interestingly, all tested enzyme variants proved to be highly stereoselective, with the conversion of 2-methyl-2-vinylmalonate by AMDase CLGIPL as only exception. This reaction produced (*S*)-2-methylbut-3-enoic acid with surprisingly low 66 %*ee* (*S*) (Table 12). In this case, either an alternative proton donor (*e.g.* water) or binding mode¹⁴⁹ was discussed to cause such impaired enantioselectivity, as control reactions with other, highly selective AMDase variants ruled out a racemizing side-reaction as source of this low enantiomeric excess. Moreover, AMDase CLGIPL decarboxylated vinyl malonates with additional substituents present on the alkene with outstanding enantioselectivity, thus reflecting the fragile interaction network responsible for ligand recognition and binding. Regarding chemical C=C bond reduction, the potential risk of double bond isomerization during transition-metal catalyzed reduction, which would cause product racemization, could be eventually eliminated by using *in-situ* generated diimide as reductant. The high enantiomeric excess observed after AMDase-catalyzed decarboxylation could be thus retained until after the final chemical reduction step (Table 12).⁴⁶

4.1.7.3 Immobilization of AMDase

Considering the limited stability of purified AMDase under process conditions ($t_{1/2} = 1.2$ h),¹⁷⁸ differing immobilization strategies were tested recently, including site-specific¹⁷⁹ or conservative covalent immobilization,^{42,177,178} adsorption, complexation,¹⁷⁸ entrapment or combined approaches (Table 13).¹⁸⁰ Wong *et al.* studied the phosphopantetheinyl transferase (Sfp)-catalyzed immobilization of ybbR-tagged proteins (12-mer *N*-terminal tag) on CoA-functionalized polystyrene nanoparticles (Table 13, entry 1).¹⁷⁹ The covalent and site-specific linkage was efficiently achieved under mild conditions, and the obtained biocatalyst revealed high operational stability over four cycles (approx. 7 % loss of activity). Yet, the observed K_M values were approximately three times higher compared to the free enzyme, whereas K_M values for AMDase and ybbR-AMDase were similar, thus

indicating that unfavorable interactions between the nanoparticle and substrate are causing this observed loss in efficiency rather than the immobilization itself.¹⁷⁹

Table 13. Overview on immobilization techniques studied in AMDase catalyzed decarboxylation. Purif.: purified enzyme; CFE: Cell-free extract.

Entry	Carrier	Interaction	Enzyme (formulation)
1	Polystyrene nanoparticle, CoA functionalized	Covalent	AMDase WT (purif.) ¹⁷⁹
2	LentiKats® (PVA gel)	Entrapment	AMDase WT (CFE) ¹⁸⁰
3	Activated MMP (magnetic microparticles)	Covalent	AMDase WT (CFE) ¹⁸⁰
4	MMP-LentiKats®	Combined	AMDase WT (CFE) ¹⁸⁰
5	Amino C2 acrylate	Covalent	AMDase WT (purif.) ¹⁷⁸ AMDase CLGIPL (purif.) ⁴² AMDase CLGIPL (CFE) ^{177,178}
6	Sepabeads EC-EP (polymethacrylate)	Covalent	AMDase WT (purif.) ¹⁷⁸
7	Sepabeads EC-HA (polymethacrylate)	Covalent	AMDase WT (purif.) ¹⁷⁸
8	Trisoperl® (porous glass)	Adsorption	AMDase WT (purif.) ¹⁷⁸
9	Trisoperl® amino (porous glass)	Adsorption	AMDase WT (purif.) ¹⁷⁸
10	EziG1™ (Fe ^{III} or Co ^{II}) (porous glass with longchain aminoalkyl coating)	Complex	AMDase WT (purif.) ¹⁷⁸ AMDase CLGIPL (purif.) ¹⁷⁸
11	EziG2™ (Fe ^{III} or Co ^{II}) (porous glass with vinylbenzyl-chloride coating)	Complex	AMDase WT (purif.) ¹⁷⁸ AMDase CLGIPL (purif.) ¹⁷⁸
12	EziG3™ (Fe ^{III} or Co ^{II}) (porous glass with styrol/acrylonitrile copolymer coating)	Complex	AMDase CLGIPL (purif.) ¹⁷⁸

Enzyme immobilization can significantly contribute to an economical use of the biocatalyst (e.g.: stability, reuse) and facilitates downstream processing.¹⁸¹

Recently, Markošová *et al.* attempted entrapment of AMDase in polyvinyl alcohol (PVA) gel (Table 13, entry 2).¹⁸⁰ While AMDase showed enhanced stability, the formulation suffered from severe activity loss after repeated use, probably due to leaching of the enzyme. In a combined approach (Table 13, entry 4), the combination of the operational stability of covalently bound AMDase on magnetic microparticles (MMP) and easy handling of PVA gel beads afforded a biocatalyst, which was stable over eight biocatalytic reactions and upon storage at 4°C (64 % activity retained after three months).¹⁸⁰

Covalent immobilization on well-established amino C2 acrylate carrier proved practical to obtain a stable biocatalyst (approx. 20 % activity yield) with a half-life of 16.5 h and a total turnover number (TTN) of over 20 000.⁴² The same strategy was further studied and applied by Aßmann *et al.* in an upscaled process for (S)-Naproxen synthesis (TTN of 83 000-107 000 over five batches).^{177,178} Also here, the immobilization increased the K_M value towards formation of Naproxen from 0.08 to 22 mM, thus underlining the strong effect of mass-transfer limitation. At the same time different other support materials and binding strategies for immobilization were evaluated (Table 13, entry 5-12). Overall, covalently attached enzyme preparations were clearly preferred in terms of enzyme loading and long-term stability in repeated batches. For example, the activity of AMDase immobilized on EziG1™ was formidable, yet the coordinative binding between the enzyme's 6xHis-tag and Fe(III) or Co(II) on the carrier is considerably weaker and thus led to extensive leaching during repeated experiments. Immobilization on amino C2 acrylate was further optimized by using CFE instead of purified protein, which certainly led to a decreased catalytic activity (30 % compared to purified AMDase), yet to an enhanced long-term operational stability ($t_{1/2} = 8.6$ d).¹⁷⁸

4.1.8 AMDases from other organisms

In search of organisms with similar reactivities and characteristics as AMDase from *Alcaligenes bronchisepticus* (now: *Bordetella bronchiseptica*) KU1201, Ohta and coworkers were able to identify two further strains from soil samples by their ability to degrade phenylmalonate.^{182,183} The strains KU1311¹⁸² and KU1313¹⁸³ were the most active ones and identified as *Achromobacter* sp. and *Enterobacter cloacae*, respectively. Respective genes were amplified from the genomic DNA

and heterologously expressed in *E. coli* for characterization. Both genes encoded a protein of 240 amino acids length, consistent with the originally identified AMDase from strain KU1201¹³⁵ and both shared a high sequence homology of 94 % (KU1311) and 85 % (KU1313). Accordingly, the observed substrate selectivity and activity were comparable to the original AMDase and all shared the same strict enantiopreference.^{131,135,182,183}

While those three AMDases show considerable similarity, the shifted pH optimum (pH 5.5) towards acidic conditions of AMDase from strain KU1313 was intriguing. It was the only microorganism, which showed AMDase activity when soil samples were screened at acidic pH, others were rather active at neutral to slightly basic pH (pH 7-8.5) (Table 14).

Table 14. Overview on AMDases identified from different organisms. n.d.: not determined.

Organism	Maximum activity	Optimum pH	Ref.
<i>Bordetella bronchiseptica</i> KU1201 (formerly: <i>Alcaligenes bronchisepticus</i> KU1201)	45°C	8.5	131,135
<i>Achromobacter</i> sp. KU1311	40°C	8.5	182
<i>Enterobacter cloacae</i> KU1313	35°C	5.5	183
<i>Chelativorans</i> sp. BNC1 (formerly: <i>Mesorhizobium</i> sp. BNC1)	n.d.	n.d.	133
<i>Variovorax</i> sp. HH01	34°C	6.0	184
<i>Variovorax</i> sp. HH02	30°C	7.0	184
<i>Polymorphum gilvum</i> SL003B-26A1	37°C	7.0	184

Okrasa *et al.* tried to identify novel AMDase enzymes from protein sequences with a similarity of 30-52 % and a requisite cysteine residue at the same relative position as in the enzyme from *B. bronchiseptica*. The putative AMDases were characterized, however only the enzyme from *Mesorhizobium* sp. (now: *Chelativorans* sp.) BNC1, exhibiting the highest similarity, showed satisfactory decarboxylase activity, while the other candidates rather acted as racemases.¹³³ The difficulty of identifying novel AMDases accompanied by the issue of inappropriate gene annotation was addressed by the work of Maimanakos *et al.*,

who developed a sequence-based search algorithm for a more reliable prediction of enzymes possessing AMDase activity.¹⁸⁴ Sequence information of confirmed AMDases was used to specify 12 conserved sequence patterns (e.g. residues of the hydrophobic pocket, dioxyanion hole, aryl binding pocket, and residues C188 and G74 including their surroundings), which, in turn, were used for concise database screening. In this way, 58 additional ORFs encoding putative AMDase-like enzymes were found and used for generating a Hidden Markov Model (HMM), which furnished six specific motifs (Table 15) and was implemented in the applied search algorithm. To prove the reliability and applicability of these search criteria, the enzyme from *Polymorphum gilvum* SL003B-26A1, identified via a data base search, was characterized together with two enzymes from *Variovorax* sp. HH01 and HH02 identified via screening of soil samples. All three enzymes showed the desired AMDase activity within the common mesophilic temperature range and around neutral to slightly acidic pH (Table 14).¹⁸⁴

Table 15. Identified conserved sequence motifs and residues involved in the catalytic mechanism of AMDase (in bold). Unconserved residues are denoted with x. Residue numbering according to *B. bronchiseptica* AMDase.¹⁸⁴

No.	Sequence pattern	Residues	Description
1	GLIVPPAxGxVPxE (res.10-23)	P14, P15	Part of large pocket, aryl binding
2	GLGLxxVxxxGY (res. 37-48)	L40, V43, Y48	Part of hydrophobic pocket
3	GAxVxLM G TSLSFYRG (res. 66-82)	G74; T75, S76	C74 characteristic for racemases; part of dioxyanion hole
4	RVAVxTAY (res. 119-126)	Y126	Part of dioxyanion hole
5	LxLxxVxx M (res. 151-159)	V156, M159	Part of hydrophobic pocket
6	DALLIS C GxL (res. 182-191)	G189; C188	Part of dioxyanion hole; catalytic active residue

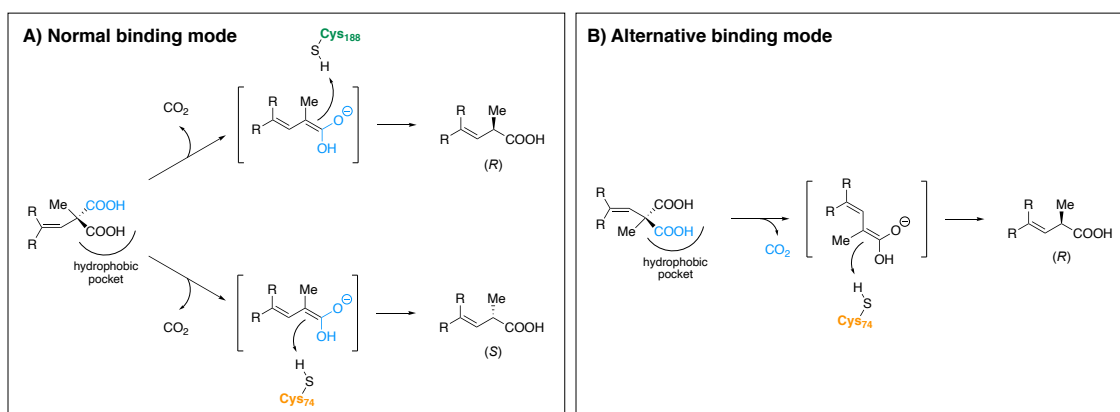
Interestingly, all putative AMDase-encoding ORFs identified by Maimanacos *et al.* belonged to the class of α -, β -, and γ -proteobacteria. When AMDase flanking regions were analyzed, genes encoding mandelate racemase/muconate lactonizing enzyme or several transporters (e.g. tripartite tricarboxylate

transporter TTT, TRAP or ABC transporters) were commonly found. According to these flanking regions and sequence similarities, AMDases were classified into eight enzyme clusters. While the natural role of these enzyme remains elusive, the fact, that close relatives of AMDase-producing bacteria often do not encode this enzyme, strongly indicated, that this gene was frequently lost in evolution and might be therefore only temporarily advantageous for organisms under certain physiologic conditions.¹⁸⁴

4.2 Aim of the project

Within this work, the contribution of a putative alternative binding mode during AMDase-catalyzed decarboxylation was investigated (Scheme 22). When Micklefield and coworkers first investigated the conversion of 2-methyl-2-alkenylmalonic acids by AMDase wildtype, the corresponding decarboxylation products were produced in high enantiopurities (99 %*ee* (*R*) for 2-methyl-2-vinylmalonic acid), thus confirming the outstanding enantioselectivity of AMDase towards the class of non-aromatic substrates.¹⁴⁵ Yet, during investigating the decarboxylation of 2-methyl-2-vinylmalonic acid catalyzed by AMDase CLGIPL, the enantioselectivity observed by Enoki *et al.* was only moderate (66 %*ee* (*S*); Table 12, entry 2). Notably, this deterioration was only observed for the smallest of all tested substrates and only for the CLGIPL variant.⁴⁶ This inexplicable behavior was assumed to result either from contribution of water or from a binding mode in which the pro-(*S*) carboxylate is accommodated in the hydrophobic pocket instead of the usually found pro-(*R*) group, thus resulting in an overall flipped orientation of all other substituents, consequently leading to the formation of the opposite enantiomer (Scheme 22).⁴⁶

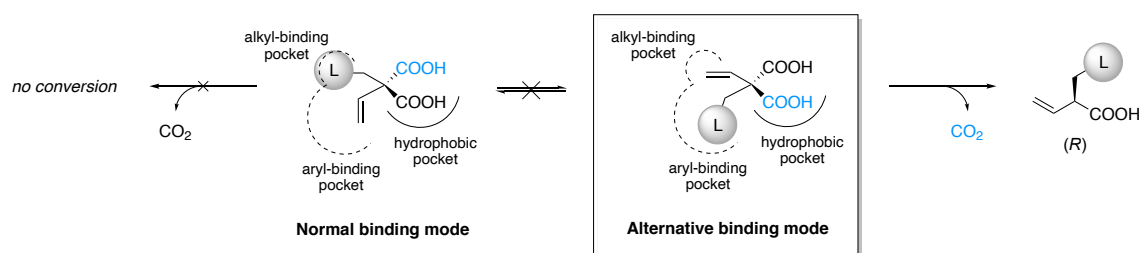
Scheme 22. Proposed binding modes of AMDase for small substrates. (A) Normal binding mode, where the pro-(*R*) carboxylate (black) is accommodated in the hydrophobic pocket and thus cleaved off. (*R*)-selective (Cys188) or (*S*)-selective (Cys74) AMDase variants would lead to the (*R*)- or (*S*)-product, respectively. (B) Putative alternative binding mode, where the pro-(*S*) carboxylate (blue) is instead accommodated in the hydrophobic pocket and thus cleaved. (*S*)-selective (Cys74) AMDase variants would lead to the (*R*)-product.



Particularly, it was hypothesized that a vinyl group might be small enough to occupy the spatially limited alkyl-binding pocket in the AMDase active site, usually seized by the methyl substituent (Scheme 23).

To test this hypothesis, the following strategies were pursued: (1) The influence of variations of active site residues of AMDase CLGIPL (V43I/G74C/A125P/V156L/M159L/C188G) was studied. Specifically, residues Ile43, Leu156 and Leu159 were exchanged by other hydrophobic amino acids (Ala, Ile, Leu, Met, Phe, Val) and residues at the reactive center (Gly190 and Gly188) were exchanged to alanine or serine. Also, combinations of several variants were tested. (2) By engineering the substrate scaffold in a way, where the methyl group is substituted by a larger substituent (*e.g.* ethyl, *iso*-butyl, *iso*-pentyl, 2,2-dimethylpropyl, benzyl), binding in the normal mode should be prevented. Instead, the substrate is expected to be constrained in the putative alternative binding mode, where the vinyl group occupies the alkyl-binding pocket, and the large substituent is accommodated in the larger aryl-binding pocket. Thus, the formation of the respective (*R*)-configured products from the typically (*S*)-selective AMDase CLGIPL and variants thereof is expected in case the substrate is exclusively bound in the alternative binding mode (Scheme 23).

Scheme 23. Proposed rationale for constraining the substrate in the putative alternative binding mode by substrate engineering. By installing a large substituent (gray sphere) in place of the methyl group, the vinyl group should be forced into the small alkyl-binding pocket. Thus, only binding in the alternative mode should be possible due to steric restrictions in the small pocket. The pro-(*S*) carboxylate is highlighted in blue.

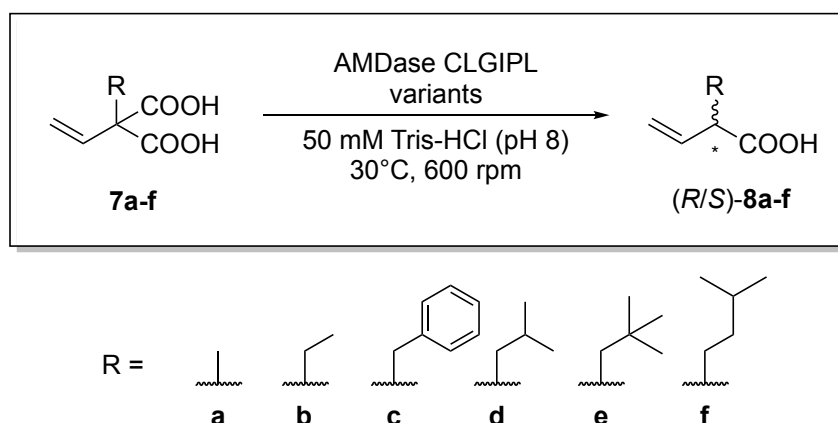


Further, computational studies aiming at elucidating the determinants for AMDase activity and selectivity were executed in cooperation with Prof. Lynn Kamerlin (Department of Chemistry–BMC, Uppsala University).

4.3 Results and discussion

This work focused mainly on characterizing the newly created AMDase CLGIPL variants in the decarboxylation of 2-methyl-2-vinylmalonic acid (**7a**), to enable direct comparison to literature data⁴⁶ and to deduce the influence of specific amino acid exchanges onto product enantiopurity (see 4.3.1-4.3.3 and 4.3.5). Furthermore, synthesis of a series of 2-alkyl-2-vinylmalonic acids **7b-f**, bearing larger alkyl substituents, was attempted in order to study the presence of the suggested alternative binding mode (Scheme 24, see 4.3.4).

Scheme 24. Decarboxylation of 2-alkyl-2-vinylmalonic acids **7a-f** catalyzed by AMDase CLGIPL and variants thereof studied in this work. R = methyl (**a**), ethyl (**b**), benzyl (**c**), *iso*-butyl (**d**), 2,2-dimethylpropyl (**e**), *iso*-pentyl (**f**).



4.3.1 Selection of target residues for site-directed mutagenesis

The AMDase CLGIPL (V43I/G74C/A125P/V156L/M159L/C188G) variant evolved from several rounds of simultaneous site-saturation mutagenesis (SSM). Compared to the initial (*S*)-selective variant G74C/C188S, the catalytic efficiency of the CLGIPL variant could be improved by 9 000-fold.¹⁵⁶ Overall, this variant turned out extremely versatile due to its high (*S*)-selectivity and activity towards a broad series of substrates.^{42,156} By transferring this set of mutations (V43I/A125P/V156L/M159L, except G74C/C188G) to AMDase WT, the powerful (*R*)-selective IPLL variant was generated (Table 16).⁴²

Table 16. Summarized residues of AMDase, which were substituted to generate the CLGIPL and IPLL enzyme variants in comparison to wildtype AMDase. The resulting stereoselectivity is determined by the position of the catalytic cysteine residue (C74 vs. C188).

AMDase variant	Res. 43	Res. 74	Res. 125	Res. 156	Res. 159	Res. 188	Selectivity	Ref.
WT	V	G	A	V	M	C	(<i>R</i>)	133
CLGIPL	I	C	P	L	L	G	(<i>S</i>)	156
IPLL	I	G	P	L	L	C	(<i>R</i>)	42

Yet remarkably, when small vinylmalonic acid substrates – in particular 2-methyl-2-vinylmalonic acid (**7a**) – were studied, AMDase IPLL converted these compounds with unchanged high enantioselectivity, while enantioselectivity of the CLGIPL variant was impaired (98 %*ee* (*R*) vs. 66 %*ee* (*S*); Table 12).⁴⁶ This was quite unexpected, as both variants share all amino acid substitutions except for the catalytic cysteine (position 74 or 188) and a glycine at the respective other position (Table 16).

We were curious, if the observed behavior of the CLGIPL variant could be alleviated by modifying active site residues of the hydrophobic pocket or close to the reactive center. As to date no crystal structure of the CLGIPL variant is available, interesting residues were selected by visual inspection of the active site of the (*S*)-selective AMDase G74C/C188S variant (PDB: 3IXL) (Figure 20).

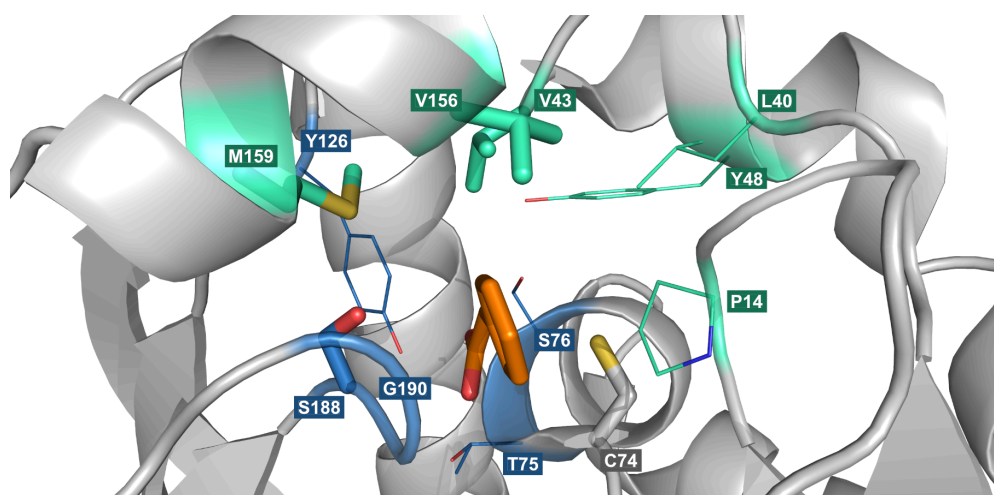


Figure 20. Representation of the active site of AMDase G74C/C188S in association with 2-phenyl acetate (orange) (PDB: 3IXL). The hydrophobic pocket is highlighted in cyan and the dioxanion hole in blue. The catalytic Cys74 and residues selected for site-directed mutagenesis (res. 43, 156, 159, 188, 190) are represented as sticks. Structure visualized with PyMOL.⁸⁸

From that, residues in the hydrophobic pocket and closest to the substrate were selected (Ile43, Leu156 and Leu159). These residues were exchanged with the original wildtype amino acid and with a series of other hydrophobic amino acids (Ala, Ile, Leu, Met, Phe, Val). Additionally, Gly188 seemed to be an interesting target, as it presents the only difference to AMDase IPLL. We envisaged an exchange to alanine or serine, even though, in the past, variants of position 188 proved extremely low in activity.^{139,148} Furthermore, the wildtype G190A variant was reported to convert 2-methyl-2-vinylmalonic acid with highest activity, which could be also beneficial in our studied case.

4.3.2 Generation of AMDase CLGIPL variants

Selected AMDase variants were generated by site-directed mutagenesis *via* a modified QuikChange protocol. Depending on the intended variant, different plasmids, based on pET28a[AMDase_CLGIPL], were used as a template for QuikChange (Table 17).

Table 17. Summary of selected AMDase CLGIPL variants, which were created by site-directed mutagenesis.

Variant	Mutations ^[a]	Template ^[b]	Variant	Mutations	Template ^[b]
1	I43A	CLGIPL	9	L159M	CLGIPL
2	I43F	CLGIPL	10	G188A	CLGIPL
3	I43L	CLGIPL	11	G188S	CLGIPL
4	I43M	CLGIPL	12	G190A	CLGIPL
5	I43V	CLGIPL	13	L159M/G190A	Variant 9
6	L156A	CLGIPL	14	G188A/G190A	Variant 12
7	L156I	CLGIPL	15	G188S/G190A	Variant 12
8	L156V	CLGIPL			

[a] All mutations are stated relative to AMDase CLGIPL (G74C/M159L/C188G/ V43I/A125P/V156L).

[b] Template refers to the respective pET28a expression construct used for site-directed mutagenesis.

Due to the overall high GC content of the AMDase CLGIPL gene (~60 %), a gradient was chosen for primer annealing (63-70°C). After analyzing aliquots of the QuikChange reaction by agarose gel electrophoresis (Figure 21 and

Appendix E, Figure S 1-Figure S 7), positive samples were digested with *DpnI*, pooled and used for transformation of *E. coli* TOP10.

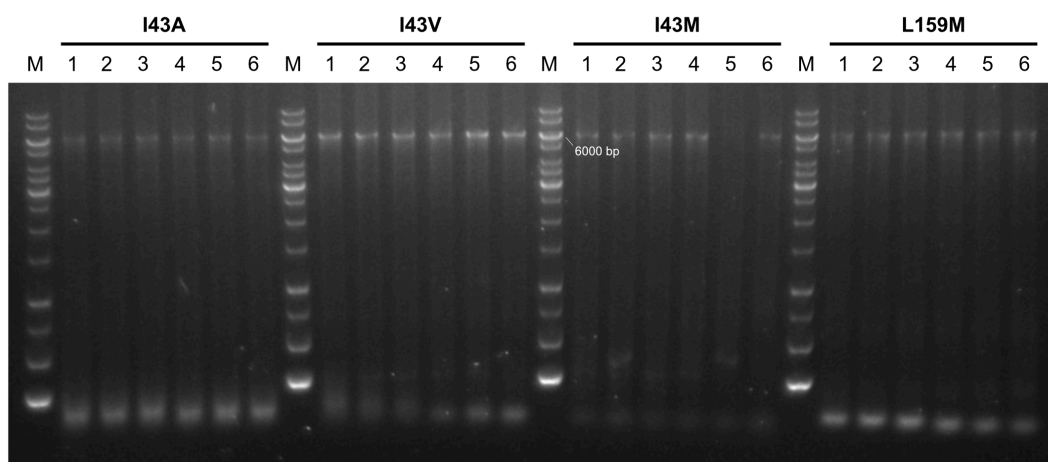


Figure 21. Agarose gel analysis of QuikChange reaction for generation of AMDase CLGIPL variants with temperature gradient ($T_1 = 63^\circ\text{C}$, $T_2 = 63.6^\circ\text{C}$, $T_3 = 65.3^\circ\text{C}$, $T_4 = 67.7^\circ\text{C}$, $T_5 = 69.3^\circ\text{C}$, $T_6 = 70^\circ\text{C}$). pET28a[AMDase_CLGIPL] (6005 bp) was used as template. M: GeneRuler 1 kb DNA Ladder (Thermo Scientific).

Transformants were selected and the corresponding plasmids analyzed by Sanger sequencing to confirm the incorporation of the desired mutations. Confirmed plasmid constructs were used to transform *E. coli* BL21(DE3), which was used as expression host.

4.3.3 Heterologous production and purification of AMDase and variants

Heterologous production of AMDase was already well established by taking advantage of the convenience of the pET vector series and thus the strong T7 promoter.^{42,44,46} The usability of T7-controlled heterologous expression systems was already discussed earlier in this work (see 3.3.2). As in previous reports, *E. coli* BL21(DE3) harboring pET28a[AMDase] was cultivated in LB-Kan⁴⁰ medium at 37°C up until an optical density (OD_{600}) of 0.5. Cultures were induced by addition of IPTG (1 mM final concentration) and further incubated at 30°C overnight. As can be seen from Figure 22, satisfactory amounts of AMDase were present in the cell-free lysate in a soluble form. It has to be noted that these results are merely presented by way of illustration and comparable results were obtained for all other studied variants of AMDase (Appendix F, Figure S 13-Figure S 17).

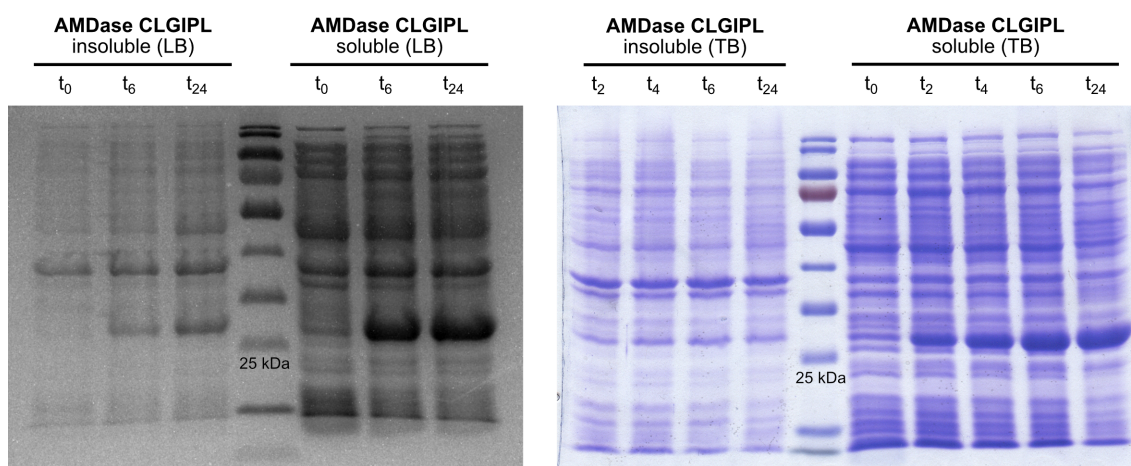


Figure 22. SDS-PAGE analysis of AMDase CLGIPL (MW = 25.8 kDa) expression study in LB (left) and TB medium (right). Samples were taken 0 h (t₀), 2 h (t₂), 4 h (t₄), 6 h (t₆) and 24 h (t₂₄) after induction. Soluble and insoluble fraction were separated after sonication. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

As AMDase was mostly produced in a soluble form under control of the T7 promoter, it was suspected that the amount of soluble enzyme could be further improved by using terrific broth (TB) medium for cultivation instead. Indeed, the overall yield of AMDase could be increased, mainly due to higher amounts of cell biomass obtained per cultivation volume (Figure 22; typically 10-15 g_{CWW} L⁻¹ in LB medium vs. 25-30 g_{CWW} L⁻¹ in TB medium). For storage purposes, the obtained biomass was washed with buffer before keeping the cell pellets at -20°C. Mostly, cell free extracts and purified enzyme were freshly prepared upon demand but could be stored at 4°C for several days, whereby the extent of loss of activity remained undetermined.

Routinely, cell-free extracts of AMDase and variants were prepared by suspending 75 mg cells per milliliter buffer before sonication. This procedure typically yielded cell-free extracts with a total protein content of 5-6.5 mg mL⁻¹.

If required, AMDase and variants were purified from the cleared lysates by Ni-affinity chromatography *via* a C-terminal His₆-tag. Standard protocols for purification proved adequate to isolate sufficient amounts of AMDase (Figure 23). Comparable results were obtained for other AMDase variants (Appendix F, Figure S 34-Figure S 36). Whenever necessary, the obtained purified fractions were further concentrated before use.

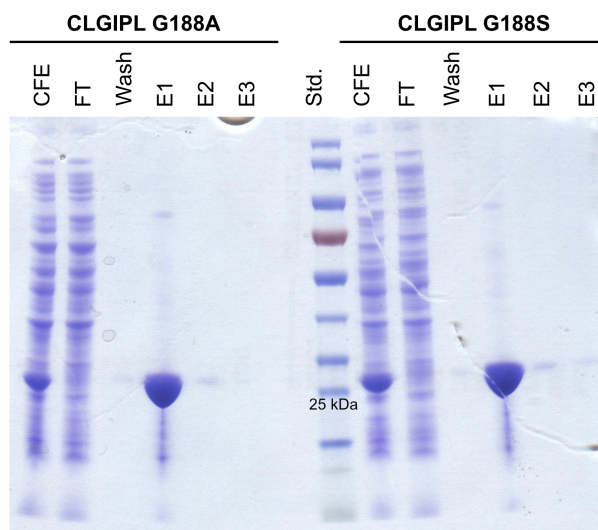


Figure 23. SDS-PAGE analysis of different fractions obtained from Ni-affinity purification of AMDase CLGIPL G188A (left) and G188S (right) (MW = 25.8 kDa). CFE: cell-free extract, FT: flow through, E: elution fraction. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

In the here presented example, 20.0 mg (G188A) and 22.4 mg (G188S) purified AMDase were obtained from 20.0 mL (5.6 mg mL^{-1} total protein) and 18 mL (5.4 mg mL^{-1} total protein) CFE, respectively, which accounts for 18 and 23 % AMDase of total protein. This further corresponds to a purification yield of around 320 mg AMDase per liter cultivation medium (Table 18).

Table 18. Yields of AMDase CLGIPL G188A and G188S obtained by purification *via* Ni-affinity chromatography from the respective cell free extracts. Purification yield (mg L^{-1} medium) was calculated based on 20 mL (G188A) and 18 mL (G188S) CFE obtained from 62 mL and 70 mL culture, respectively.

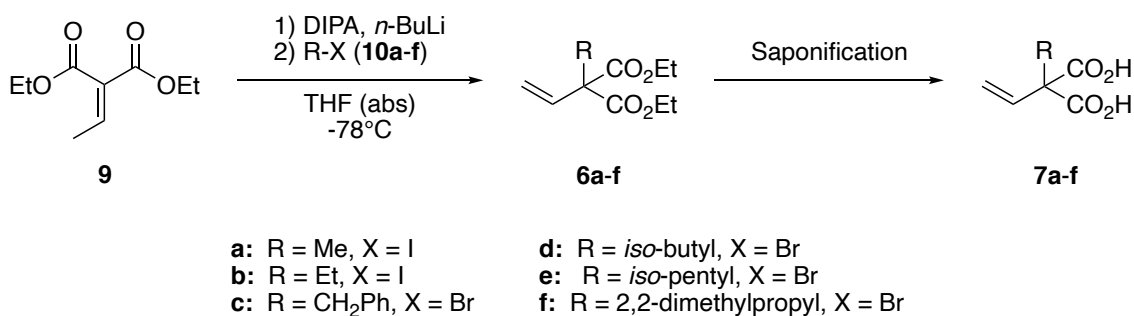
Variant	Purified protein (mg mL^{-1} CFE)	Total protein (mg mL^{-1} CFE)	PAD per total protein	Purification yield (mg L^{-1} medium)
G188A	1.0	5.6	18 %	322
G188S	1.2	5.4	23 %	320

4.3.4 Substrate synthesis

The 2-substituted 2-vinylmalonic acids **7a-f** were not commercially available and had to be synthesized. For this, the general synthetic strategy included intermediary formation of a diester, before the quite unstable malonic acid was released by deprotection. Previously, Gaßmeyer *et al.* reported the appealing

strategy of protecting arylmalonic acids as benzyl esters. The reaction conditions used for deprotection (*i.e.* hydrogenation) are mild enough to yield the exceedingly labile arylmalonic acids in satisfactory yields and without undesired spontaneous decarboxylation.⁴² Yet, for the substrate scaffold used in this work, this strategy is not applicable, due to presence of the terminal alkene, which would undergo reduction during hydrogenation. Therefore, the here applied approach is taking advantage of the generally higher stability of vinylmalonic acids compared to their aryl-counterparts, thus allowing use of ordinary ethyl esters, which can be cleaved by saponification under basic conditions (Scheme 25).

Scheme 25. General scheme for the synthesis of 2-substituted 2-vinylmalonic acids **7a-f** via alkylation of diethyl 2-ethylidenemalonate (**9**) and saponification of the corresponding diethyl esters **6a-f**. Different alkyl halides **10a-f** were used as alkylating agents.



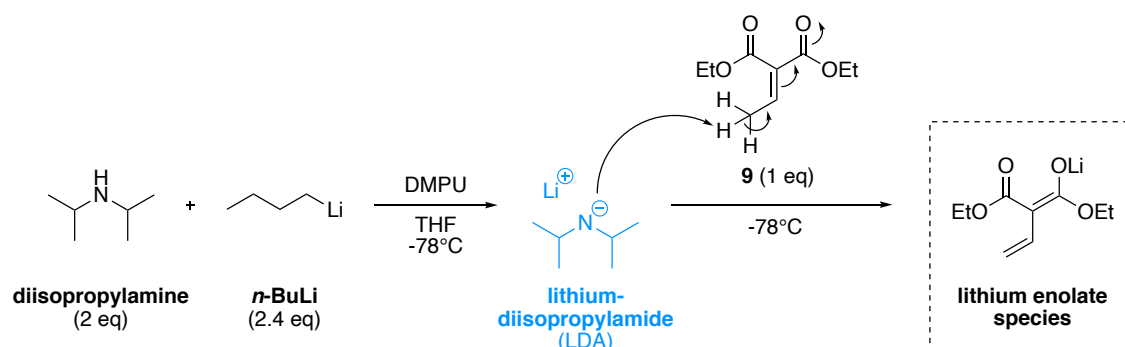
This route was already established previously for a series of 2-methyl-2-vinylmalonic acid derivatives with variations on the vinyl moiety.^{46,145} However, modifications of the alkyl sidechain had not been reported. In this work, we envisaged application of the same synthetic strategy, yet by using different alkyl halides as electrophile, different alkyl groups should be introduced (Scheme 25, first step).

4.3.4.1 General synthetic route towards **6a-f** and **7a-f**

Preparation of lithium diisopropylamide (LDA) followed previously reported procedures.⁴⁶ Under inert conditions, absolute THF and dry diisopropylamine (2 eq) were mixed and cooled to -78°C (acetone/dry ice). An excess *n*-BuLi (2.4 eq) was added dropwise followed by addition of absolute *N,N'*-Dimethylpropyleneurea (DMPU). DMPU is a popular additive in lithiation reactions and can modulate the reactivity of LDA. It was long thought that it acts

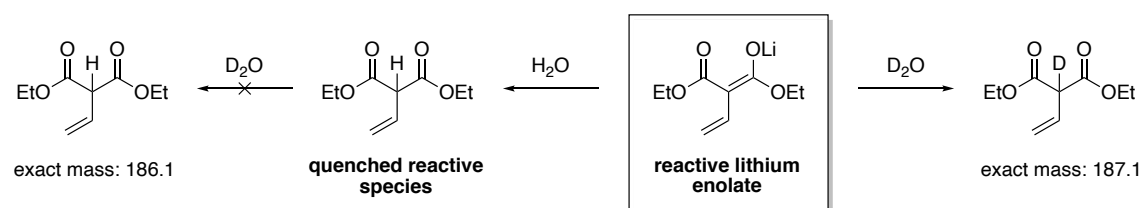
merely *via* deaggregation of organolithium clusters, but recent studies have rather drawn a more complex picture of the actual role of such additives.^{185,186} The formed LDA is able to abstract a proton from diethyl 2-ethylidenemalonate (**9**) in a kinetic fashion, which causes rearrangement into a reactive lithium enolate species (Scheme 26).

Scheme 26. Preparation of lithium diisopropylamide (LDA) (in blue) from diisopropylamine and *n*-BuLi in absolute THF at -78°C . *N,N'*-Dimethylpropyleneurea (DMPU) is added as cosolvent. LDA can further abstract a proton from diethyl 2-ethylidenemalonate (**9**) to yield the reactive lithium enolate species (in dashed box).



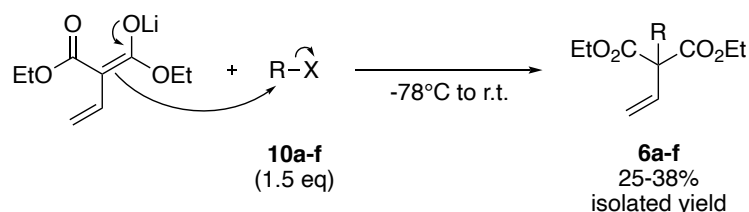
In the following reaction step, this lithium enolate can perform nucleophilic attack on an alkyl halide, overall resulting in alkylation of the central carbon atom (see 4.3.4.2-4.3.4.5). A simple approach proved useful for detecting residual reactive lithium enolate species in the reaction mixture (e.g. when incomplete conversion to the final products **6a-f** was observed). By quenching a small aliquot of the reaction mixture in D_2O , followed by small-scale work-up, remainders of reactive species would yield a deuterated product, whereas quenched reactive species (*i.e.* already protonated) would not react with D_2O . These two compounds could be easily differentiated by GC-MS analysis (Scheme 27).

Scheme 27. D_2O quench for detection of reactive lithium enolate species. Remaining reactive species in the reaction mixture reacts with D_2O and yields a deuterated product (right, exact mass: 187.1). Already quenched reactive species in the mixture cannot react with D_2O (left, exact mass: 186.1). Both quenched products can be differentiated by GC-MS analysis.



In case the lithium enolate was still present in the reaction mixture, more alkyl halide **10a-f** could be added to obtain full conversion of malonate **9**. Due to their high volatility, alkyl halides **10a-f** were easily purged by the nitrogen counterflow used for maintaining dry conditions in the reaction vessel, thus often resulting in incomplete conversion of the limiting diethyl 2-ethylidenemalonate (**9**) (Scheme 28).

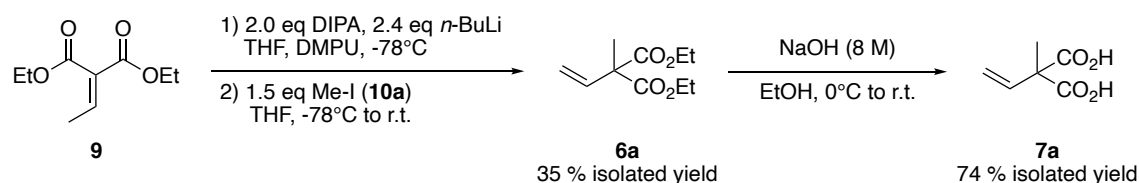
Scheme 28. Generalized reaction scheme for the preparation of diethyl 2-alkyl-2-vinylmalonates **6a-f** via nucleophilic attack of alkyl halides **10a-f** by the previously formed lithium enolate.



4.3.4.2 Synthesis of 2-methyl-2-vinylmalonic acid (**7a**)

The substrate 2-methyl-2-vinylmalonic acid (**7a**) was prepared as previously reported^{46,145} and the synthetic route is summarized in Scheme 29.

Scheme 29. Synthetic route towards 2-methyl-2-vinylmalonic acid (**7a**) in two steps.



Initially, LDA was prepared as described earlier (see 4.3.4.1), diethyl 2-ethylidenemalonate (**9**) was added (typically 1.0 g, 5.4 mmol) and the resulting mixture stirred at -78°C for 45 min. Then, iodomethane (**10a**) was added in a slight excess (1.5 eq) and the slightly turbid mixture was gradually warmed to room temperature overnight. Conversion was monitored by TLC and GC-MS and, in case of incomplete conversion, presence of reactive lithium enolate species was additionally confirmed by performing a D_2O quench and GC-MS analysis (Scheme 27 and Figure 24).

Typically, additional 0.5 eq of MeI (**10a**) and stirring overnight at r.t. were required to obtain full conversion. After work-up, column chromatography was required to purify the crude product. It has to be noted, that detection of product **6a** via TLC

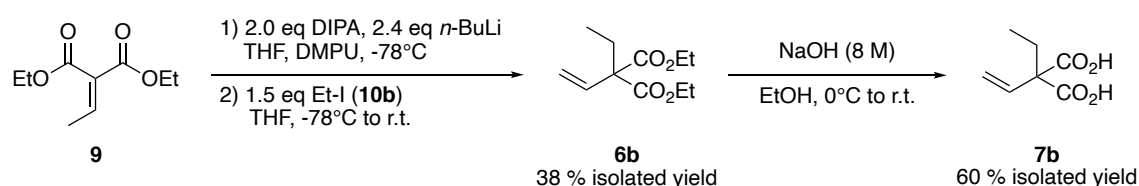
was quite difficult, due to weak staining with KMnO_4 and inability to detect with alternative staining reagents or UV-light. It can thus not be excluded, that a certain loss of product occurred during chromatographic purification. Further, diester **6a** is quite volatile, which complicates removal of solvents used for purification (cyclohexane and ethyl acetate). Hence, isolated yields ranged only between 25 % to 35 %, which is yet comparable to previous reports.^{46,145}

Saponification of diester **6a** was also done according to published procedures.⁴⁶ After dissolving **6a** in EtOH, an excess of 8 M NaOH was added in a dropwise fashion to the cooled solution (0°C). The mixture was stirred and allowed to warm to room temperature overnight. The suspension was acidified to pH 4 (HCl) and the solvents were evaporated. After extraction with ethyl acetate, crystallization of product **7a** was attempted as previously described, but all efforts remained unsuccessful. Instead, the extract was evaporated to dryness, leaving a slightly brownish solid as crude product behind. The crude was taken up in *n*-pentane which washed away impurities, while the polar diacid **7a** remained undissolved. The pentane layer was withdrawn by pipetting and the procedure was repeated 2-3 times until the retrieved solid product was almost colorless. This procedure furnished 2-methyl-2-vinylmalonic acid (**7a**) sufficiently pure in up to 74 % yield.

4.3.4.3 Synthesis of 2-ethyl-2-vinylmalonic acid (**7b**)

Similarly to the route executed for the synthesis of **7a**, synthesis of 2-ethyl-2-vinylmalonic acid (**7b**) was intended (Scheme 30).

Scheme 30. Synthetic route towards 2-ethyl-2-vinylmalonic acid (**7b**) in two steps.



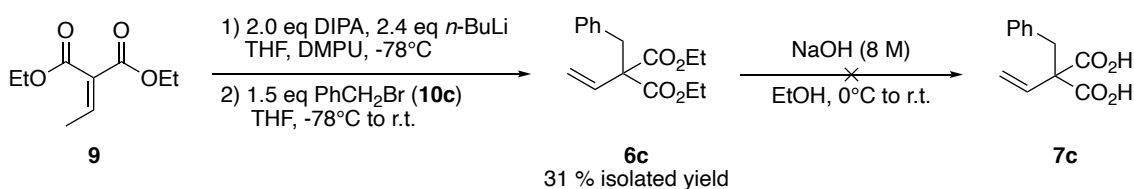
The experimental procedure followed the one described in Section 4.3.4.2, but this time using ethyl iodide (**10b**) as alkylating agent, correspondingly. As it was also the case for synthesis of **6a**, quantitative conversion of malonate **9** was not obtained when using 1.5 eq of **10b** and stirring overnight, but additional 0.5 eq of EtI (**10b**) were required. Again, full conversion was essential to partly compensate the moderate yields also observed during synthesis of **6a**. The target

compound **6b** could be isolated in 38 % yield as colorless oil after purification by column chromatography. Also, saponification of **6b** was straightforward and furnished the desired malonic acid **7b** as faint yellow solid in 60 % yield.

4.3.4.4 Synthesis of 2-benzyl-2-vinylmalonic acid (**7c**)

In accordance with the synthetic procedures applied for **7a-b**, synthesis of **7c** was attempted *via* the same route (Scheme 31). The overall intention was to prepare a substrate suitable to confirm the putative alternative binding mode, that is with an alkyl substituent too large to fit in the alkyl binding pocket during AMDase-catalyzed decarboxylation of vinylmalonic acids (Scheme 23).

Scheme 31. Synthetic route towards 2-benzyl-2-vinylmalonic acid (**7c**) in two steps.



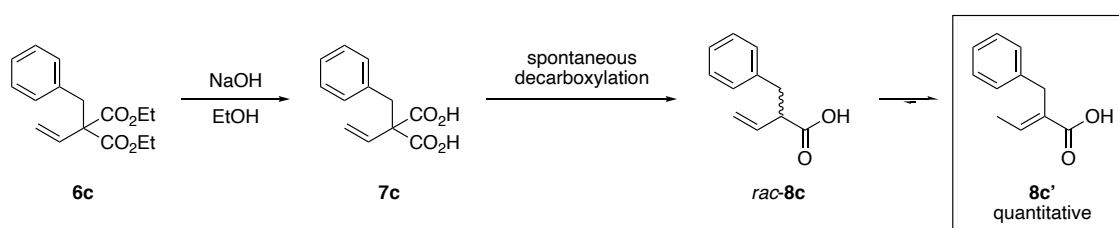
The aromatic substituent attached to the targeted compounds **6a** and **7a** proved advantageous when preparing these compounds, particularly during detection and purification (*i.e.* UV absorbance and staining behavior). Also, isolation of **6c** was less critical due to the considerably higher molecular weight compared to **6a-b**. Yet, also in this case, supplementation with additional 0.5 eq of benzyl bromide (**10c**) was necessary to reach full conversion. Overall, diester **6c** could be isolated in 31 % yield as a colorless oil after purification *via* flash column chromatography.

Saponification of **6c** was also attempted after the established procedure used for **6a-b**. The diester **6c** was dissolved in EtOH and 8 M NaOH was added to the cooled solution. After 5 h at room temperature, full conversion was detected by TLC analysis, however the observed retardation factor (R_f) of 0.54 of the formed products was quite unexpected. Typically, the R_f observed for diacids was lower than 0.2 under the applied conditions (CH/EA = 1:1). Further analyses *via* GC-MS, after derivatization with BSTFA, and NMR were also inconclusive.

In another attempt, a literature procedure from Arnold *et al.* was followed.¹⁸⁷ They reported successful saponification of **6c** in an aqueous solution of potassium

hydroxide, which was heated to reflux for 9 h and furnished the target compound **7c** in 98 % yield. Yet there was no reported evidence (*i.e.* NMR or mass analysis), that the isolated product was indeed the desired malonic acid **7c**. In our hands, this procedure furnished a single product, which however could not be identified as the desired product **7c** by NMR analysis. Another look into literature revealed, that deprotection of **6c** might be quite problematic. As early as in 1935, Ingold and Rogers reported that hydrolysis of **6c**, under either acidic or basic conditions, always furnished the decarboxylation product 2-benzylbut-3-enoic acid (*rac*-**8c**), which further isomerized to yield the more stable 2-benzylbut-2-enoic acid (**8c'**) (Scheme 32).¹⁸⁸

Scheme 32. Saponification of diethyl 2-benzyl-2-vinylmalonate (**6c**) to yield 2-benzyl-2-vinylmalonic acid (**7c**) followed by spontaneous decarboxylation to 2-benzylbut-3-enoic acid (*rac*-**8c**). The intermediately formed *rac*-**8c** completely isomerizes to 2-benzylbut-2-enoic acid (**8c'**).



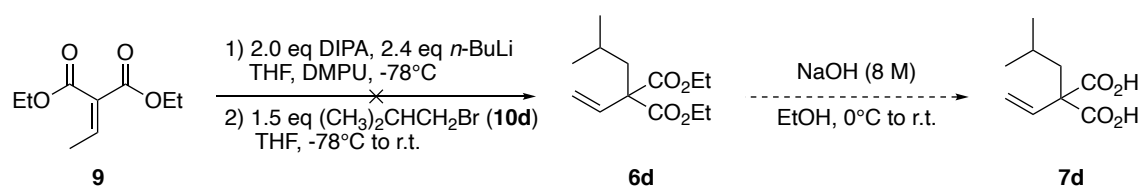
Indeed, NMR spectra obtained from the previously attempted saponification reactions strongly indicated presence of **8c'**. Particularly when employing the conditions reported by Arnold *et al.* (KOH, H₂O, reflux), diethyl ester **6c** was quantitatively converted to **8c'**, whereas mixtures of different compounds were obtained when using milder conditions (NaOH, EtOH, 0°C) (Appendix K, Figure S 64). Admittedly, the beforementioned publications suffer from incomplete analytical evidence from a current perspective but were yet the only reports found in literature dealing with compound **6c**. In particular, the instability of **6c** during hydrolysis observed by Ingold and Rogers¹⁸⁸ was consistent with the results obtained in this work. Due to the reported sensitivity of **6c** towards both acidic and basic hydrolytic conditions, the use of an entirely different protection strategy is advisable, which would further require synthesis of a different, suitable 2-ethylidenemalonate derivative as starting material.

4.3.4.5 Synthesis of other 2-substituted 2-vinylmalonic acids

As synthesis of our preferred substrate for studying the putative alternative binding mode of AMDase, 2-benzyl-2-vinylmalonic acid (**7c**), proved problematic *via* the established and here applied route, preparation of other compounds with larger alkyl-chain substituents (**7d-f**) was envisaged.

At first, synthesis of 2-isobutyl-2-vinylmalonic acid (**7d**) was attempted according to the established procedure (Scheme 33).

Scheme 33. Synthetic route towards 2-isobutyl-2-vinylmalonic acid (**7d**) in two steps.



As also previously observed, after stirring overnight, residual starting material was still detected by GC-MS. Additional **10d** (1 eq) was added after 24 h, and the reaction was again left overnight. After 48 h, conversion was checked again with GC-MS. Additionally, a small aliquot was quenched with D₂O in order to evaluate the presence of reactive species (Figure 24).

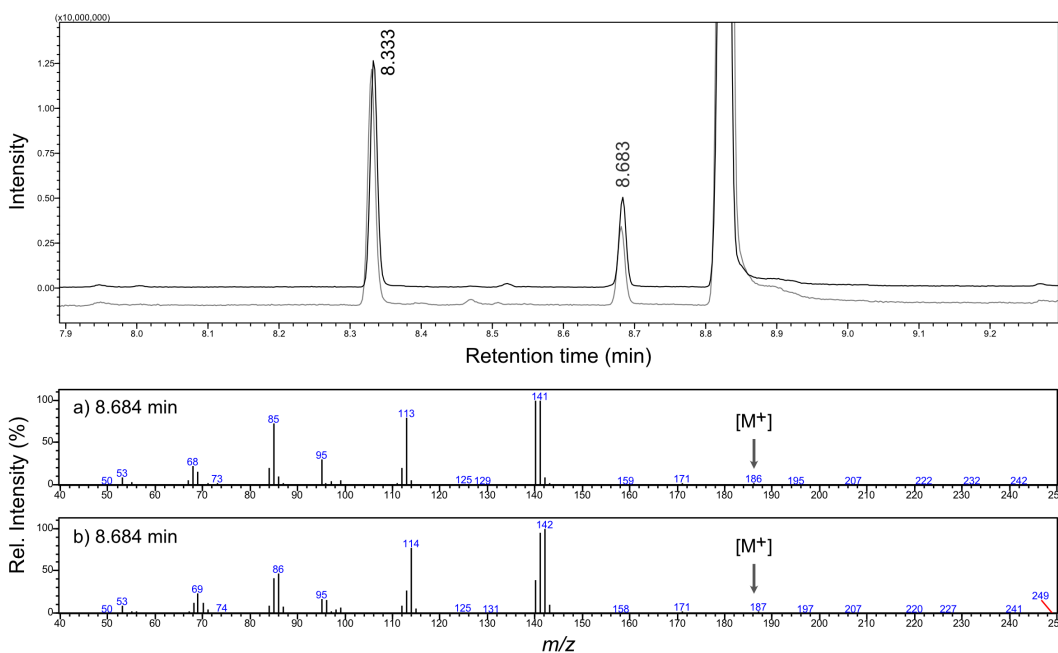


Figure 24. Top: TIC obtained from GC-MS analysis of the reaction mixture of **6d** after 48 h and aqueous (black) workup or D₂O quench (gray). Bottom: Mass spectra at 8.684 min corresponding to the undeuterated (a) and deuterated (b) starting material **9**.

Due to considerable amounts of (still reactive) starting material **9** in the reaction mixture, it was decided to leave the mixture for additional 72 h at room temperature. Yet, after aqueous workup and subsequent analysis of the obtained extract by GC-MS, the peak which supposedly corresponded to our target product **6d** ($t_R = 8.3$ min) could not be detected anymore. Instead, a series of unidentified compounds obviously arose from the workup procedure ($t_R > 9$ min) (Figure 25).

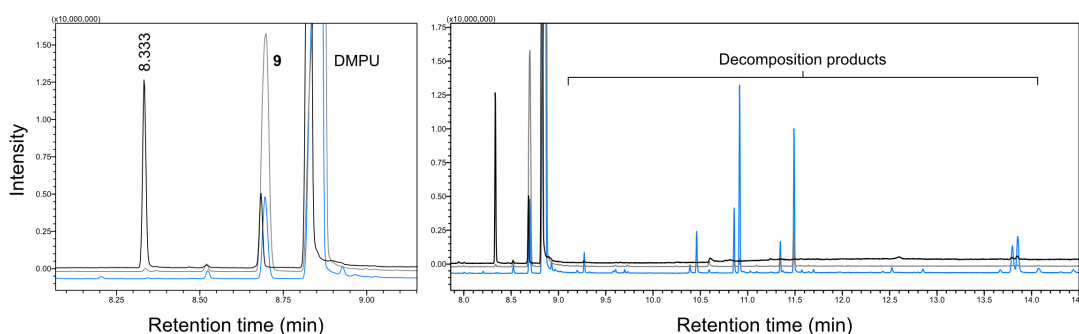
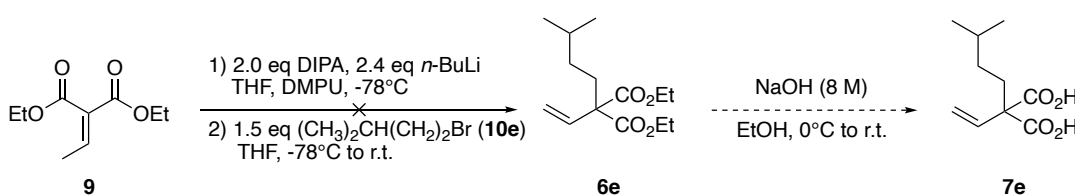


Figure 25. TIC obtained from GC-MS analysis of the reaction mixture of **6d** after 24 h (gray), 48 h (black) and after workup (blue). Left: While a clear decrease of **9** is observed between 24 and 48 h, the presumed product peak ($t_R = 8.3$ min) vanished after workup. Right: Various unidentified decomposition products were detected after workup.

Alternatively, we envisaged synthesis of 2-isopentyl-2-vinylmalonic acid (**7e**) with a slightly larger substituent (Scheme 34).

Scheme 34. Synthetic route towards 2-isopentyl-2-vinylmalonic acid (**7e**) in two steps.



Routinely, conversion was monitored by GC-MS after 24 h, but again, incomplete conversion of the starting material **9** required addition of supplementary **10e** (1 eq). Quite interestingly, also this time, the peak at 8.3 min, which was assumed to correspond to product **6d**, was also monitored. This observation made the previous peak assignment highly doubtful, but yet no other peaks were detected in the reaction mixture of **6d** before the workup. It has to be noted that in general, assignment of peaks solely due to mass spectra obtained from compounds like **6a-f** is challenging as they readily fragment, and smaller fragments of the core

structure are similar for this series of compounds **6a-f** and also the starting material **9**. It is thus also likely, that the commonly observed peak at 8.3 min (Figure 24-Figure 26) can be correlated to a structural derivative of **9**. After 48 h, relative peak intensities of starting material **9** and the peak at 8.3 min were switched compared to 24 h (*i.e.* “less” conversion after 48 h), thus making the aforementioned presumed misinterpretation highly likely.

As also observed for **6d**, a plethora of additional products could be detected following reaction workup. Careful comparison of the GC-MS chromatograms obtained after 24 h, 48 h and workup revealed that only two peaks ($t_R = 9.3$ min and 11.4 min) were already significantly present before workup (after 48 h) and could thus possibly be connected to product formation (Figure 26).

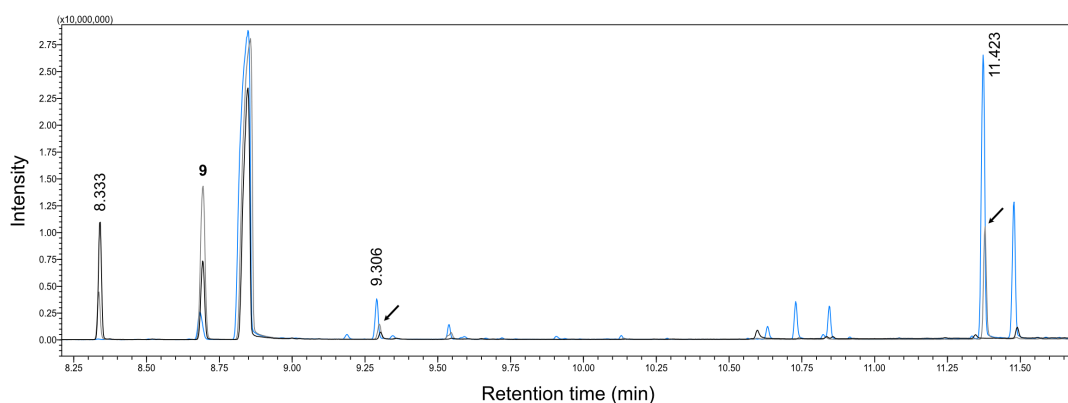
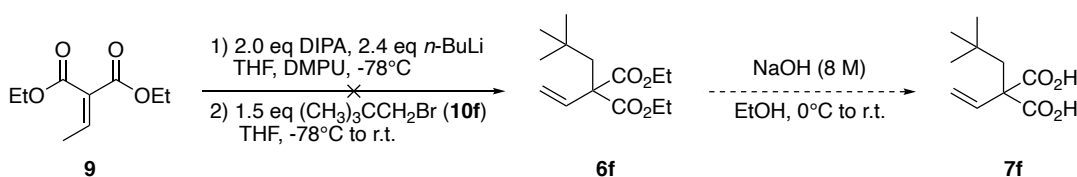


Figure 26. TIC obtained from GC-MS analysis of the reaction mixture of **6e** after 24 h (black), 48 h (gray) and after workup (blue). Peaks marked with an arrow (at $t_R = 9.3$ min and 11.4 min) significantly increased after 48 h and workup.

Yet, the respective mass spectra were inconclusive and the presence of **6e** could not be confirmed. In any case, only a low degree of conversion was reached, which would, along with the observed formation of unidentified decomposition products during workup, impede successful isolation of target product **6e**. In order to counter potential instability issues faced during synthesis **6d-e**, the strategy was slightly adapted for preparation of **6f** (Scheme 35).

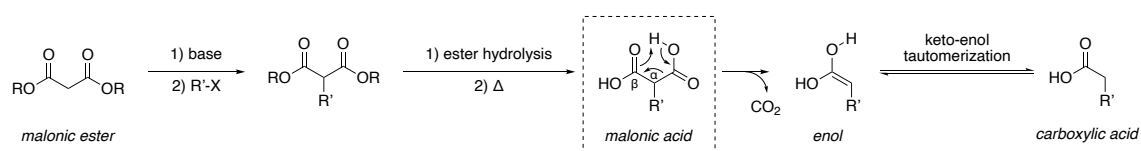
Scheme 35. Synthetic route towards 2,2-dimethylpropyl-2-vinylmalonic acid (**7f**) in two steps.



During stirring the reaction mixture overnight, the reaction vessel was protected from light, due to the potential decomposition of alkyl bromides when exposed to light. Additionally, a sudden rise of temperature during reaction or workup steps was particularly avoided by slow addition of reagents at constant cooling or cooling during quenching the reaction mixture. When incomplete conversion was detected after 24 h, the reaction mixture was supplemented with only 0.5 eq of **10f**. Yet again, only peaks at 8.3 min and 8.7 min (malonate **9**) were detected and potential product formation could not be confirmed.

Considering the potential instability of **6d-f** towards the applied conditions during reaction workup, that is quenching with saturated NH_4Cl solution, sufficing stability of these compounds at conditions required for ester cleavage to form **7d-f** is doubtful. Issues of this kind were already faced during attempted saponification of **6c** (see 4.3.4.4). It has to be noted that the ease of transferring substituted malonic esters into their relatively unstable malonic acid derivatives is synthetically exploited for the synthesis of substituted carboxylic acids (Scheme 36).

Scheme 36. General reaction scheme of a malonic ester synthesis. Initially, the ester is deprotonated by a strong base, followed by nucleophilic attack of an alkyl halide ($\text{R}'\text{-X}$) electrophile ($\text{S}_{\text{N}}2$). Aqueous ester hydrolysis takes place under acidic conditions, followed by decarboxylation under heating. Decarboxylation occurs *via* a six-membered cyclic transition state to yield an instable enol intermediate which readily tautomerizes into a carboxylic acid.

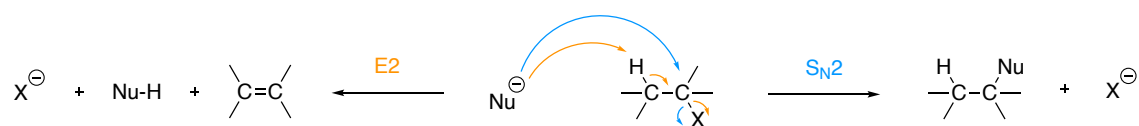


In this reaction sequence termed “malonic ester synthesis”, malonic esters are initially alkylated in a $\text{S}_{\text{N}}2$ -type reaction. This step can also be repeated to obtain (unsymmetric) disubstituted malonic esters. After ester hydrolysis at acidic conditions, the formed malonic acid can undergo decarboxylation at elevated temperatures *via* a six-membered cyclic transition state. The formed enol is relatively instable and readily undergoes tautomerization to yield a substituted carboxylic acid. While the ease of decarboxylation depends on the overall stability of the malonic acid, it has been demonstrated that the electronic nature of

introduced substituents is at least not entirely responsible, due to the relative non-polar character (*i.e.* little to no separation of charges) of this transition state.¹⁸⁹ Yet, in the present case of compound **6d-f**, it seems much more plausible that the nucleophilic substitution itself is largely hindered by the constitution of the used electrophilic alkyl halides **10d-f**. In an S_N2-type reaction, accessibility of the electrophilic carbon is essential due to back-side nucleophilic attack and loss of the leaving group in a concerted fashion.¹⁹⁰ Dostrovsky and Hughes already described a decreasing reactivity within a series of alkyl bromides with increasing substituent size and number following methyl > ethyl > *n*-propyl > *iso*-butyl >> neopentyl (*i.e.* 2,2-dimethylpropyl). While the relative rate decrease was according to expectations within the first four members of this series with 18 (Me): 1 (Et) : 0.28 (*n*Pr) : 0.030 (*i*Bu), the relative rate observed for neopentyl bromide (4.2×10^{-6}) was extraordinarily low.¹⁹¹ This low reactivity renders neopentyl bromide virtually inert against nucleophilic attack in a S_N2 reaction. Indeed, this fact is detrimental for the attempted synthesis of **6d-f** *via* the previously established route. While the synthesis of 2-methyl- (**6a**) and 2-ethyl-2-vinylmalonic acid (**6b**) was successful, formation of **6d-f** could not be confirmed, probably due to the aforementioned low reactivity of the employed alkyl halides **10d-f**. Further, in case of **6a-b**, the respective alkyl iodides **10a-b** were used, whereas only the corresponding, less reactive, alkyl bromides **10d-f** were available for the synthesis of **6d-f**.¹⁹⁰ Interestingly, synthesis of 2-benzyl-2-vinylmalonic acid (**6c**) was successful despite the quite bulky benzyl substituent. Yet, a quite inconvenient behavior of benzyl substituents was described in literature as “benzylic effect”, where the presence of a benzyl group can lead to enhanced rates compared to alkyl groups despite their bulkiness.¹⁹⁰

It also has to be kept in mind that E2-type elimination is a competitive reaction of S_N2-type nucleophilic substitutions. This competing reactivity is particularly promoted in cases where the α-carbon is difficult to access for nucleophilic attack. Instead, the nucleophile could abstract a better accessible proton from β-C which would furnish an alkene after elimination of the leaving group (Scheme 37).

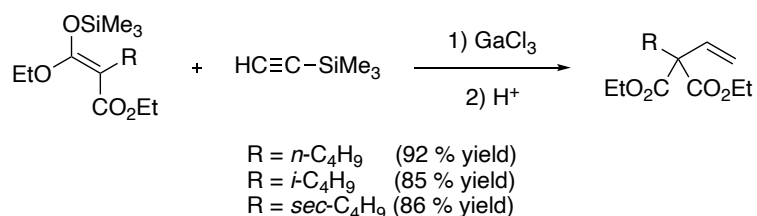
Scheme 37. Competition between nucleophilic substitution (S_N2 , in blue) and elimination (E2, in orange). During S_N2 the nucleophile (Nu^-) attacks the electrophilic carbon thereby inducing loss of the leaving group (X), whereas during E2, the nucleophile instead abstracts a proton at β -C and leads to formation of an alkene.



Yet, the actual extent of E2 elimination in the attempted synthesis of **6d-f** is purely speculative, as Nu-H would be also obtained from unreacted lithium enolate species after aqueous workup and the formed alkenes could neither be detected by TLC or GC-MS.

In view of the many inconclusive hurdles along the here attempted synthetic pathway towards malonic acids **7c-f**, establishment of an entirely different alternative synthesis strategy seems inevitable. Yet, evidence of compounds **6d-f** and **7d-f** in literature is, if at all, scarce and particularly the behavior of **6d-f** during ester hydrolysis and stability of the resulting malonic acids **7d-f** is unclear. Above-mentioned reactivity issues of alkyl halides **10d-f** towards nucleophilic attack generally prevent application of “classical” substitution chemistry. For instance, solely one specific synthesis was described for butylated 2-vinylmalonates in literature. In this approach, silyl enol ethers were ethenylated by trimethylsilylethyne in presence of $GaCl_3$ (Scheme 38).¹⁹² While this procedure comprises specialized chemistry, success of subsequent malonic ester hydrolysis is yet uncertain.

Scheme 38. Synthesis of diethyl 2-alkyl-2-vinylmalonates *via* ethenylation of silyl enol ethers by trimethylsilylethyne in presence of $GaCl_3$. Scheme adapted from Arisawa *et al.*¹⁹²



In view of the aforementioned issues faced during synthesis of alternative AMDase substrates **7c-f**, pursuing investigations regarding characterization of AMDase variants were limited to 2-methyl- (**7a**) and 2-ethyl-2-vinylmalonic acid (**7b**).

4.3.5 Characterization of enantioselectivity of AMDase CLGIPL variants

Before further characterizing the decarboxylation of 2-alkyl-2-vinylmalonic acids **7a-b** catalyzed by the newly created AMDase CLGIPL variants, activity of AMDase CLGIPL towards 2-methyl-2-vinylmalonic acid (**7a**) needed to be confirmed. For this purpose, decarboxylation of **7a** was attempted with cell-free extract containing AMDase CLGIPL as previously described (Table 19, entry 1).⁴⁶ Alternatively, the same reaction was also performed with purified AMDase CLGIPL (Table 19, entry 2).

Table 19. Comparison of product ee obtained when using cell-free extracts (CFE) or purified enzyme for AMDase-catalyzed decarboxylation of 2-methyl-2-vinylmalonic acid (**7a**). Reaction conditions: 10 mM **7a** in 50 mM Tris-HCl (pH 8), CFE or purified enzyme, 30°C, 600 rpm. Results obtained from chiral GC-FID analyses are presented in Appendix I (Figure S 41-Figure S 42).

Entry	AMDase variant	Enzyme preparation	Enzyme amount	Reported selectivity	Experimental ee
1	CLGIPL	CFE	50 % (v/v)	(S) ¹⁵⁶	72 % (S)
2	CLGIPL	purified	1 mg mL ⁻¹	(S) ¹⁵⁶	71 % (S)
3	IPLL	CFE	5 % (v/v)	(R) ⁴²	98 % (R)

These initial results clearly demonstrated that undesired side-reactivities originating from native *E. coli* proteins present in the cell lysate are insignificant under the studied conditions. The observed enantiomeric excess of the obtained reaction product 2-methylbut-3-enoic acid (**8a**) was similar with 72 %ee (S) and 71 %ee (S) in case of reactions performed with CFE and purified AMDase CLGIPL, respectively. As an additional control reaction, and to confirm the different enantioselectivity, decarboxylation of **7a** was also performed with the (R)-selective IPLL variant (Table 19, entry 3). Indeed, these results are consistent with previously published data (98 %ee (R) for IPLL and 66 %ee (S) for CLGIPL; Table 12, entry 1-2).⁴⁶

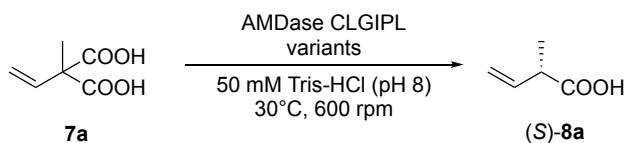
4.3.5.1 Conversion of 2-methyl-2-vinylmalonic acid (**7a**)

In a first round of site-directed mutagenesis, several variants of AMDase CLGIPL were created by modification of residue I43, L156, L159 and G190. Instead of Ile43, other hydrophobic amino acids were introduced, namely alanine, leucine,

methionine and valine (WT). The same strategy was followed for Leu156, where alanine, isoleucine and valine (WT) were introduced instead. These two positions were considered quite important, as they are situated right “on top” of the substrate binding site and are central constituents of the hydrophobic network covering the active site (Figure 20). Leu159 was exchanged to the wildtype residue methionine. While this residue is also part of the hydrophobic pocket, it is located in a less central position of the cavity. Further, we were interested to additionally introduce the mutation G190A into the CLGIPL scaffold, as Micklefield and coworkers reported the superior activity of AMDase G190A towards the smallest substrate **7a**.¹⁴⁵ These initially created variants of AMDase CLGIPL were tested in the conversion of 2-methyl-2-vinylmalonic acid (**7a**). Decarboxylation was initiated by addition of 50 % (v/v) cell-free extract (75 mg_{cww} mL⁻¹) and the reaction products **8a** were extracted from the mixture as soon as full conversion was detected by TLC. The extracts were then directly analyzed by chiral GC-FID without prior derivatization as previously established.⁴⁶ This procedure ensured comparable results obtained from chiral analyses and further prevented issues caused by residual substrate, as will be discussed later in this section.

Variants of position 43 were generally comparably active as AMDase CLGIPL and full conversion was observed after about 20 h in all cases (Table 20, entry 2-5). While the enantioselectivity of variants I43A and I43V was identical to CLGIPL (72 %ee (S); Table 20, entry 2-3), a slightly higher selectivity was observed for I43M and I43L (77 %ee (S) and 83 %ee (S); Table 20, entry 4-5). From these results, a generally broad tolerance of different amino acids at position 43 was deduced and an increased optical purity of reaction products was observed when introducing larger side chains (Leu, Met > Ala, Val). Following from this, it was hypothesized, that by introducing even larger, hydrophobic amino acids at this position (e.g. phenylalanine), the product ee might be further enhanced. Yet, the activity observed for the I43F variant was extremely low, which prevented further characterization due to overall low levels of conversion (Table 20, entry 6).

Table 20. Summarized results obtained from conversion of 2-methyl-2-vinylmalonic acid (**7a**) by AMDase CLGIPL (C₇₄L₁₅₉G₁₈₈I₄₃P₁₂₅L₁₅₆) and several variants thereof. Reaction conditions: 10 mM **7a** in 50 mM Tris-HCl (pH 8), 50 % (v/v) CFE, 30°C, 600 rpm. n.d.: not determined. Relative activity is divided in higher (++++), similar (+++), lower (++) and negligible (+) activity compared to CLGIPL.



Entry	AMDase variant	ee	Relative activity	Full conversion ^[a]
1	C ₇₄ L ₁₅₉ G ₁₈₈ I ₄₃ P ₁₂₅ L ₁₅₆	72 % (S)	+++	20 h
2	C ₇₄ L ₁₅₉ G ₁₈₈ A ₄₃ P ₁₂₅ L ₁₅₆	72 % (S)	+++	20 h
3	C ₇₄ L ₁₅₉ G ₁₈₈ V ₄₃ P ₁₂₅ L ₁₅₆	72 % (S)	+++	20 h
4	C ₇₄ L ₁₅₉ G ₁₈₈ M ₄₃ P ₁₂₅ L ₁₅₆	77 % (S)	+++	20 h
5	C ₇₄ L ₁₅₉ G ₁₈₈ L ₄₃ P ₁₂₅ L ₁₅₆	83 % (S)	+++	19 h
6	C ₇₄ L ₁₅₉ G ₁₈₈ F ₄₃ P ₁₂₅ L ₁₅₆	n.d. ^[b,c]	+	-
7	C ₇₄ L ₁₅₉ G ₁₈₈ I ₄₃ P ₁₂₅ A ₁₅₆	72 % (S)	++	42 h
8	C ₇₄ L ₁₅₉ G ₁₈₈ I ₄₃ P ₁₂₅ V ₁₅₆	61 % (S)	+++	19 h
9	C ₇₄ L ₁₅₉ G ₁₈₈ I ₄₃ P ₁₂₅ I ₁₅₆	70 % (S)	+++	19 h
10	C ₇₄ M ₁₅₉ G ₁₈₈ I ₄₃ P ₁₂₅ L ₁₅₆	72 % (S) ^[d]	++	> 25 h ^[e]
11	C ₇₄ L ₁₅₉ G ₁₈₈ I ₄₃ P ₁₂₅ L ₁₅₆ + G190A	45 % (S)	++++	20 h
12	C ₇₄ M ₁₅₉ G ₁₈₈ I ₄₃ P ₁₂₅ L ₁₅₆ + G190A	62 % (S)	++	42 h
13	C ₇₄ L ₁₅₉ A ₁₈₈ I ₄₃ P ₁₂₅ L ₁₅₆	69 % (S) ^[c,d]	+	-
14	C ₇₄ L ₁₅₉ S ₁₈₈ I ₄₃ P ₁₂₅ L ₁₅₆	81 % (S) ^[c,d]	+	-
15	C ₇₄ L ₁₅₉ A ₁₈₈ I ₄₃ P ₁₂₅ L ₁₅₆ + G190A	n.d. ^[b,c]	+	-
16	C ₇₄ L ₁₅₉ S ₁₈₈ I ₄₃ P ₁₂₅ L ₁₅₆ + G190A	n.d. ^[b,c]	+	-

[a] As judged by TLC. [b] Not further studied due to extremely low activity and resulting low levels of conversion. [c] Incomplete conversion can cause decomposition of malonate in GC injection unit, which produces racemic product, thereby negatively affecting product ee. [d] Reaction products were separated from remaining substrate *via* semi-preparative TLC for determination of ee. [e] During reevaluation of product ee *via* semi-preparative TLC, small residues of substrate were detected.

In the case of residue Leu156, introduction of an alanine residue was less tolerated, which was reflected in the lower activity as compared to the original CLGIPL variant (Table 20, entry 7). In contrast, by exchanging the present leucine residue to valine or isoleucine, enzyme activity was not adversely affected (Table 20, entry 8-9). Moreover, no obvious trend regarding the influence on enantioselectivity could be observed with similar results obtained for L156A and L156I (72 %ee (S) and 70 %ee (S); Table 20, entry 7 and 9) and only 61 %ee (S) for L156V (Table 20, entry 8). Interestingly, an exchange of Leu159 to methionine, as present in AMDase wildtype, just resulted in a loss of activity with almost unchanged enantioselectivity (72 %ee (S); Table 20, entry 10). In this particular case of the L159M variant, residual substrate was detected in the reaction extract during reevaluation of product ee. It should be thus noted, and will be also discussed later in this section, that the here stated ee was obtained after purification of the reaction products *via* semi-preparative TLC.

The positive impact of the mutation G190A onto AMDase activity towards 2-methyl-2-vinylmalonic acid (**7a**) was also observed when introducing this additional mutation to AMDase CLGIPL. Okrasa *et al.* previously observed a 4.6-fold increase in activity during conversion of **7a** compared to the WT, while maintaining perfect product enantiopurity of 99 %ee (R).¹⁴⁵ Yet in the here presented case, the observed optical purity of 45 %ee (S) was impaired compared to CLGIPL (Table 20, entry 11). This rather strong influence was yet unexpected and thus subject of supplementary studies (see 4.3.5.3). In view of the low activity of the L159M variant, we were also curious, if combined with the G190A mutation, activity towards decarboxylation of **7a** could be improved. Yet also then, both activity and enantioselectivity were only moderate (62 %ee (S); Table 20, entry 12).

As the here studied variations of residues mainly within the hydrophobic pocket had little to no effect on enantioselectivity, we were curious how an exchange of residue Gly188 would affect the product ee. This position seems to be a particularly interesting target, as it virtually presents the only difference between CLGIPL (C74 and G188) and IPLL (G74 and C188) variant. Yet, previously reported variations at this site (*e.g.* Ala or Ser) proved detrimental to activity.^{139,148} Indeed, also in the present work, activity of both CLGIPL G188A and G188S variants was extremely low. Under standard conditions (*i.e.* 50 % (v/v) CFE), only

low levels of conversion were detected and reactions were discontinued after 9 d, where residual enzymatic activity in the CFEs was already unlikely. Also, when the experiment was repeated with quite high dosage of purified enzyme (1 mg mL^{-1}), full conversion was not reached within 4 d, though G188S was unambiguously more active than G188A.

As mentioned earlier, remainders of substrate in the reaction extracts might pose a risk during GC analysis, as the malonic acid could decarboxylate spontaneously in the injection unit of the gas chromatograph, thus resulting in distortion of the product enantiopurity. Indeed, this was confirmed during chiral GC-FID analysis of malonic acid **7a**, which furnished an almost racemic mixture of **8a** (22 %*ee* (*R*)) and residual **7a** (Figure 27).

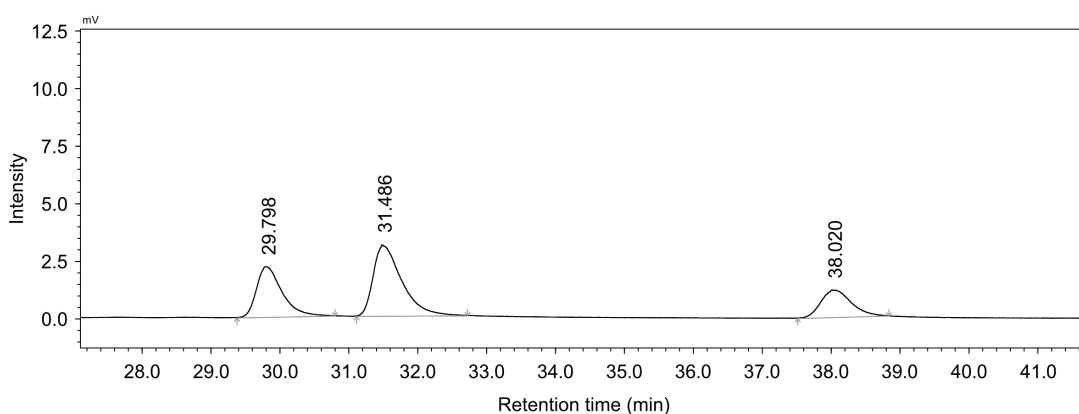


Figure 27. Chromatogram obtained from chiral GC-FID analysis of 2-methyl-2-vinylmalonic acid (**7a**). Due to thermally induced decarboxylation in the injection unit, malonic acid **7a** appears as almost racemic mixture of **8a** ((*S*)-**8a**: $t_R = 29.8 \text{ min}$, (*R*)-**8a**: $t_R = 31.5 \text{ min}$; 22 %*ee* (*R*)) and residual **7a** ($t_R = 28.0 \text{ min}$).

To address this issue, the obtained reaction extracts were purified by semi-preparative TLC which was easily possible due to considerable differences in polarity (Appendix G, Table 60). The retrieved purified products were then subjected to chiral GC-FID analysis. Particularly in case of G188A, the impact of residual substrate was prominent. While in the non-purified extract the optical purity was determined to range only between 5-17 %*ee* (*S*), it was increased to 69 %*ee* (*S*) after purification, thus strongly indicating presence of a quite large portion of unreacted **7a** in the reaction extract (Table 20, entry 13). Although a similar effect was observed for G188S, where the optical purity was increased from 71-74 %*ee* (*S*) to 81 %*ee* (*S*) upon purification, it was way less pronounced due to overall higher levels of conversion (Table 20, entry 14).

Eventually, combination of the G188A or G188S variant with the G190A mutation was attempted to potentially increase activity. Yet, also in these cases, this hypothesis did not prove true, and the observed activity was even lower than for the corresponding single mutants. Hence, the mentioned G188A/G190A and G188S/G190A variants were not investigated further (Table 20, entry 15-16).

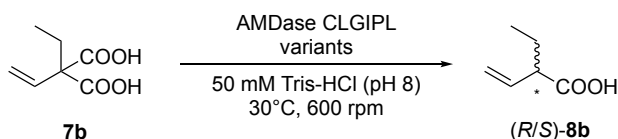
4.3.5.2 Conversion of 2-ethyl-2-vinylmalonic acid (**7b**)

In order to study the presence of a putative alternative binding mode during decarboxylation of 2-methyl-2-vinylmalonic acid (**7a**), it was originally intended to synthesize a series of similar substrates **7b-f**, but with a larger alkyl substituent in place of the methyl group. It was hypothesized, that such a substrate scaffold would trigger binding in the alternative binding mode, that is with the vinyl group accommodated in the small alkyl binding pocket, as accommodation of the larger alkyl substituent is prevented by spatial limitation. This binding mode would lead to the formation of (*R*)-enantiomers from the typically (*S*)-selective CLGIPL variants. Yet, preparation of these substrates was largely prevented by too low reactivity of the envisaged alkyl halides in a S_N2 reaction (see 4.3.4.5), and only 2-ethyl-2-vinylmalonic acid (**7b**) was successfully prepared from this series of substrates.

Initially, conversion of **7b** was studied with AMDase WT, IPLLL, CLGIPL and first-generation variants thereof (Table 21). Quite interestingly, while AMDase CLGIPL converted **7b** as expected, also the IPLLL variant showed some activity towards this compound (Table 21, entry 2-3). Wildtype AMDase however did not accept **7b** as a substrate (Table 21, entry 1). This latter observation was not very surprising, as the related aromatic substrate 2-ethyl-2-phenylmalonic acid was neither converted by AMDase WT,¹³¹ thus underlining the inability to accommodate substituents larger than a methyl group in the alkyl binding pocket of the wildtype enzyme. Quite remarkably, the AMDase racemase variant G74C possessed some activity towards this compound (0.05 U mg^{-1})^{142,157} and the activity could be almost increased five-fold to 0.24 U mg^{-1} by additionally introducing the V43A mutation.¹⁵⁷ Considering the extensive modifications of active site residues present in IPLLL, CLGIPL and variants thereof, accommodation of either a vinyl or ethyl group in the alkyl binding pocket seems plausible as well. It could be further assumed that a vinyl group, compared to an

ethyl group, should be slightly more compact, due to shorter C=C bond distance, and overall less flexible, thus conceivably slightly favoring its accommodation over an ethyl moiety.

Table 21. Summarized results obtained from conversion of 2-ethyl-2-vinylmalonic acid (**7b**) by AMDase WT and several variants (first generation). Reaction conditions: 10 mM **7b** in 50 mM Tris-HCl (pH 8), 50 % (v/v) CFE, 30°C, 600 rpm. Relative activity is divided in higher (++++), similar (+++), lower (++) and negligible (+) activity compared to CLGIPL.



Entry	AMDase variant	Relative activity	Conversion ^[a]
1	WT	-	No conversion
2	IPLL	++	Incomplete (48 h)
3	C ₇₄ L ₁₅₉ G ₁₈₈ I ₄₃ P ₁₂₅ L ₁₅₆	+++	Complete (42 h)
4	C ₇₄ L ₁₅₉ G ₁₈₈ A ₄₃ P ₁₂₅ L ₁₅₆	++++	Complete (24 h)
5	C ₇₄ L ₁₅₉ G ₁₈₈ V ₄₃ P ₁₂₅ L ₁₅₆	++++	Complete (24 h)
6	C ₇₄ L ₁₅₉ G ₁₈₈ M ₄₃ P ₁₂₅ L ₁₅₆	++++	Complete (24 h)
7	C ₇₄ M ₁₅₉ G ₁₈₈ I ₄₃ P ₁₂₅ L ₁₅₆	+	Low conversion (42 h)
8	C ₇₄ L ₁₅₉ G ₁₈₈ I ₄₃ P ₁₂₅ L ₁₅₆ + G190A	++	Incomplete (42 h)

[a] As judged from TLC.

When studying conversion of **7b** by the newly created variants of AMDase CLGIPL, it became apparent that all variants of positions 43 were more active towards this substrate than the original CLGIPL variant (Ile43) (Table 21, entry 4-6). These results emphasize once more that residue 43 resembles a hotspot position for accommodating larger substituents in the alkyl binding pocket. In contrast, extremely low levels of conversion were observed with variant L159M (Table 21, entry 7) and also the outstanding activity of the G190A variant towards **7a** was unmet with this substrate.

Yet, during chiral GC-FID analysis of the obtained reaction products **8b**, severe technical limitations were faced. Due to the overall chemical similarity of both the

vinyl and ethyl substituent, separation of enantiomers could not be achieved with the given instrumental setup (Figure 28).

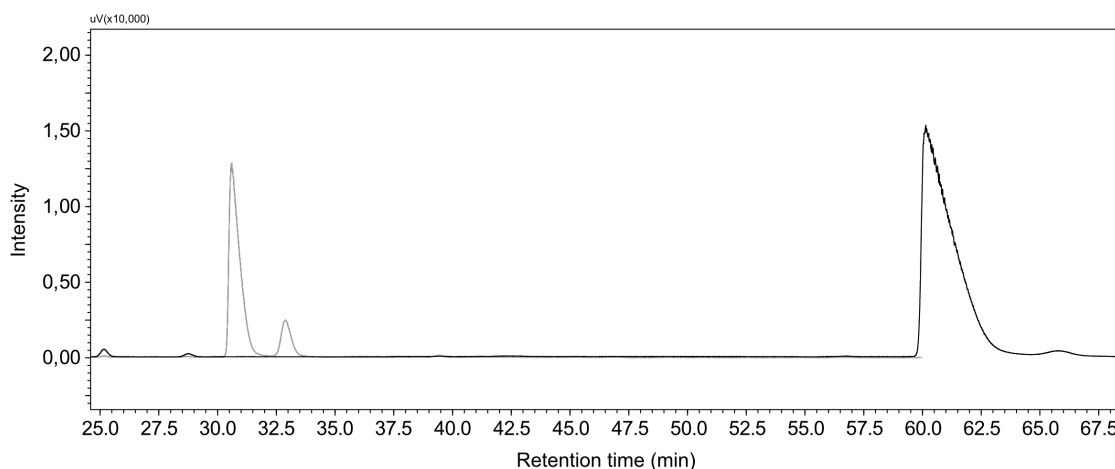


Figure 28. Chromatogram obtained from chiral GC-FID analysis of the reaction products obtained from decarboxylation of 2-ethyl-2-vinylmalonic acid (**7b**) by AMDase CLGIPL I43A (black) in comparison to (*S*)-**8a** ($t_R = 30.5$ min) and (*R*)-**8a** ($t_R = 33.0$ min) (gray).

As a consequence, neither the presence nor the identity of the target product enantiomers (*R/S*)-**8b** could be confirmed solely by GC-FID analysis. Yet, formation of 2-ethylbut-3-enoic acid (**8b**) is likely as judged from the typical retention behavior of malonic acids compared to monoacids during TLC analysis (**7b**: $R_f = 0.18$, **8b**: $R_f = 0.68$; Appendix G, Table 60). To eventually confirm the identity of the reaction products, further characterization *via* NMR spectroscopy is highly recommended.

4.3.5.3 Determination of temperature dependency of enantioselectivity

We were also interested, if different reaction temperatures would have an influence on the optical purity of the obtained reaction products. For this purpose, AMDase CLGIPL G190A and I43M variants were selected and 2-methyl-2-vinylmalonic acid (**7a**) was used as a substrate. It was particularly interesting that the lowest observed enantiomeric excess was obtained when using the most active variant, AMDase CLGIPL G190A (45 %*ee* (*S*); Table 20, entry 11).

Generally, formation of each product enantiomer is connected to a distinct reaction pathway and thereby determined by a certain activation energy barrier ($\Delta\Delta G^\ddagger$). Composition of the final product is thus determined by the energetic accessibility of the transition states leading to either of the two product

enantiomers. In the case of AMDase, imperfect optical purities most likely originate from the contribution of an additional binding mode, as formation of a certain product enantiomer is predetermined by the position of the proton donating cysteine residue relative to the substrate (see Section 4.2).

Phillips described the thermodynamic correlation of temperature and stereochemistry as follows:^{193,194} In general, the optical purity of a chiral compound can be defined by the enantiomeric ratio E according to Equation (1).¹⁹⁵ In a kinetically controlled reaction, the ratio of formed enantiomers equals the ratio of their rate constants (k_{cat}/K_M).

$$E = \frac{R}{S} = \left(\frac{k_{cat}}{K_M}\right)_R / \left(\frac{k_{cat}}{K_M}\right)_S \quad (1)$$

The resulting difference in activation free energy $\Delta\Delta G^\ddagger$ can be thus formulated as shown in Equation (2), whereby the temperature dependency becomes apparent.

$$\Delta\Delta G^\ddagger = -RT \ln E = \Delta\Delta H^\ddagger - T\Delta\Delta S^\ddagger \quad (2)$$

Following from this, there exists a temperature T_r – termed “racemic temperature” – where there is no enantioselectivity observed ($E = 1$), due to $\Delta\Delta G^\ddagger = 0$, that is formation of either enantiomer is energetically equal (Equation (3) and (4)).

$$\Delta\Delta H^\ddagger = T\Delta\Delta S^\ddagger \quad (3)$$

$$T_r = \Delta\Delta H^\ddagger / \Delta\Delta S^\ddagger \quad (4)$$

As a consequence, at temperatures below T_r , the enthalpic contribution ($\Delta\Delta H^\ddagger$) is dominating, thus resulting in increased enantioselectivity at decreasing reaction temperatures ($\Delta\Delta G^\ddagger > 0$ due to $\Delta\Delta H^\ddagger > T\Delta\Delta S^\ddagger$).

In enzymatic catalysis, differences in activation enthalpy ($\Delta\Delta H^\ddagger$) mainly originate from non-covalent Van-der-Waals interactions of the substrate with the enzyme, whereas activation entropy differences ($\Delta\Delta S^\ddagger$) are a result of restricted (rotational) motion of substrate and active site residues. A temperature dependency of enantioselectivity can be expected if the terms $\Delta\Delta H^\ddagger$ and $T\Delta\Delta S^\ddagger$ are rather balanced. In contrast, if bonding of the substrate in the active site is very pronounced (e.g. hydrogen bonding or ionic interactions), subtle changes of temperature cannot compensate the strong influence of $\Delta\Delta H^\ddagger$.¹⁹⁴ In other words, the racemic temperature T_r would be considerably high in such a case, thus

rendering the impact imposed by small temperature changes onto enantioselectivity rather imperceptible.

In the here studied examples of AMDase CLGIPL I43M and G190A, a diverging behavior was observed. While the product ee remained mostly constant within the studied temperature range (10°C-40°C) when using the I43M variant (72-79 %ee (S)), the G190A variant produced **8a** in a more variable composition (41-56 %ee (S)) (Figure 29). It has to be noted that, originally, a reaction temperature of 4°C was also included in this study, but the apparent enzyme activity at this temperature was too low to obtain significant results. The fact, that the feasible temperature range for influencing stereochemistry of an enzymatic reaction is clearly limited by the operability of the biocatalyst, was also mentioned by Phillips.¹⁹⁴

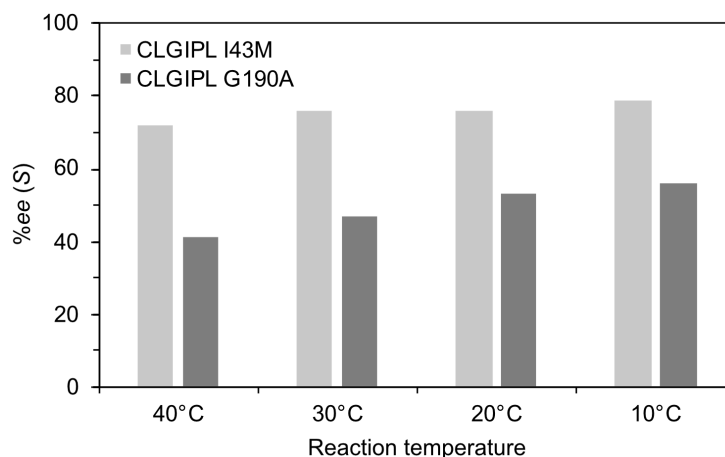


Figure 29. Determined enantiomeric excess of the product 2-methylbut-3-enoic acid (**8a**) obtained from AMDase CLGIPL I43M (light gray bars) or G190A-catalyzed (dark gray bars) decarboxylation of 2-methyl-2-vinylmalonic acid (**7a**) at different reaction temperatures. Raw data is shown in Appendix H (Table 61).

In the case of AMDase CLGIPL I43M, it seems that the transition state leading to the (S)-enantiomer is clearly more stabilized than the one leading to the opposite enantiomer, which is also reflected in the relatively high average ee. An enantiomeric excess of 76 % (S), as observed at 30°C, corresponds to a mixture of 88 % (S)- and 12 % (R)-enantiomer. According to previous discussions, we rather assume an exceeding contribution of $\Delta\Delta H^\ddagger$ in this example, thus leading to a low sensitivity towards different reaction temperatures. In comparison, the ee observed for G190A at 30°C is only 45 % (S) (corresponding to 73 % (S)- and 27 % (R)-enantiomer). This could be attributed to a generally more dynamic

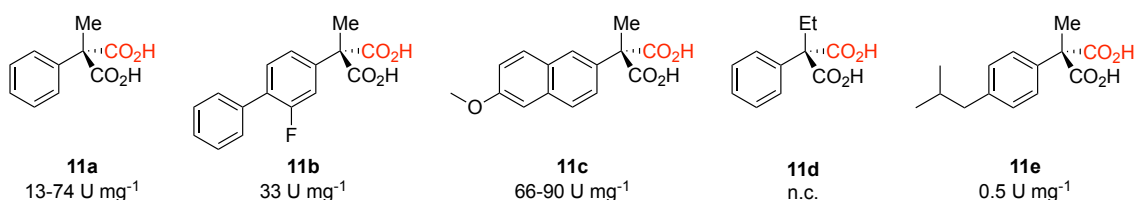
interchange between transition states resulting from an overall smaller energy difference $\Delta\Delta G^\ddagger$. This tendency to interchange is also reflected in the observed temperature dependency of stereoselectivity, where the overall activation free energy difference $\Delta\Delta G^\ddagger$ increases at decreasing temperature.

4.3.6 Computational studies

Computational studies were conducted in cooperation with Prof. Lynn Kamerlin (Department of Chemistry–BMC, Uppsala University), which aimed at a deeper understanding of AMDase activity, substrate scope and selectivity. In this regard, efforts to predict and design improved variants of AMDase are still hampered by the limited understanding of the underlying molecular principles. As a basis of this work, empirical valence bond (EVB) simulations¹⁹⁶ were performed and complemented by additional metadynamics simulations.¹⁹⁷

Specifically, we studied decarboxylation of the substrates shown in Scheme 39 by either AMDase wildtype or variants G74C/C188A, G74C/C188G and CLGIPL.

Scheme 39. Model substrates **11a–e** used in this study. The stated specific activity (U mg^{-1}) values for each substrate refer to AMDase wildtype.^{42,131,148,156} The pro-(*R*) and pro-(*S*) carboxylate are shown in black and red, respectively. n.c.: not converted.



It should be noted that the following section merely aims at summarizing and discussing the obtained results, rather than providing an in-depth description of the applied computational methods. These are summarized in detail in our recent publication.¹⁹⁸

4.3.6.1 Empirical valence bond (EVB) simulations

To start with, suitable structures of AMDase variants G74C/C188A, G74C/G188S and CLGIPL were created starting from the WT crystal structure (PDB: 3IP8) and substrates were docked into the active site. This procedure essentially afforded two main binding modes (Mode I and Mode II), whereby only Mode I resulted in viable activation free energies from EVB simulations (Figure 30).

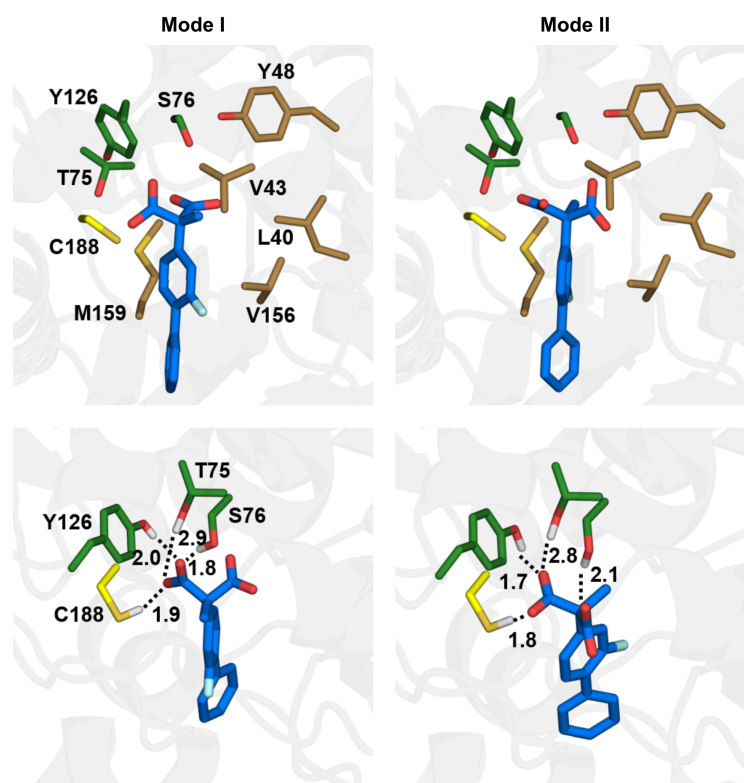


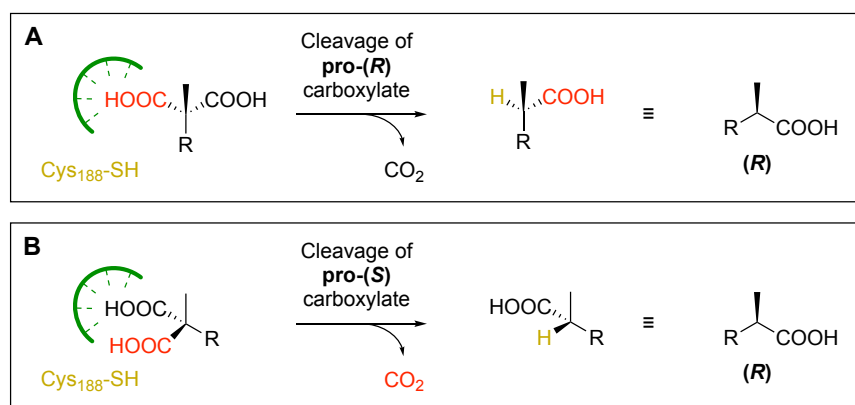
Figure 30. Representative binding modes (Mode I and Mode II) of substrate **11b** in the active site of AMDase WT obtained from docking. In this figure, substrate **11b** (blue), residues of the dioxanion hole (T75, S76, Y126; in green), residues of the hydrophobic pocket (L40, V43, Y48, V156, M159; in brown) and Cys188 (yellow) are highlighted. Top: Overview of the active site residues. Bottom: Detailed interactions between substrate (blue) and dioxanion hole (green). Figure taken and adapted from Biler *et al.*¹⁹⁸

In Mode I, the pro-(*S*) carboxylate is stabilized in the dioxanion hole, whereas the pro-(*R*) carboxylate is facing the hydrophobic pocket. Upon system equilibration, this carboxylate groups slightly rotates to end up fully exposed to the hydrophobic cavity. During this study, the cleavage of both the pro-(*R*) and pro-(*S*) carboxylate was simulated, despite earlier isotope labelling studies suggested exclusive cleavage of the pro-(*R*) carboxylate for (*R*)- and (*S*)-selective AMDase variants.^{132,133,141}

In agreement with experimental observations, the EVB models only showed turnover of compounds **11a-c** and **11e**, while activation free energies obtained for **11d** (for cleavage of either carboxylate) were very high. Also, the loss of activity upon introducing the G74C/C188X motif was reflected in the EVB simulations, whereby, in agreement with experiments, G74C/C188A was even less active than G74C/C188G.^{139,148}

As AMDase-catalyzed decarboxylation proceeds *via* a common planar intermediate, the formed product enantiomer is only determined by the orientation of this intermediate relative to the proton donating residue (res. 74 or 188), regardless of which carboxylate is cleaved beforehand. During EVB simulations, Michaelis complexes according to Pose A were typically observed during modelling cleavage of the pro-(*R*) carboxylate, whereas Pose B was mainly observed during simulating cleavage of the pro-(*S*) carboxylate (Scheme 40).

Scheme 40. Different possible binding poses (Pose A and Pose B) of substrates in the active site of AMDase wildtype and the resulting product configurations. The dioxanion hole is illustrated in green, the pro-(*R*) carboxylate in black and the pro-(*S*) carboxylate in red. The formed product is determined by the position of the Cys188 residue of the WT and other (*R*)-selective variants or the Cys74 (on opposite face of the intermediate compared to WT) in (*S*)-selective variants (e.g. CLGIPL or G74C/C188X). R: aryl or alkenyl. Figure taken and adapted from Biler *et al.*¹⁹⁸



In case of AMDase wildtype, preferential cleavage of the pro-(*R*) carboxylate was observed for all substrates studied, even for those which were described to be not (**11d**) or only poorly (**11e**) converted by AMDase.^{131,156} According to the poses presented in Scheme 40, this would lead to formation of the corresponding (*R*)-enantiomers. Also, when studying the (*S*)-selective variants G74C/C188X, preferential cleavage of the pro-(*R*) group (*via* Pose A) was shown, which would yet result in formation of (*S*)-enantiomers due to protonation by Cys74 from the opposite site of the intermediate. Both observations are in perfect agreement to isotope labelling studies, where cleavage of the pro-(*R*) carboxylate was confirmed for (*R*)- and (*S*)-selective variants.^{132,133,141} Yet quite surprisingly, for AMDase CLGIPL, EVB simulations showed preferential elimination of the pro-(*S*) carboxylate group, which would still result in formation of (*S*)-products due to reacting *via* Pose B. Despite this quite unexpected reaction behavior detected for

CLGIPL, the observed (*S*)-selectivity, as well as the calculated activation free energies were in agreement with experiments.^{42,44,145,148,156}

4.3.6.2 *Molecular-level determinants of the observed activation free energies*

At first, factors influencing the low activity of AMDase towards compound **11e** and the inability to convert **11d** were analyzed. Generally, presence of alkyl substituents on both substrates could complicate accommodation of these compounds in the hydrophobic pocket, thus rather leading to non-productive binding modes. During the simulations performed in this work, **11d-e** were relatively motile within the active site as compared to well-accepted substrates. In addition, also the extended alkyl substituents themselves were fluctuating substantially. As a consequence of less precise positioning of these substrates, more solvent could penetrate the active site, thus potentially interfering with the destabilizing interactions within the hydrophobic pocket.

When analyzing active site volumes upon substrate binding, the observed volume was directly correlating with substrate size, thus indicating a certain flexibility of the active site to provide a “tight fit” depending on each substrate.

As mentioned before, accessibility of the active site by solvent turned out crucial, as typically less than one water molecule was observed in the close vicinity of the carboxylate to be cleaved for well-accepted substrates. In case of substrates **11d-e** and AMDase variants G74C/C188X, the observed low reactivity could be clearly connected to an increased number of water molecules present close to the reactive center. The vital ability to exclude water from the active site is a principle often observed in enzyme catalysis.^{199–202}

Also, electrostatic contributions of particular amino acids to the calculated overall activation free energies were considered. Yet, while some residues had explicit stabilizing (*e.g.* T75 and Y126 from the dioxyanion hole) and others destabilizing properties (*e.g.* hydrophobic pocket residues except M159), the actual interaction network within the AMDase active site is indeed quite complex and a clear discrimination between stabilizing and destabilizing sets of amino acids was unfeasible.

4.3.6.3 Contribution of ground-state effects

To study participation of ground-state destabilization during AMDase-catalyzed decarboxylation, molecular dynamics simulations were performed and analyzed by consulting the grid inhomogeneous solvation theory (GIST).^{203–205} This method allowed for calculation of local solvation free energies which were used as a measure of local hydrophilicity and hydrophobicity within the active site. This grid map of hydrophobic and hydrophilic areas was projected onto the active site of AMDase WT and CLGIPL, each containing substrate **11b** in both reactive Pose A and B (Figure 31). It became quite obvious that the majority of the AMDase active site cavity was hydrophobic (in magenta), which is ideally suited to accommodate the extended aromatic substrate moieties.

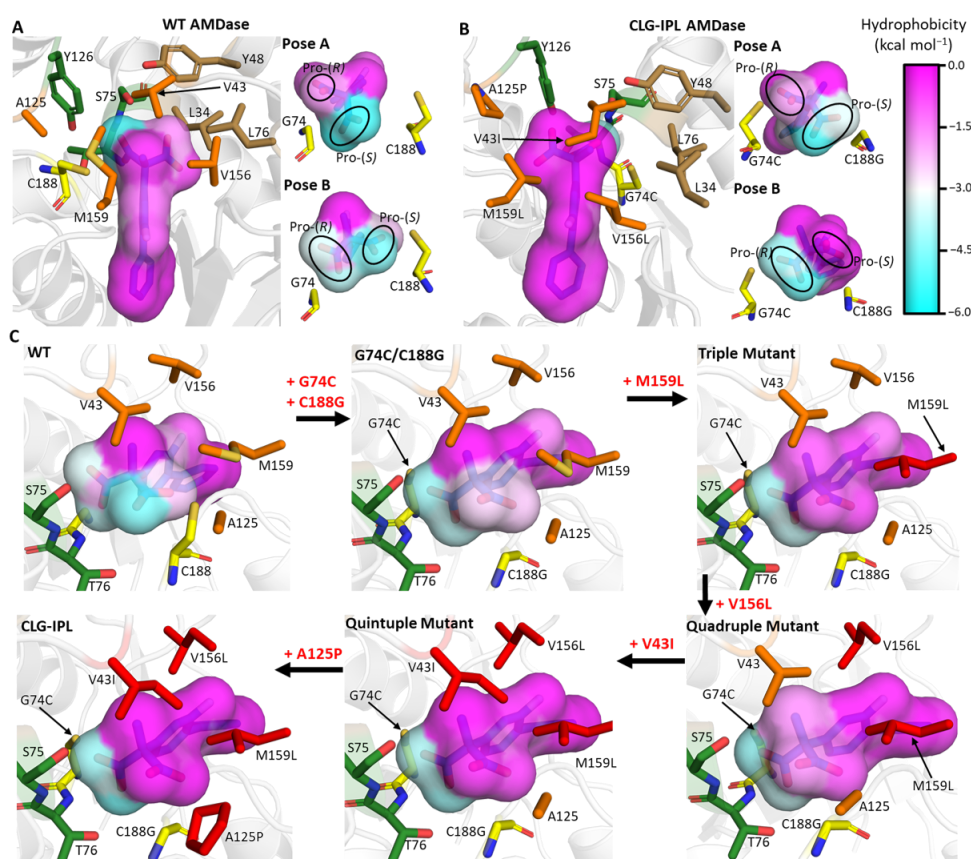


Figure 31. Local active site hydrophobicity projected onto substrate **11b** in the AMDase active site. Hydrophobic areas are represented in magenta and hydrophilic areas in cyan. (A) Binding of **11b** in the WT AMDase active site (Pose A) and a close-up of both carboxylates in either Pose A or B. (B) Binding of **11b** in the AMDase CLGIPL active site (Pose B) and a close-up of both carboxylates in either Pose A or B. (C) Active site hydrophobicity over the course of AMDase CLGIPL engineering (Pose B). Mutations introduced in each step are shown in red. In this figure, residues of the dioxanion hole (green), the original hydrophobic pocket (brown) and target residues during engineering AMDase CLGIPL (orange) are highlighted. Figure taken from Biler *et al.*¹⁹⁸

In case of AMDase wildtype, which is supposed to react through Pose A (see 4.3.6.1), the pro-(*S*) carboxylate is clearly stabilized within the highly hydrophilic dioxyanion hole (in cyan), whereas the pro-(*R*) group is positioned in a hydrophobic area. In contrast, in the less preferred Pose B, the pro-(*R*) group gets rather stabilized than destabilized, thus impeding successful decarboxylation *via* this pose (Figure 31A). As suggested by the previously performed EVB simulations, CLGIPL curiously favored decarboxylation of the pro-(*S*) group *via* Pose B. By the GIST analysis performed here, this was indeed confirmed, as ground-state destabilization of the pro-(*S*) carboxylate in Pose B was clearly evident. It seemed that by creating the six-fold variant CLGIPL, an additional hydrophobic pocket emerged, thus allowing this unusual cleavage of the pro-(*S*) group (Figure 31B). It should be noted that reaction *via* Pose A appears still plausible in this case, a fact already observed during the performed EVB simulations.

We were also curious if emergence of this additional hydrophobic pocket could be followed over the engineering pathway of the CLGIPL variant, which resulted from several rounds of simultaneous site-saturation mutagenesis (SSM).¹⁵⁶ In the first step, G74C/C188G responsible for (*S*)-selectivity was introduced. These mutations removed steric clashes exerted by C188 onto the pro-(*S*) group when bound in Pose B, thus allowing a more favorable positioning of the substrate residues. When introducing the substitution M159L, the hydrophobicity close to the pro-(*S*) carboxylate was increased significantly (Figure 31C). This is in conjunction with the substantial activity increase observed in experiments after introducing this mutation.¹⁵⁶ All subsequently introduced substitutions had a rather weak impact on hydrophobicity as these positions were generally more distant to the reactive center. Yet, the overall hydrophobicity was still gradually increasing over all mutagenesis steps (Figure 31C). This evolutionary pathway clearly demonstrated that ground-state destabilization by modulation of active site hydrophobicity can directly influence both AMDase activity and selectivity for cleaving a particular carboxylate group. While development of an additional hydrophobic pocket in the CLGIPL variant was not by design, the targeted engineering of pockets still represents a viable strategy for future engineering of AMDase.

To complement the present work, differences in population of ground-state binding poses were studied by performing well-tempered metadynamics simulations (WT-MetaD).¹⁹⁷ These differences can originate from differing stabilization of Michaelis complexes in both observed Pose A and B. In all studied cases, two energy minima were detected (corresponding to Pose A and B), which were separated by a transition state barrier. This barrier represents the unfavorable interconversion between both poses which is accompanied by breaking and forming of interactions between the dioxyanion hole and either carboxylate group. Once again, for WT AMDase and G74C/C188X variants, the energetically preferred state was Pose A, while Pose B was mostly preferred for CLGIPL. Binding of substrate **11c** in the CLGIPL active site presented the only exception, but the difference in free energy between both poses was only ~ 1.5 kcal mol⁻¹, thus allowing easy interconversion. The performed simulations thus underline, that the determined free energy minima are congruent with the observed ideal reactive binding poses.

It can be thus concluded that cleavage of a certain carboxylate group is determined at several stages along the reaction coordinate: (1) through binding of the substrate in a preferred binding pose, followed by (2) selective destabilization of the carboxylate to be cleaved through placing in a hydrophobic pocket and eventually through (3) differing stabilization of the transition states leading to cleavage of either the pro-(*R*) or pro-(*S*) group.

4.4 Summary

In this work, arylmalonate decarboxylase (AMDase) from *B. bronchiseptica*, more specifically the CLGIPL (V43I/G74C/A125P/ V156L/M159L/C188G) variant, was the main target of investigation. In particular, the peculiar behavior of the CLGIPL variant during decarboxylation of 2-methyl-2-vinylmalonic acid (**7a**), which led to the desired product 2-methylbut-3-enoic acid (**8a**) but only in 66 %ee (S), attracted our attention. These findings led to our working hypothesis according to which a putative alternative binding mode adversely affects the otherwise exceptional enantioselectivity of this enzyme, yet curiously only in this specific substrate-enzyme constellation, as related substrates with additional substituents on the alkenyl moiety were converted with unchanged high selectivities.⁴⁶

AMDase CLGIPL was thus subjected to rational amino acid exchanges mainly within the hydrophobic and alkyl binding pocket and their impact on enantioselectivity was investigated. While our intention was to prevent binding of the substrate in the suggested alternative binding mode by rationally re-engineering the active site cavity, only marginal changes in enantioselectivity were observed upon exchanging single amino acids (Figure 32).

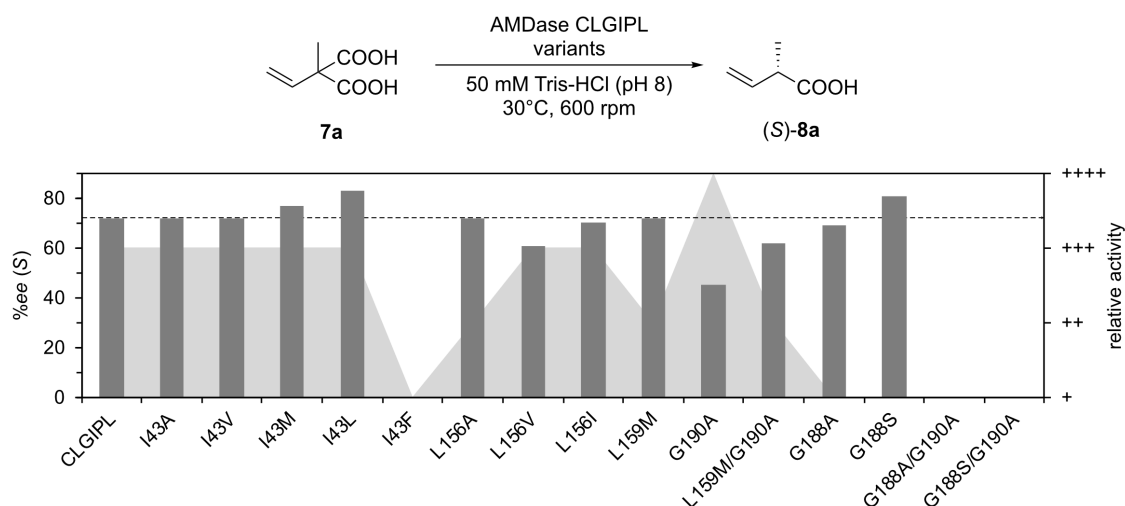


Figure 32. Summarized enantioselectivity (dark gray bars) and relative activity (light gray area) of AMDase CLGIPL and variants thereof towards the decarboxylation of 2-methyl-2-vinylmalonic acid (**7a**) as represented in the reaction scheme. In case there is no entry for enantioselectivity, the respective ee was not determined. Relative activity is divided in higher (++++), similar (+++), lower (++) and negligible (+) activity compared to CLGIPL.

While some slight improvements of enantioselectivity could be observed, for instance when gradually increasing the size of the substituent at position 43 (from

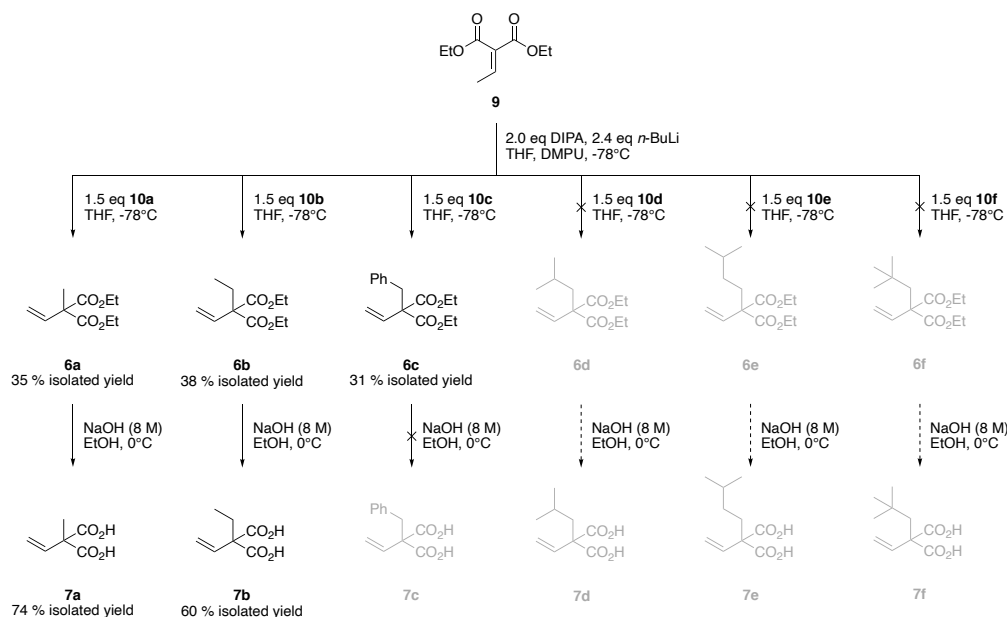
I43A to I43L), installation of an even bulkier phenylalanine (I43F) led to a completely inactive variant. Similarly, the slightly enhanced *ee* observed for the G188S variant was offset by the extraordinarily low activity of this variant. These results underline once more the significance of synergistic effects within the hydrophobic pocket to drive catalysis of AMDase.⁴³

While some of the created variants were also able to accept 2-ethyl-2-vinylmalonic acid (**7b**) as a substrate, further investigations were hampered by technical limitations regarding separation of the potentially formed enantiomers during chiral GC-FID analysis.

We also turned our attention to engineering the substrate scaffold in a way, where the small methyl group usually attached to the substrate should be substituted by bulkier alkyl groups. We proposed that such a structural composition could enforce binding of the substrate in the alternative binding mode, as then only the vinyl substituent would be small enough to fit in the spatially restricted alkyl binding pocket. As evidence of our conceived compounds diethyl 2-alkyl-2-vinylmalonates **6b-f** and 2-alkyl-2-vinylmalonic acid **7b-f** in literature was scarce, we intended to adapt the already established two-step synthetic route towards 2-methyl-2-vinylmalonic acid (**7a**) to our purpose (Scheme 41). In the first step, alkylation of diethyl 2-ethylidenemalonate (**9**) was attempted *via* LDA-mediated formation of a reactive lithium enolate species of **9**, followed by nucleophilic attack of the respective alkyl iodides (**10a-b**) or bromides (**10c-f**). In the second step, the formed 2-alkyl-2-vinylmalonates **6a-f** should be saponified to yield the corresponding 2-alkyl-2-vinylmalonic acids **7b-f** (Scheme 41).

Yet, in terms of the alkylation reaction, only compounds **6a-c** could be successfully isolated in 31-38 % yield. In case of compounds **6d-f**, formation of the target products was presumably prevented by the low reactivity of the employed alkyl bromides **10d-f** in a nucleophilic substitution reaction (S_N2). Issues of a different kind were faced during ester cleavage of diethyl 2-benzyl-2-vinylmalonate (**6c**). Due to instability of the released malonic acid **7c**, the corresponding decarboxylation product 2-benzylbut-2-enoic acid (**8c'**) was the main compound recovered from this reaction. In view of the issues encountered on the way of preparing 2-alkyl-2-vinylmalonic acids **7c-f**, development of an entirely different synthesis strategy is unavoidable.

Scheme 41. Summarized synthetic route towards diethyl 2-alkyl-2-vinylmalonate **6a-f** and 2-alkyl-2-vinylmalonic acid **7b-f**.



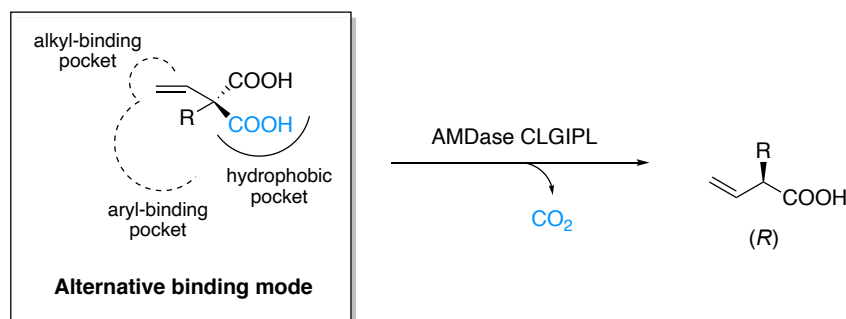
Complementary computational studies helped to further uncover molecular level determinants for the observed activity and selectivity of AMDase and its variants towards different aromatic model compounds. The performed EVB calculations could reproduce experimentally observed activation free energies, enantioselectivity and substrate scope of AMDase and variants. The substrate scope itself seemed to be mainly determined by adequate substrate binding while simultaneously assuring the exclusion of water from the active site, the latter of which can destroy the vulnerable hydrophobic network needed for successful conversion. Both factors are closely connected as for non-converted substrates, an increased motion or flexibility of substrate moieties was often accompanied by a greater solvent exposure of the active site. Analysis of electrostatic interactions has highlighted the complex interplay within the AMDase active site. While these effects seemed to play a role in stabilizing individual transition states, ground-state destabilization was the main determinant of selectivity between different possible transition states. Pursuing GIST analysis evidenced this ground-state destabilization by entrapment of the cleaving carboxylate group in a hydrophobic cavity. Engineering of AMDase selectivity by manipulation of hydrophobic pockets thus represents a useful approach as evidenced by the yet unintentionally emerged additional hydrophobic pocket during generation of the CLGIPL variant.

4.5 Outlook

In view of the promising insights obtained from computational studies performed on AMDase with different aromatic model substrates,¹⁹⁸ we are also striving to deepen our knowledge regarding the smaller alkenyl substrates.

In this respect, we are currently contriving further computational studies in cooperation with Prof. Lynn Kamerlin (Department of Chemistry–BMC, Uppsala University) in order to investigate contribution of a putative alternative binding mode during decarboxylation of 2-alkyl-2-vinylmalonic acids **7**, as for instance suspected in the case of AMDase CLGIPL during conversion of 2-methyl-2-vinylmalonic acid (**7a**). As outlined already earlier in this work, placement of the substrate in such an orientation would lead to formation of (*R*)-enantiomers when using the conventionally (*S*)-selective variant CLGIPL (Scheme 42).

Scheme 42. Proposed alternative binding mode of AMDase CLGIPL during decarboxylation of 2-alkyl-2-vinylmalonic acids. In this binding mode, the vinyl group is supposed to be located in the small alkyl-binding pocket, which is usually only occupied by a methyl group. This relative orientation of the substrate and consequently also of the formed enediolate intermediate would result in formation of the respective (*R*)-enantiomer after protonation by Cys74.

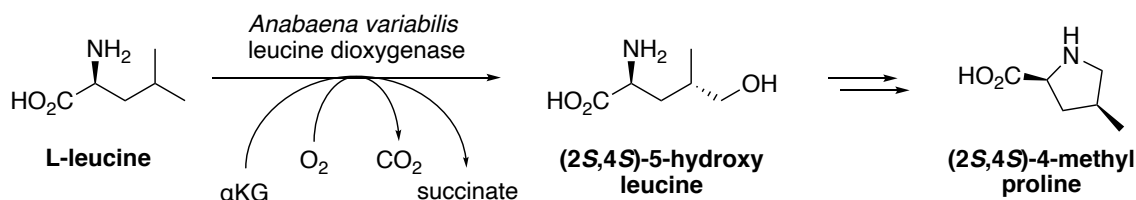


As previously demonstrated for AMDase, EVB simulations together with thorough analysis of involved contributing factors such as active site volume and hydrophobicity, electrostatic interactions or population of ground-state binding poses can help to elucidate experimentally observed ambiguities.

5 L-Leucine dioxygenase from *Anabaena variabilis* (AvLDO)

A non-heme Fe(II) and α -ketoglutarate dependent L-leucine dioxygenase was recently discovered in the cyanobacterium *Anabaena variabilis* (AvLDO). The enzyme can catalyze δ -specific hydroxylation and sulfoxidation of L-amino acids in a regio- and stereoselective fashion. The gene encoding AvLDO is located in a gene cluster involved in (2S,4S)-4-methylproline biosynthesis, which itself is a building block for bioactive peptides, thus indicating its native functional role in cyanobacteria (Scheme 43). Yet, the substrate scope of AvLDO is strictly limited to the aliphatic amino acids L-leucine and L-norleucine for hydroxylation and L-methionine and ethionine for sulfoxidation, thus currently hampering broader applicability of this enzyme.

Scheme 43. Fe(II)/ α KG dependent L-Leucine dioxygenase from *Anabaena variabilis* (AvLDO) catalyzes the δ -specific hydroxylation of L-leucine to yield (2S,4S)-5-hydroxyleucine. AvLDO is encoded in a gene cluster involved in biosynthesis of (2S,4S)-4-methylproline.



5.1 Theoretical background

5.1.1 Classification of oxygenases

For a long time, it has been thought that biological oxidation processes proceed exclusively *via* the transfer of electrons or hydrogen atoms between substrate and acceptor molecules. These enzymes were termed “dehydrogenases” or “oxidases” in case oxygen served as an electron acceptor. According to these early proposals, oxygen incorporated into substrate molecules could only originate from water, and direct incorporation of molecular oxygen was considered irrelevant.²⁰⁶

However, in 1955, isotope labelling studies using $^{18}\text{O}_2$ revealed two enzymes, which were able to incorporate oxygen stemming exclusively from molecular oxygen.^{207,208} This was in contrast to early-known oxidation reactions, where oxygen was reduced to H_2O or H_2O_2 . This newly discovered class of enzymes was termed “oxygenases”, which were further subclassified into mono- and dioxygenases (Figure 33).²⁰⁹

Monoxygenases incorporate only one oxygen atom into the substrate, whereas the other one is reduced to water with the aid of reducing agents (mostly NAD(P)H). Typically, they also rely on cofactors such as flavin, non-heme iron, heme or copper.²⁰⁹

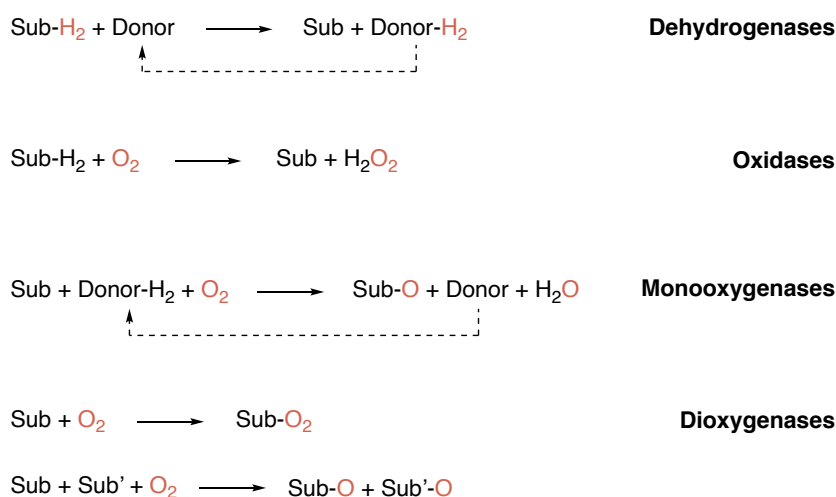


Figure 33. Enzymatic oxidation reactions as catalyzed by dehydrogenases, oxidases, monoxygenases and dioxygenases. The dashed arrow indicates the requirement for a cofactor recycling system. Sub: substrate. Figure adapted from Faber.²¹⁰

In contrast, dioxygenases incorporate both oxygen atoms either into one substrate or into two different substrate molecules. In the latter case, α -ketoglutarate always constitutes one of the substrates. The majority of dioxygenases uses either heme or non-heme iron as a cofactor, whereby the latter is clearly predominant.²⁰⁹

Specifically, non-heme Fe(II)/ α -ketoglutarate dependent dioxygenases catalyze two-electron substrate oxidation, coupled to oxidative decarboxylation of the cosubstrate α -ketoglutarate producing succinate and CO_2 . While one oxygen atom is typically incorporated into the substrate (e.g. during substrate hydroxylation), the other one ends up in the coproduct succinate.²¹¹

5.1.1.1 Biological role of Fe(II)/ α KG dependent dioxygenases

Besides their most prominent hydroxylation activity, Fe(II)/ α KG dependent dioxygenases catalyze a wide array of different other oxidative transformations, including halogenation, demethylation, ring closure or expansion, desaturation and epoxidation or epimerization reactions (Figure 34).^{212,213}

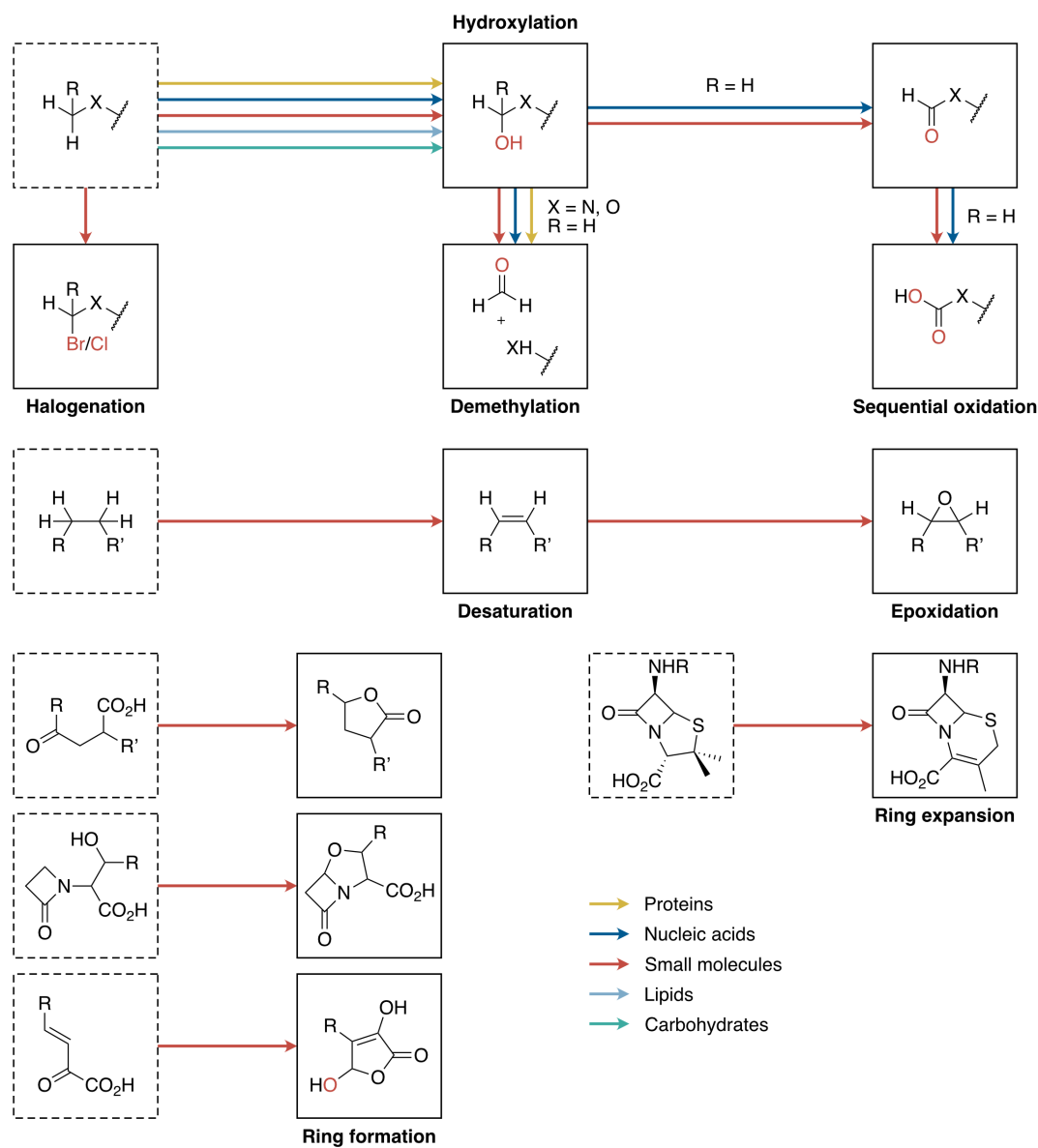


Figure 34. Overview on reactions catalyzed by Fe(II)/ α KG dependent oxygenases. Besides hydroxylation, also halogenation, demethylation, and sequential oxidation can occur. Further, this enzyme class can catalyze desaturation and epoxidation or ring formation and expansion. The color of arrows represents the respective target substrates being either proteins (yellow), nucleic acids (dark blue), small molecules (red), lipids (light blue) or carbohydrates (cyan). Figure adapted from Islam *et al.*²¹²

Due to their catalytic versatility, these enzymes are involved in numerous biological processes and are ubiquitously found in animals, plants and microbes.^{212,214}

In primary metabolism, Fe(II)/ α KG dependent dioxygenases are mainly involved in regulatory or sensing processes, including regulation of transcription, nucleic acid repair, hypoxia sensing, lipid metabolism and collagen biosynthesis.^{212,215–217} For instance, *E. coli* AlkB mediates repair of DNA or RNA damage induced by methylation *via* hydroxylation of the *N*-methylated nucleobases followed by deformylation.²¹⁸ Furthermore, the diverse regulatory functions covered by Fe(II)/ α KG dependent dioxygenases in humans also make them an attractive targets for therapeutic intervention.²¹⁹ In plant and microbe secondary metabolism, Fe(II)/ α KG dependent dioxygenases mainly participate in small molecule biosynthesis or degradation, including reactivities towards carbohydrates, peptides or terpenoids and alkaloids.²¹² Intriguingly, these enzymes do not act exclusively on (monomeric) building blocks (e.g. hydroxylation of amino acids²²⁰) but can also diversify compounds during oligomerization (e.g. halogenation of tethered peptides²²¹) or even after complete assembly (e.g. peptide modifications²¹²).

5.1.2 Structure and mechanism of Fe(II)/ α KG dependent dioxygenases

The core structure of Fe(II)/ α KG dependent dioxygenases is defined by variations of a double-stranded β -helix fold (DSBH), also known as jellyroll fold. This core structure provides binding sites for Fe(II) and α -ketoglutarate and can be surrounded by several additional structural elements that mark different sub-families. Particularly, less structurally defined and conserved regions are thought to be involved in substrate binding, that is, besides β I, β II and β VIII strands, loop inserts between β II- β III or β IV- β V or *N*- and *C*-terminal regions (Figure 35).^{211,222} Sequentially consecutive β -strands of the DSBH are located in different sheets, except β IV and β V, which are both part of the minor β -sheet.²²² Additional parallel or anti-parallel β -strands are mostly present, extending either of the two core β -sheets (Figure 35).²²² The structural asymmetry of the distorted barrel- or sandwich-like structure formed by the major and minor β -sheet creates a suitable active site cavity, where β I and β VIII strands form the entrance.²¹¹ The Fe(II)

coordination site is characterized by the conserved HXD/E...H motif, which is localized at the wider end of the DSBH core within the minor β -sheet (β II and β VII).^{211,222}

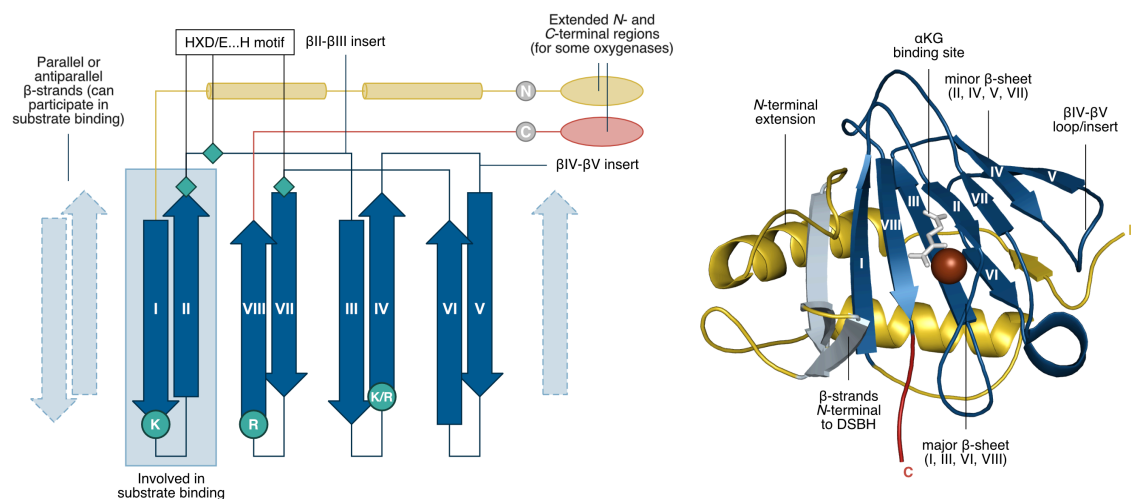


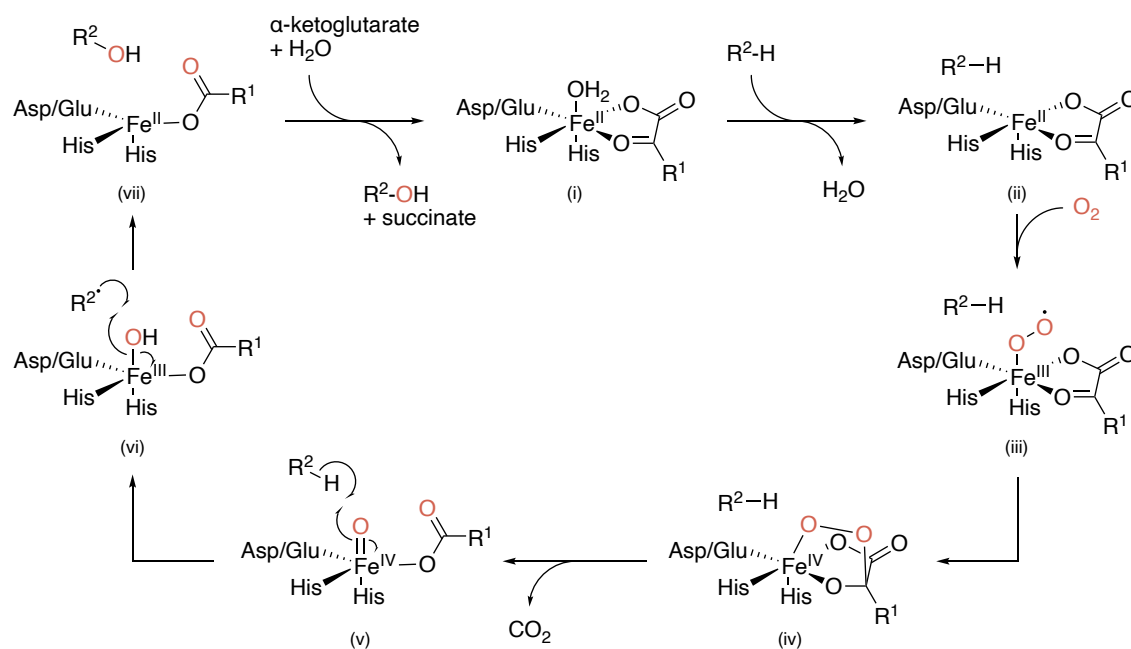
Figure 35. General topology observed for Fe(II)/ α KG dependent dioxygenases. The conserved DSBH core (dark blue) consists of eight antiparallel β -strands (I-VIII) arranged in two four-stranded β -sheets (major and minor sheet). The distorted barrel-like structure contains conserved Fe(II) binding sites (cyan diamonds) and α KG binding residues (cyan circles), which are subfamily characteristic. The core structure is surrounded by the N-terminal extension (yellow), typically featuring α -helices and C-terminal extension (red). Additional β -strands, as well as inserts between adjacent β -strands can be present (left). Regions according to the general topology projected onto the structure of AlkB (PDB: 3BIE) (right). Structure visualized with PyMOL.⁸⁸ Figure adapted from Islam *et al.*²¹² and Aik *et al.*²¹¹

It is generally believed that the conformational flexibility of this enzyme class during catalysis is widely underestimated.²¹¹ Also, the sequential order of (co)substrate-binding seems to play an important role. For instance, NMR structural analysis of AlkB revealed stabilization of the overall structure upon binding of both Fe(II) and α -ketoglutarate, thereby enabling subsequent binding of the prime substrate.²²³ There is also evidence that an induced fit is triggered upon substrate binding in some cases.^{211,222} In any case, accurate positioning of the prime substrate is essential to avoid substrate-uncoupled turnover of α -ketoglutarate or oxidation of miscellaneous atoms (see 5.1.2.1).^{211,212}

Substrate hydroxylation is clearly the best understood and most widespread reactivity of Fe(II)/ α KG dependent dioxygenases. Hydroxylation typically occurs at a non-activated carbon center. The 2-His-1-carboxylate facial triad typically coordinates the Fe(II) center and remaining coordination sites are occupied by three water molecules. Two of these water molecules are displaced upon binding

of the cosubstrate α -ketoglutarate, which coordinates Fe(II) in a bidentate manner, where the keto group is positioned *trans* to the Asp or Glu residue and the carboxylate *trans* to one of the His residues (Scheme 44-i).

Scheme 44. Catalytic cycle proposed for Fe(II)/ α KG dependent dioxygenases during substrate hydroxylation. The iron center is coordinated by the 2-His-1-carboxylate facial triad and by α KG in a bidentate fashion (i). Water displacement triggered by substrate binding creates a vacant coordination site for O_2 (ii), which leads to the formation of a Fe(III)-superoxo radical (iii). After attack of α KG (C2) (iv) followed by oxidative decarboxylation, the central Fe(IV)-oxo or ferryl intermediate is formed (v). This species abstracts a hydrogen atom from the substrate R^2 -H and the generated substrate radical $R^{2\cdot}$ attacks the Fe(III)-OH intermediate (vi). The hydroxylated product R^2 -OH dissociates, thereby concluding the catalytic cycle. $R^1 = (CH_2)_2COOH$.



The positioning of the primary substrate R^2 -H close to the reactive center triggers displacement of the last remaining metal-bound water molecule, thereby creating a vacant coordination site for molecular oxygen (Scheme 44-ii). The binding of O_2 leads to the formation of a Fe(III)-superoxo radical species (Scheme 44-iii), which is able to attack α -ketoglutarate at C2, thereby resulting in a bicyclic peroxohemiketal structure (Scheme 44-iv). Upon the collapse of this intermediate accompanied by oxidative decarboxylation of α -ketoglutarate, the central ferryl intermediate (Fe(IV)-oxo species) coordinated by succinate is formed (Scheme 44-v). This highly reactive species can abstract a hydrogen atom of the substrate molecule, thereby generating a Fe(III)-OH complex and a substrate radical $R^{2\cdot}$ *via* homolytic bond fission (Scheme 44-vi). The oxygen rebound mechanism is concluded through the attack of Fe(III)-OH by the substrate radical, thus leading

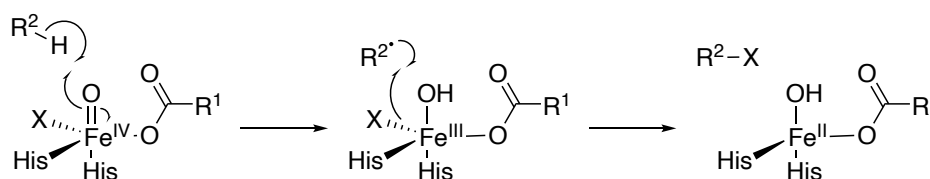
to the final hydroxylated product R^2-OH , which further dissociates to finalize the catalytic cycle (Scheme 44-vii).²¹³ Alternative pathways were also discussed regarding the later steps of the catalytic cycle, including substrate hydroxylation and dissociation of products.²²⁴

5.1.2.1 Halogenation reactivity of Fe(II)/ α KG dependent dioxygenases

While in the majority of Fe(II)/ α KG dependent dioxygenases the central Fe(II) is coordinated by a conserved HXD/E...H triad, α KG dependent halogenases present an interesting exception to this binding motif.²¹² The characteristic carboxylate residue is exchanged by a non-coordinating residue (e.g. alanine) to create space for binding of a halide ion.²²¹

Despite the slightly different coordination pattern, the initial catalytic steps for activation of molecular oxygen proceed *via* the formation of a high-energy Fe(IV)-oxo intermediate, as is also observed for the related hydroxylases (Scheme 44).²¹³ Similarly, the haloferryl intermediate abstracts a hydrogen atom of the substrate, thereby creating a substrate radical that reacts with the halide bound to the metal center (Scheme 45).

Scheme 45. Catalytic mechanism proposed for Fe(II)/ α KG dependent halogenases. Catalytic steps towards formation of the central Fe(IV)-oxo (here: haloferryl) intermediate are identical to the corresponding hydroxylases, except that a halide X (Cl or Br) replaces the carboxylate (Asp or Glu) of the facial triad. Similarly, the haloferryl intermediate (left) abstracts a hydrogen atom from the substrate whereupon a substrate radical and a *cis*-halohydroxo-ferric species (center) are formed. The substrate radical then reacts with the halide to finally result in formation of the halogenated product R^2-Cl (right). Scheme adapted from Martinez *et al.*²¹³



5.1.2.2 Substrate uncoupled turnover of α -ketoglutarate

It is generally thought that the ability of α -ketoglutarate dependent oxygenases to alternatively use only two or three metal coordinating residues offers superior catalytic flexibility compared to P_{450} enzymes, where the majority of coordination sites is already occupied by the heme porphyrin.²¹² Yet, this shielded metal center

makes P₄₅₀ enzymes generally less susceptible to oxidative damage compared to Fe(II)/ α KG dependent dioxygenases.²²⁵

Besides the possibility of the ferrous cofactor to undergo oxidation to Fe(III) mediated by O₂ in the absence of any (co)substrate, the ability of Fe(II)/ α KG dependent dioxygenases to catalyze the decarboxylation of α -ketoglutarate uncoupled from prime substrate conversion has also been long known.^{226,227} This alternative reaction pathway depends on presence of the same cosubstrates, that is α -KG and molecular oxygen, and leads to the formation of succinate and CO₂, probably also forming a reactive Fe(IV)-oxo intermediate.^{219,225} In the absence or in case of improper positioning of the prime substrate, the enzyme remains trapped in this high iron oxidation state, thereby preventing any further turnover.^{214,222} While undesired oxidation of the iron center can be reverted by the addition of reducing agents, the most prominent example being ascorbate,²¹⁴ the highly reactive Fe(IV)-oxo intermediate can conversely also react with active site residues, thereby leading to irreversible self-hydroxylation^{228–230} or even fragmentation²³¹ of the protein scaffold. Yet, the exact mechanisms involved in these auto-oxidation processes remain somewhat elusive, as isotope labelling studies using ¹⁸O₂ revealed that the oxygen incorporated in the hydroxylated sidechain of TauD was of different origin, thus indicating an involvement of more complex reaction cascades.²²⁹

5.1.3 Fe(II)/ α KG dependent amino acid hydroxylases

Among the Fe(II)/ α KG dependent hydroxylases acting on free amino acids, L-proline hydroxylases are probably best studied, due to the importance of chiral hydroxyproline building blocks for pharmaceuticals and antibiotics.²³² Three different types, which are L-proline *cis*-4-hydroxylase (*cis*-P4H), *trans*-4-hydroxylase (*trans*-P4H) and *cis*-3-hydroxylase (*cis*-P3H), have been identified in microbes so far.^{233–235} Beyond that, some microbial dioxygenases were identified to catalyze hydroxylation of L-asparagine (L-asparagine 3-hydroxylase AsnO) or L-arginine (L-arginine 3-hydroxylase), both of which are involved in the biosynthesis of peptide antibiotics in *Streptomyces* sp.^{236,237} Only in 2009, L-isoleucine 4-hydroxylase (IDO) from *Bacillus thuringiensis* was identified,²³⁸ thereby commencing discovery of further dioxygenases acting on aliphatic amino acids, early examples including L-leucine 5-hydroxylase (LdoA) from *Nostoc*

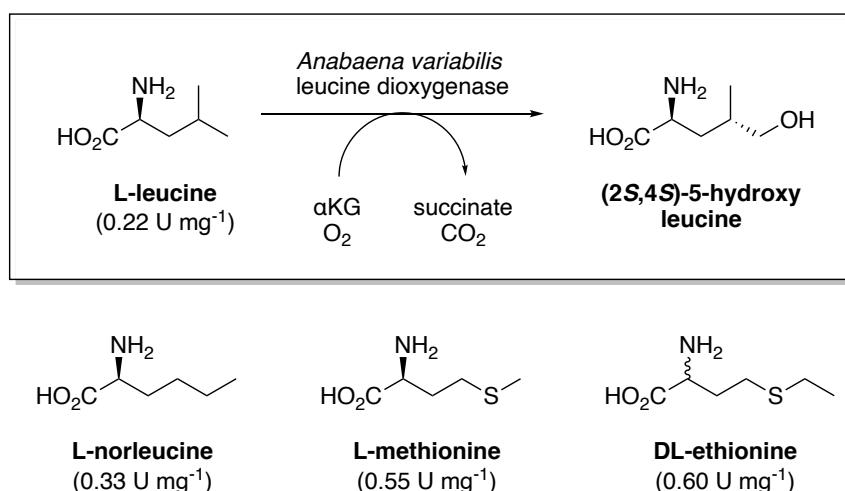
punctiforme and *N*-succinyl L-leucine 3-hydroxylase (SadA) from *Burkholderia ambifaria*.^{239,240} Resulting hydroxylated aliphatic amino acids are commonly found as secondary metabolites or integrated into proteins.²²⁰

5.1.3.1 L-Leucine dioxygenase from *Anabaena variabilis* (AvLDO)

An open reading frame within the genome of *Anabaena variabilis* encoding a putative L-leucine dioxygenase (AvLDO) was identified as a result of a protein BLAST search using L-leucine-5-hydroxylase (LdoA), previously identified in the cyanobacterium *Nostoc punctiforme*, as query sequence (82 % identity).^{240,241} Similarly, as for LdoA, AvLDO is encoded as part of a gene cluster involved in (2*S*,4*S*)-4-methylproline biosynthesis, a building block of bioactive peptides commonly found in cyanobacteria.^{240–242}

Characterization of AvLDO revealed that amongst the tested proteinogenic and non-proteinogenic amino acids, only L-leucine and L-norleucine were accepted as substrates for hydroxylation or alternatively L-methionine and DL-ethionine for sulfoxidation (Scheme 46). Regioselective δ -hydroxylation was detected for AvLDO, which is coinciding with LdoA and in accordance with the putative involvement in methylproline biosynthesis (*via* (2*S*,4*S*)-5-hydroxyisoleucine as intermediate).²⁴²

Scheme 46. δ -Specific hydroxylation of L-Leucine catalyzed by L-leucine dioxygenase from *Anabaena variabilis* (AvLDO). Additionally, AvLDO catalyzes the hydroxylation of L-norleucine or sulfoxidation of L-methionine and DL-ethionine. Specific activities (U mg⁻¹) determined for each substrate are given in parentheses.



Hydroxylation of L-leucine catalyzed by AVLDO yielded (2S,4S)-5-hydroxyisoleucine and upon sulfoxidation of L-methionine, the corresponding (*R*)-sulfoxide was obtained (75 %*de*), which is opposite to the diastereomer formed by *Btl*DO (Scheme 46).⁴⁹ AVLDO showed a clear preference for acidic reaction conditions and low to medium temperatures (optimum at pH 4 and 25°C), similar to what was observed for LdoA,²⁴⁰ while most other amino acid hydroxylases prefer a neutral pH.^{238,239,243} Specific activities determined for AVLDO were rather low (0.22-0.60 U mg⁻¹, Scheme 46), but yet comparable to other dioxygenases like LdoA or *Btl*DO.^{240,244}

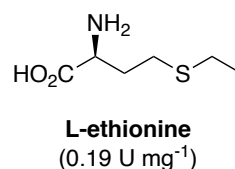
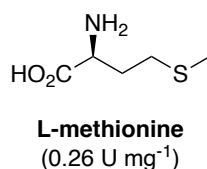
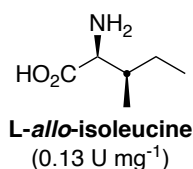
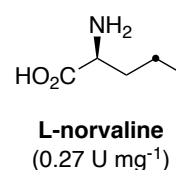
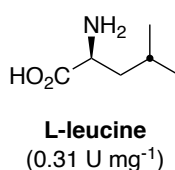
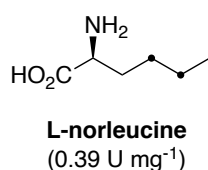
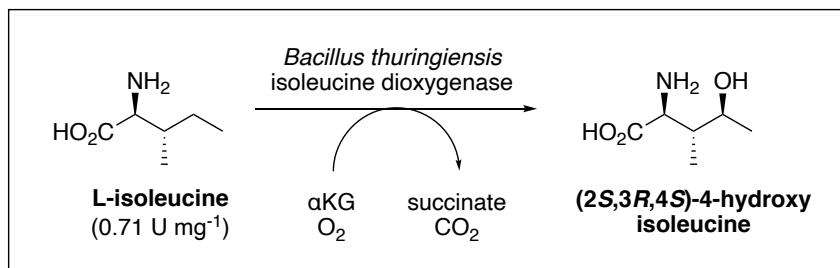
5.1.3.2 L-Isoleucine dioxygenase from *Bacillus thuringiensis*

The natural non-proteinogenic amino acid 4-hydroxyisoleucine was first identified in fenugreek seeds (*Trigonella foenum-graecum*) and has interesting insulinotropic properties.²⁴⁵ It was proposed that this compound is produced *via* hydroxylation of L-Ile, most probably mediated by a mono- or dioxygenase enzyme. Isoleucine hydroxylating activity dependent on ferrous iron, α -ketoglutarate, oxygen and ascorbate could indeed be detected in seedlings of fenugreek but the enzyme responsible for this transformation remained uncharacterized.²⁴⁶

In 2009, Kodera *et al.* identified an L-isoleucine dioxygenase in cell lysates of *Bacillus thuringiensis*, which, at that time, was the first identified Fe(II)/ α KG dependent dioxygenase acting on free aliphatic amino acids. It was demonstrated that the conversion of L-Ile by *Btl*DO resulted in (2S,3R,4S)-4-hydroxyisoleucine, which is equivalent to the product isolated from fenugreek seeds.^{238,247}

Hibi *et al.* further characterized the substrate scope of *Btl*DO, which, besides L-Ile, includes hydroxylation of several aliphatic amino acids and sulfoxidation of L-methionine and L-methionine. Interestingly, *Btl*DO showed differing regioselectivity for hydroxylation depending on the substrate. While C4-specific hydroxylation was predominant (L-isoleucine, L-norleucine, L-leucine, L-norvaline), L-norleucine was found to be alternatively hydroxylated at C5 and L-*allo*-isoleucine was curiously hydroxylated at C3 (Scheme 47).²⁴⁴

Scheme 47. C4-specific hydroxylation of L-isoleucine catalyzed by L-isoleucine dioxygenase from *Bacillus thuringiensis* (*BtlDO*). Additionally, *BtlDO* further accepts L-norleucine, L-leucine, L-norvaline and L-*allo*-isoleucine as substrates for hydroxylation (the respective regioselectivity is marked in each substrate) and L-methionine and L-ethionine for sulfoxidation. Specific activities (U mg⁻¹) determined for each substrate are given in parentheses.



Later on, the scope of sulfoxidation catalyzed by *BtlDO* was extended to *S*-methyl-, *S*-ethyl- and *S*-allyl-L-cysteine, which resulted in the formation of the respective (*S*)-sulfoxides.²⁴⁸ Quite interestingly, *BtlDO* also catalyzed dehydrogenation of 4-HIL to a limited extent (0.18 U mg⁻¹),²⁴⁴ but this reaction did not seem to be relevant under physiological conditions in *Bacillus thuringiensis*, as the majority of (2*S*,3*R*)-2-amino-3-methyl-4-ketopentanoate (AMKP) is rather produced by NAD⁺ dependent HIL dehydrogenase.²⁴⁹

5.1.4 Recent concepts

5.1.4.1 Coupling of substrate hydroxylation to central carbon metabolism

Smirnov *et al.* were the first to interconnect cosubstrate turnover of Fe(II)/αKG dependent dioxygenases and cellular carbon metabolism *via* the TCA cycle intermediates α-ketoglutarate and succinate, thereby coupling cell viability to dioxygenase activity. By blocking central metabolic routes of *E. coli* for succinate synthesis *via* knock-out of the *sucAB* (encoding α-ketoglutarate dehydrogenase)

and *aceAK* (encoding isocitrate lyase and isocitrate dehydrogenase kinase/phosphatase) genes, the carbon flux was channeled through IDO-catalyzed hydroxylation, thereby shunting the interrupted TCA cycle (Figure 36).²⁵⁰

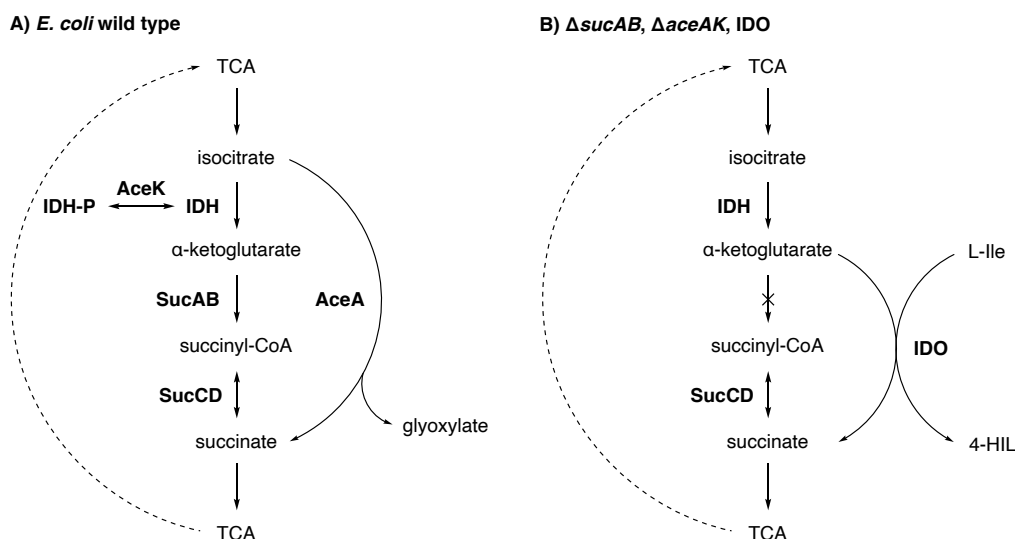


Figure 36. Central metabolic routes towards succinate in *E. coli* wild type (A) and *E. coli* Δ (*sucAB* and Δ *aceAK*) (B). In *E. coli* Δ , succinate production relies on the presence of IDO, which produces succinate from α -ketoglutarate upon turnover of the prime substrate L-isoleucine. IDH: isocitrate dehydrogenase, IDH-P: phosphorylated IDH, *AceK*: isocitrate dehydrogenase kinase/phosphatase, *SucAB*: α -ketoglutarate dehydrogenase, *SucCD*: succinyl-CoA synthetase, *AceA*: isocitrate lyase, IDO: L-isoleucine-4-hydroxylase, 4-HIL: 4-hydroxyisoleucine. Figure adapted from Smirnov *et al.*²⁵⁰

The resulting *E. coli* Δ could not grow in minimal medium supplemented with L-Ile unless IDO was present to bypass the interrupted TCA cycle. Additionally, expression of the Na^+ -dependent transporter of branched-chain amino acids (*BrnQ*) was upregulated to optimized uptake of L-isoleucine.²⁵⁰

Schmid and coworkers carried the concept of TCA cycle engineering in the context of Fe(II)/ α KG dependent dioxygenases forward and developed an optimized host strain for hydroxylation of proline catalyzed by P4H.⁵⁰ A previous study already revealed that knock-out of proline degradative pathways (Δ *putA*) could enhance both hydroxyproline formation rate and yield. Interestingly, the increased demand of α -ketoglutarate caused by P4H catalysis did not result in an enhanced flux towards α KG supply, thus indicating potential limitations in this central metabolite.²⁵¹ Similar to the concept introduced by Smirnov *et al.*, Schmid and coworkers alternately deleted *sucA* (encoding α KG dehydrogenase E1 subunit) or *sucC* (encoding succinyl-CoA synthetase β -subunit) together with

aceA (encoding isocitrate lyase) and *putA* (encoding proline dehydrogenase), resulting in the strains *E. coli* $3\Delta sucA$ and $3\Delta sucC$, respectively. While the growth of *E. coli* $3\Delta sucC$ in minimal medium did not rely on the presence of P4H due to supply of succinyl-CoA *via* the intact α KG dehydrogenase and succinate *via* fumarate reductase, growth of *E. coli* $3\Delta sucA$ was fully dependent of P4H activity. However, *E. coli* $3\Delta sucA$ suffered from reduced cell size (27 %) through shortage in lysine, methionine and diaminopimelate, the biosynthesis of which relies on succinyl-CoA. The rate of substrate hydroxylation could be further increased by overexpression of *putP* encoding the Na^+ /L-proline transporter.⁵⁰ This was in contrast to results obtained for *E. coli* $\Delta putA$, where P4H activity was limited by α -ketoglutarate availability.²⁵¹ Yet, in case of *E. coli* $3\Delta sucA$, competition for α KG could be relieved by deletion of *sucA*.⁵⁰

While in this approach, growth could be only restored to a limited extent by endogenous L-proline biosynthesis, recombinant *E. coli* $\Delta putA \Delta sucAB \Delta aceAK$ additionally overexpressing genes encoding γ -glutamyl kinase and glutamate-semialdehyde dehydrogenase (ProAB) for enhanced fermentative production of L-proline next to P4H was able to produce 31.0 g L⁻¹ *trans*-4-hydroxy-L-proline from glucose in a fed-batch approach.²⁵² Similar fermentative systems were also implemented in *Corynebacterium glutamicum*, yielding 21.7 g L⁻¹ hydroxyproline from glucose in batch fermentation.²⁵³ Sun *et al.* used a combined approach for efficient production of 5-hydroxy-leucine including enzyme engineering of LdoA from *Nostoc punctiforme* (*Np*LDO), metabolic engineering *via* TCA cycle ($\Delta sucA$, $\Delta aceA$), metabolic flux optimization and introduction of glutamate oxidase and catalase for increased α -ketoglutarate supply resulting in 5.6 g L⁻¹ of the target product.²⁵⁴

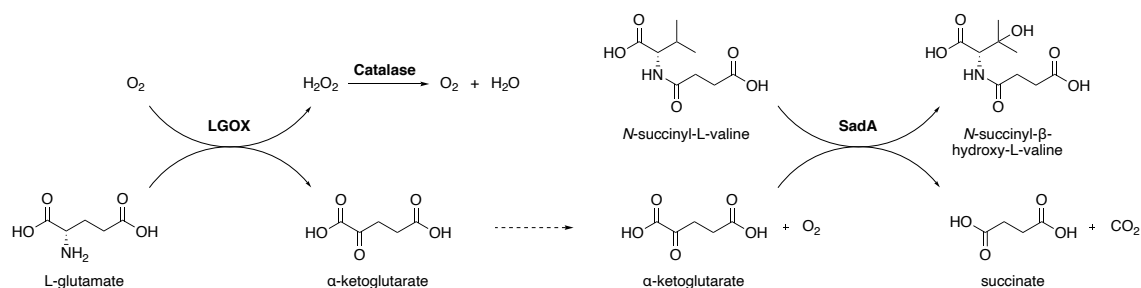
Schmid and coworkers further outlined the usability of *E. coli* knock-out strains for screening of active Fe(II)/ α KG dependent dioxygenases.⁵⁰ This was previously also demonstrated by Zhang *et al.* who screened a library of *Btl*DO variants generated by random mutagenesis *in vivo* by using *E. coli* $\Delta sucA \Delta aceA$ as selection host strain. The best identified IDO variant showed a 1.5-fold enhanced catalytic efficiency (k_{cat}/K_M) compared to the wildtype enzyme.²⁵⁵

5.1.4.2 *In-situ* cosubstrate generation for *in-vitro* applications

Despite the advantages offered by *in-vivo* approaches, including stabilization of the biocatalyst within the whole-cell environment and constant supply of cosubstrate α -ketoglutarate (see 5.1.4.1), such systems often suffer from native metabolic side-reactivities, mass transfer limitations or substrate toxicity, thus hampering operation at industrially relevant scales. In turn, *in-vitro* applications require (at least) stoichiometric supply of the cosubstrate α -ketoglutarate, which is irreversibly decarboxylated during catalysis.

In this regard, Busch *et al.* developed a one-pot/two-step enzymatic cascade combining *in-situ* production of α -ketoglutarate from L-glutamate and hydroxylation of *N*-succinyl-L-valine by the Fe(II)/ α KG dependent amino acid hydroxylase SadA from *Burkholderia ambifaria* (Scheme 48). By using immobilized enzymes, both downstream processing and recyclability of biocatalyst were significantly facilitated.²⁵⁶ A similar approach was previously also implemented by Sun *et al.* in whole-cell and cell-free systems.²⁵⁷

Scheme 48. Two-step enzymatic cascade including LGOX catalyzed oxidation of L-glutamate to yield α -ketoglutarate and subsequent hydroxylation of *N*-succinyl-L-valine catalyzed by SadA, which uses the produced α -ketoglutarate as a stoichiometric cosubstrate. Scheme adapted from Busch *et al.*²⁵⁶



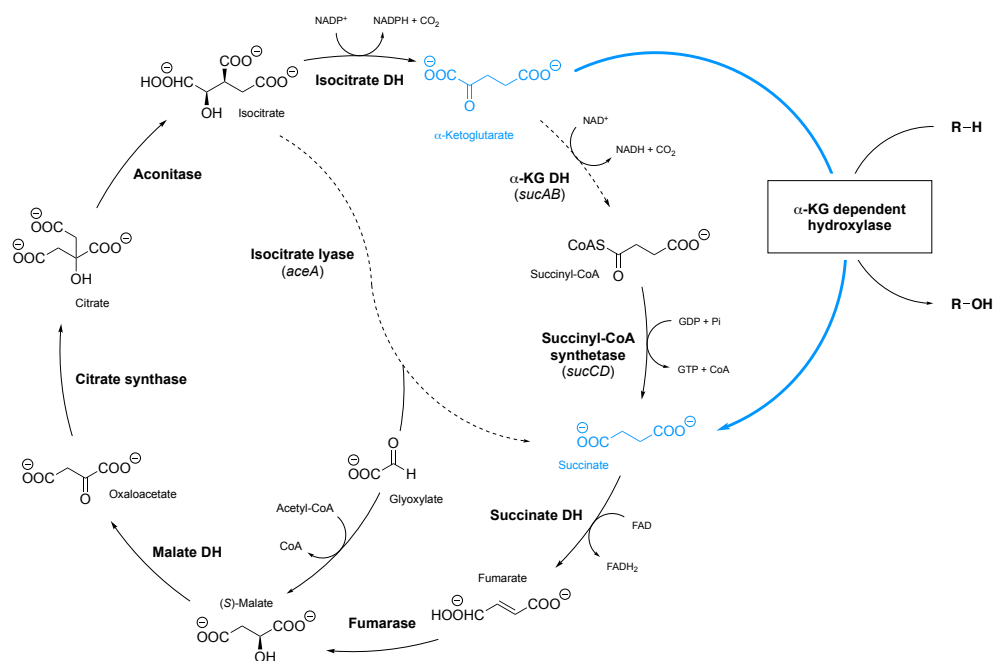
The authors demonstrated the usability of L-glutamate oxidation catalyzed by immobilized L-glutamate oxidase (LGOX) from *Streptomyces ghanaensis* on a 1 L scale at substrate concentrations of 73.6 g L⁻¹ (339 mM), overall resulting in a space-time yield of 14.2 g L⁻¹ h⁻¹. Besides the necessity of adding catalase to alleviate noxious effects caused by emerging hydrogen peroxide, optimized oxygen supply contributed significantly to the productivity of both reaction steps. Yet, the SadA catalyzed reaction step was limited to 40 mM substrate loading (9.3 g L⁻¹) and the heterogeneous biocatalyst suffered from severe loss of activity

after the first reaction cycle (10 % residual activity). Yet, separate immobilization of both biocatalysts allowed operation of both reaction steps at the respective optimal conditions in a consecutive fashion and furthermore enabled simple interchange between different dioxygenases if required. Yet, it was also pointed out that the implemented system does not amend the medium atom economy (54 %) of the overall process, which could be only overcome by a viable methodology for regenerating α -ketoglutarate directly from succinate.²⁵⁶

5.2 Aim of the project

Within this project, L-leucine dioxygenase from *Anabaena variabilis* (AVLDO), was the main target of our investigations. While Fe(II)/ α -KG-dependent dioxygenases are capable of oxy-functionalizing inactivated C-H bonds, their substrate scope is often limited or hard to rationalize. Additionally, despite the overall number of identified non-heme dioxygenases is rising steadily, structural information about this enzyme class is still limited. Even though family members mostly share a common jellyroll core fold, the detailed architecture differs considerably among different subclasses due to the large diversity of catalyzed reactions.²¹²

Scheme 49. Schematic representation of the *E. coli* Δ *sucA* selection strain, which is deficient in α -ketoglutarate dehydrogenase (Δ *sucA*) and isocitrate lyase (Δ *aceA*). The impaired formation of succinate and thus interruption of the tricarboxylic acid (TCA) cycle render the resulting strain practically unviable. In the presence of an α -KG dependent hydroxylase, the gap in the TCA cycle can be bypassed by formation of succinate as a coproduct during substrate hydroxylation (in blue).



In view of the limited knowledge about determinants of substrate scope, using a semi-rational enzyme engineering approach is unavoidable but requires the availability of a suitable selection and screening platform. In this respect, amongst others, Schmid and coworkers reported use of an *E. coli* knock-out strain, which, in itself, is unable to produce succinate, a central intermediate of the TCA cycle.

Thus, viability of this strain relies on the presence of an active α -KG dependent dioxygenase to bypass the interrupted TCA cycle by producing succinate as coproduct during substrate hydroxylation (Scheme 49).⁵⁰ It was claimed that by using this strain, enzyme libraries could be screened for dioxygenase variants converting a substrate of interest by supplementing this compound during cultivation.

A 3DM database was created based on the AvLDO crystal structure to approach better comparability of more loosely related dioxygenases, which should facilitate analysis of conserved structural features among distant homologues.²⁵⁸ Based on structural analysis, residues that might interfere with the accommodation of L-isoleucine were selected for site-saturation mutagenesis. By using degenerate codons only encoding a limited set of amino acids at each of the selected target residues, the diversity obtained at each position could be largely restricted to both chemically and spatially appropriate amino acids. The created library should be further subjected to *in-vivo* selection and enrichment by using the strain *E. coli* BL21(DE3) Δ *sucA* Δ *aceA* Δ *putA* (3Δ *sucA*). Based on their ability to grow on L-Ile containing minimal medium, putative isoleucine-converting AvLDO variants were selected and further characterized. In particular, the suitability of different methods to quickly assess dioxygenase activity, including *in-vitro* and resting cell reactions combined with TLC or GC-MS analysis, were examined.

5.3 Results and discussion

While previous studies aimed at implementing the selection strain *E. coli* 3Δ*sucA* and developing suitable conditions for enrichment of dioxygenase variants,^{259,260} the focus of this work was to test the established system. For this purpose, a library of AvLDO covering residues Y67, I70, F78, E158 and Y206 was generated by site-saturation mutagenesis (see 5.3.1-5.3.2.3). The created library was further subjected to selection based on their ability to utilize L-isoleucine as a substrate (see 5.3.3-5.3.4). Eventually, characterization of enriched variants was attempted and different strategies to confirm dioxygenase activity were tested (see 5.3.5-5.3.6).

5.3.1 Selection of target residues for site-saturation mutagenesis

Interesting target residues for site-saturation mutagenesis were selected by structural analysis of the AvLDO active site. The active site is made up of a reactive iron center, which is coordinated by H151, D153 and H237 in a characteristic His-X-Asp/Glu-X_n-His motif²⁶¹ and additionally by the co-substrate α-ketoglutarate. In case of the AvLDO crystal structure, the coproduct succinate was instead found (Figure 37). The remaining coordination sites are available for activation of oxygen, necessary for hydroxylating the substrate sidechain. This means that the substrate sidechain is facing the active center but binding of the substrate occurs at the opposite side of the cavity. In the related dioxygenase GriE, the amino acid group of the substrate is stabilized by Tyr59, Gln93, Trp116, Arg242 and backbone interactions to Val170.²⁶² However, corresponding spatially equivalent residues are absent in AvLDO. Yet, polar contacts between Tyr63, Arg142 and the sulfonic acid moiety of the buffer component *N*-Cyclohexyl-2-aminoethanesulfonic acid (CHES) are present, thus indicating their presumable involvement in substrate binding (Figure 37).

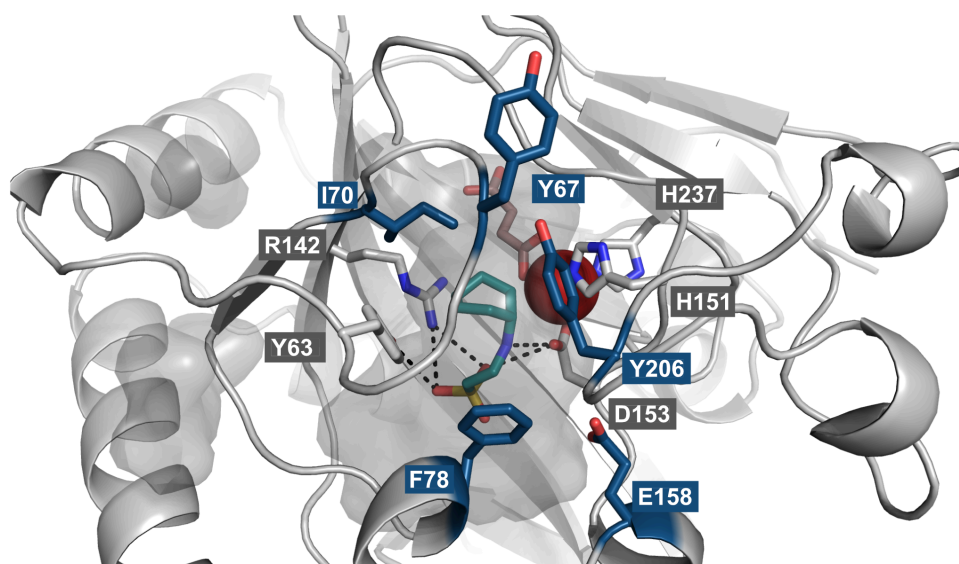


Figure 37. Representation of the AVLDO active site containing Fe^{2+} (red) coordinated by H151, D153, H237 (gray sticks) and succinate (pink). The buffer component *N*-Cyclohexyl-2-aminoethanesulfonic acid (CHES, in cyan) presumably takes the place where the substrate amino acid is usually located (polar contacts to Y63 and R142). The gray surface delimits the active site cavity. Residues selected for SSM are highlighted in blue (Y67, I70, F78, E158, Y206). Structure visualized with PyMOL.⁸⁸

Coming from the assumption that residues 63 and 142 are involved in substrate binding, residues lining the cavity between this site and the active iron center can potentially influence the accommodation of the substrate sidechain. In this regard, however, only Ile70 prominently points into the pocket, whereas most other residues in this area are presenting only their backbone. We further selected Tyr67 and Tyr206 which are located at opposing loops, which might be involved in closing-off the active site after substrate binding. Additionally, the bulky residues Phe78 and Glu158 located close to this assumed substrate entrance were selected (Figure 37, Table 22). Further information about the selected positions was sourced from the 3DM database,²⁵⁸ which was created based on the AVLDO structure (Table 22).

Except for Tyr67, all other selected residues were aligned in the core regions defined by the 3DM database. While this usually indicates the presence of spatially equivalent positions throughout the whole superfamily, closer examination revealed that 52 % (Ile70) or over 99 % (Phe78, Glu158, Tyr206) of all aligned sequences in the database have a gap (*i.e.* no structurally equivalent residue) at these positions. At the selected residues, amino acids present in

AvLDO fully match with the subfamily consensus (CP01QA subfamily) but slightly differ compared to the overall consensus (Table 22).

Table 22. Selected residues of AvLDO, their location and various information about these positions obtained from the 3DM database.

AvLDO res.	Y67	I70	F78	E158	Y206
Location in AvLDO	Loop at entrance	Facing binding pocket	Helix near entrance	Helix near entrance	Loop at entrance
3DM alignment position	Variable	55	63	143	191
Overall consensus^[a]	-	K (26 %)	F (74 %)	E (31 %)	Y (97 %)
Subfamily consensus^[b]	Y (36 %)	I (84 %)	F (86 %)	E (65 %)	Y (100 %)
Sequences with gaps^[c]	- (56 %)	52 % (14 %)	> 99 % (0 %)	> 99 % (4 %)	> 99 % (4 %)

[a] Consensus (and occurrence in %) of all sequences aligned in the 3DM database superfamily (107 403 sequences). [b] Consensus (and occurrence in %) of all sequences aligned in the CP01QA subfamily also including AvLDO (50 sequences). [c] Percentage of all superfamily and subfamily (in parentheses) sequences having a gap at this position.

Closer analysis of the amino acid distribution within the superfamily and subfamily CP01QA at all selected sites revealed that I70 (3DM residue 55) and E158 (3DM residue 143) are somewhat variable, but F78 (3DM residue 63) and particularly Y206 (3DM residue 191) were highly conserved (Figure 38). While this high degree of conservation can indicate the importance of these particular residues, it has to be kept in mind, that the majority of sequences (>99 %) possessed a gap at the respective positions. While the underlying structure-based multiple sequence alignment (SB-MSA) of the 3DM database should facilitate knowledge transfer of spatially and structurally equivalent positions also within distant homologs of the superfamily, residues lining the active site cavity often do not exhibit specific functions or interactions. They are thus not conserved both structurally and sequentially in many cases. This was also reflected by the fact that most of the residues selected upon visual inspection of the AvLDO active site cavity were either aligned in variable regions (Y67) or had a gap (*i.e.* no

structurally equivalent residue) at the respective positions (F78, E158, Y206; >99 % of all sequences).

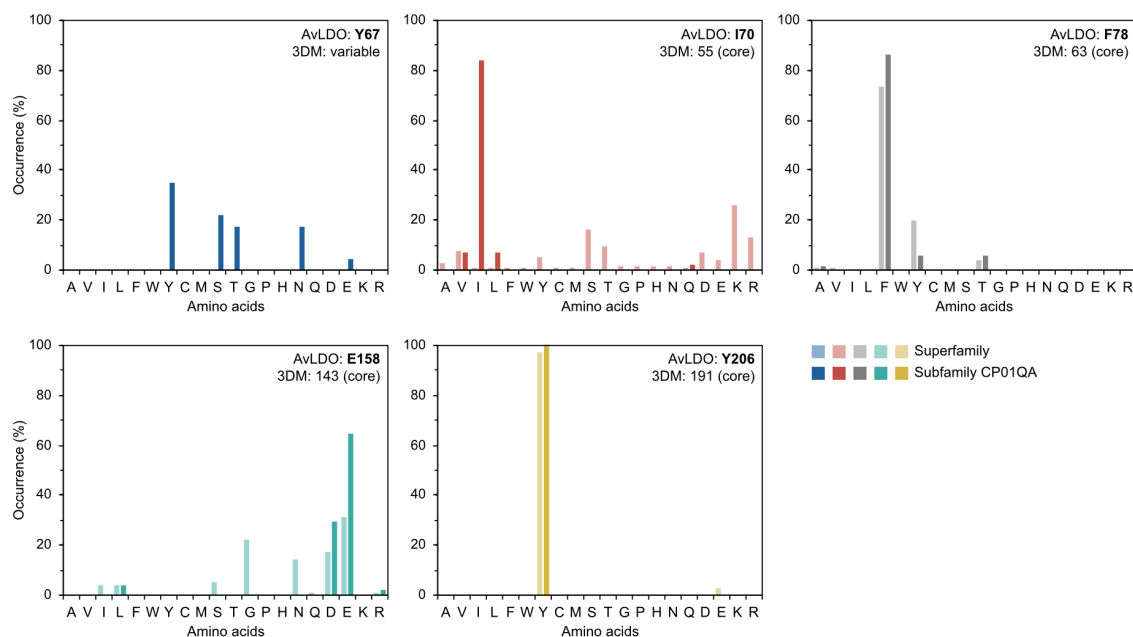


Figure 38. Occurrence (%) of different amino acids within the 3DM superfamily or subfamily CP01QA alignment at positions corresponding to AvLDO Y67 (3DM: variable), I70 (3DM: 55), F78 (3DM: 63), E158 (3DM: 143) and Y206 (3DM: 191).

For instance, the related and also δ -specific, dioxygenase GriE has a slightly more relaxed substrate scope than AvLDO but is also unable to convert L-isoleucine. Curiously, GriE accepts L-*allo*-Ile to a certain extent (over 40-fold reduced TTN) but seems to be otherwise sensitive towards substitutions at β -C.²⁶³ Structural analysis of GriE revealed that V115 and W116 potentially cause steric clashes during accommodation of L-Ile. Yet again, these residues were not aligned as part of the 3DM core structure. Also, L-isoleucine dioxygenase from *B. thuringiensis* (BtlDO) would be a particularly interesting structure to compare with.²⁴⁴ However, it has not been included in the here applied 3DM database, and a crystal structure was published only recently (PDB: 6LNH).

5.3.2 Preparation of an AvLDO library

Previously, already a small library of AvLDO variants was created.²⁵⁹ The small library included variation of Ile70 (*via* DYW codon) and Phe78 (*via* NTT + TAT codons). In total, 35 different amino acid double variants were comprised in this library (60 different in terms of codons). Due to the close proximity of both

targeted sites, mutations could be introduced at the same time by using a single set of mutagenesis primers in a QuikChange reaction.

Table 23. Overview of targeted mutation sites of AvLDO, the corresponding degenerate codons introduced at each position and the resulting number of codon and amino acid variants. All included amino acids and the respective codons used are listed below. Asterisks mark the wildtype codons (if included).

	Y67	I70	F78	E158	Y206
Degenerate codon	WMT	DYW	NTT + TAT	RVK	TDK
Codon variants	4	12	5	12	6
Amino acid variants	4	7	5	9	5 ^[a]
A (alanine)		GCA, GCT		GCG, GCT	
C (cysteine)					TGT
D (aspartic acid)				GAT	
E (glutamic acid)				GAG*	
F (phenylalanine)		TTT	TTT*		TTT
G (glycine)				GGG, GGT	
I (isoleucine)		ATA, ATT*	ATT		
K (lysine)				AAG	
L (leucine)		TTA	CTT		TTG
N (asparagine)	AAT			AAT	
R (arginine)				AGG	
S (serine)	TCT	TCA, TCT		AGT	
T (threonine)	ACT	ACA, ACT		ACG, ACT	
V (valine)		GTA, GTT	GTT		
W (tryptophan)					TGG
Y (tyrosine)	TAT*		TAT		TAT
* (stop)					TAG

[a] TDK includes six codons including a Stop codon, thus only encoding five different amino acids.

The here intended extensive library of AVLDO (Y67X/I70X/F78X/E158X/Y206X) encompassed 6300 unique amino acid quintuple variants (in 17 280 codon configurations) with the targeted mutational sites spanning over 139 residues in total (Table 23). Thus, the approach for site-saturation mutagenesis (SSM) had to be adapted accordingly. Our method of choice was an overlap extension PCR (OE-PCR) to create a megaprimer library (see 5.3.2.1) which is then used in a subsequent megaprimer PCR of whole plasmid (MEGAWHOP) to introduce the mutations into the expression plasmid (see 5.3.2.2).

Different degenerate codons, encoding only a limited set of amino acids, were chosen for each of the mutated positions (Table 23). For residue 67, the codon WMT (encoding N, S, T and Y) was chosen as the diversity of amino acids observed at this position among other subfamily members (Figure 38) was best represented. Instead of I70, which directly points into the active site cavity, a set of mostly small and hydrophobic amino acids encoded by DYW (A, F, I, L, S, T and V) was selected. Hydrophobic amino acids encoded by the NTT codon (F, I, L and V) in combination with tyrosine (TAT) were selected to substitute F78 sitting close to the substrate entry and binding site. For residue 158, a set of mostly charged and hydrophilic amino acids encoded by RVK (A, D, E, G, K, N, R, S and T) was selected, thereby respecting potential polar contacts established by the glutamic acid residue present in the wildtype. Finally, Y206, which was found to be highly conserved (Figure 38), was substituted by other mostly hydrophobic amino acids (C, F, L, W, Y and Stop) encoded by TDK. In any case, the selected codons also included the wildtype amino acid.

5.3.2.1 Megaprimer generation by overlap extension PCR (OE-PCR)

To perform multi-site SSM, a megaprimer library containing all targeted mutational sites was generated conceptually following a modified OE-PCR protocol described by An *et al.* (Figure 39).²⁶⁴ As illustrated in Figure 39, for each mutational site, a set of primers encoding the chosen restricted degenerate codons was designed, whereby position 67, 70 and 78 were treated collectively. Each primer pair (F_n/R_n , approx. 30-60 bp) shared a complementary region of 34-51 bp, including the aforementioned codons, which later provided compatible overlaps between the generated adjacent fragments. Additionally, two flanking primers F_0 and R_0 were required but contained no mutation sites (Figure 39).

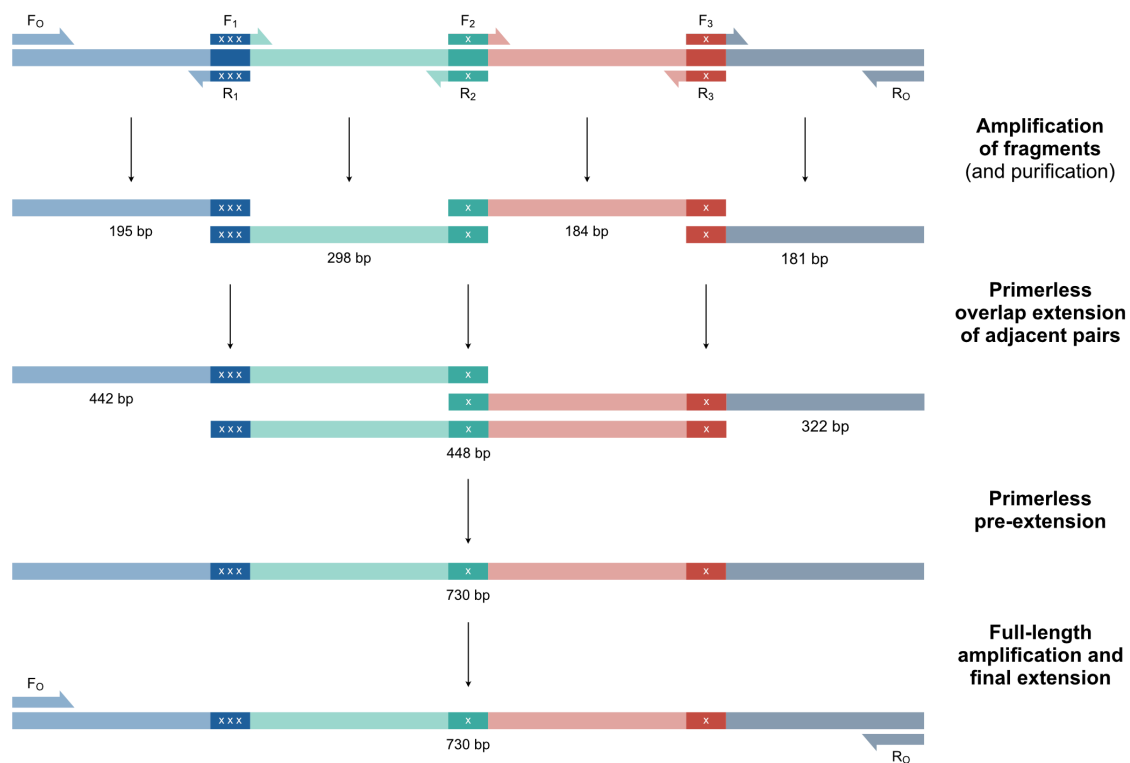


Figure 39. Schematic overview of the here applied overlap extension PCR (OE-PCR) procedure. In the first step, the individual fragments are amplified and purified. The primers used in this step contain degenerate codons (indicated by x) to introduce desired mutations in the overlapping regions of the adjacent fragments. In the second step, two adjacent fragments are mixed, respectively, and extended without primers. Next, all reaction mixtures obtained from step two are mixed and primerless pre-extension is performed. In the last step, the flanking primers F₀ and R₀ are added to the mixture for full-length amplification and final extension. For visual clarity, DNA is only presented as one strand.

In the first step, all four fragments (181-298 bp) were amplified separately in a standard PCR reaction initially using a gradient during annealing (55.0-64.1°C) (Figure 40). Nevertheless, a common annealing temperature and elongation time of 62°C and 20 s, respectively, turned out practical for all fragments.

After separating parental DNA by preparative agarose gel electrophoresis, the isolated fragments were mixed in a pairwise fashion in three separate reactions, and primerless overlap extension was executed in 15 cycles of denaturation (94°C, 20 s), annealing (56°C, 30 s) and elongation (68°C, 30 s). All three reactions were mixed right after, and the fragments were pre-extended in 20 cycles of denaturation and elongation before both flanking primers were added to the mixture. In this final part, the full-length assembly was amplified in 30 cycles of denaturation, annealing (60°C) and elongation (68°C, 1 min) followed by additional 10 cycles of post-extension (denaturation and extension). Again, the

resulting full-length fragment (730 bp) was purified by preparative agarose gel electrophoresis to separate truncated fragments potentially present in the mixture (Figure 40).

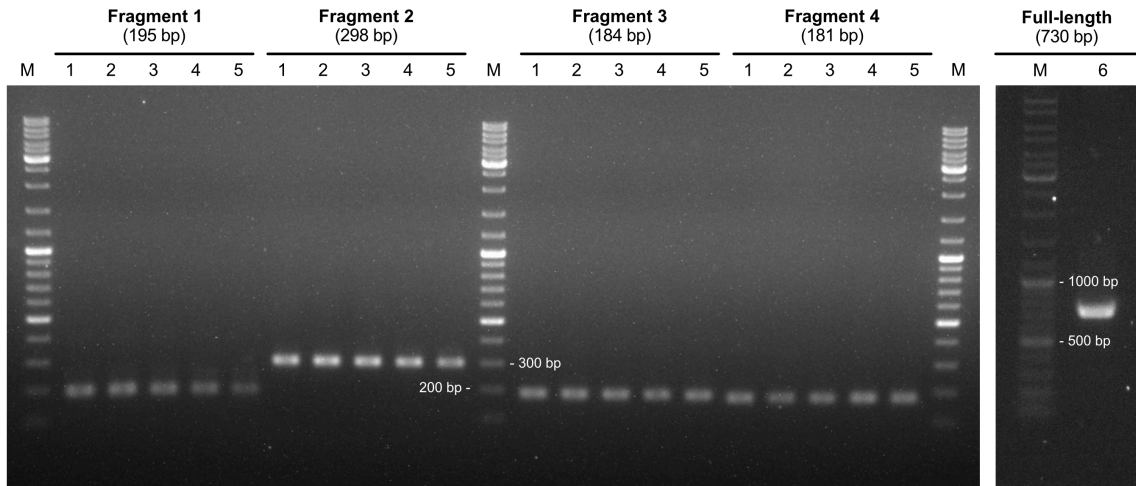


Figure 40. Left: Agarose gel analysis of amplified fragments of AvLDO required for OE-PCR. During annealing, a temperature gradient of 55.0 (1), 55.9 (2), 58.4 (3), 61.7 (4) and 64.1°C (5) was employed. pET28a_AvLDO_WT (6091 bp) was used as template. Right: Agarose gel analysis of full-length amplicon after OE-PCR. M: GeneRuler DNA Ladder Mix (Thermo Scientific).

It should be highlighted that by following this modified OE-PCR procedure, preparation of such a multi-fragment assembly can be easily accomplished within one day and considerably less time is needed if the individual fragments were already amplified and purified before.

To verify successful incorporation of desired mutations, the prepared megaprimer library was cloned into pJET2.1/blunt cloning vector and colonies were randomly picked after transformation. Selected clones were screened for correctly sized inserts *via* colony PCR (Figure S 8) and, except for clone 2 and 8, amplification of the inserted AvLDO megaprimer was successful. Clones 1 and 3-5 were further subjected to Sanger sequencing in order to analyze the incorporated base pair exchanges (Table 24).

Following these sequencing results, the generation of a megaprimer library *via* OE-PCR was successful. That is, while some clones contain the wildtype codon at single positions actually by design, no full wildtype sequence was detected. This also indicates successful removal of parental DNA, which would otherwise largely prevail and lead to almost sole detection of wildtype sequences.

Table 24. Summarized sequencing results of clones 1, 3, 4 and 5 in comparison to AVLDO WT. Wildtype codons are highlighted in gray.

Clone	Res. 67	Res. 70	Res. 78	Res. 158	Res. 206
WT	Y (TAT)	I (ATT)	F (TTT)	E (GAG)	Y (TAC)
1	Y (TAT)	A (GCT)	Y (TAT)	T (ACG)	L (TTG)
3	N (AAT)	S (TCA)	Y (TAT)	T (ACT)	F (TTT)
4	T (ACT)	A (GCT)	F (TTT)	S (AGT)	C (TGT)
5	N (AAT)	I (ATT)	I (ATT)	D (GAT)	Stop (TAG)

5.3.2.2 Megaprimer PCR of whole plasmid (MEGAWHOP)

A megaprimer PCR of whole plasmid (MEGAWHOP) was intended as described by Miyazaki.²⁶⁵ This methodology allows the introduction of a specific target DNA fragment (e.g. mutated megaprimers) by replacing a corresponding homologous section on a template plasmid *via* whole-plasmid PCR. After the removal of methylated parental DNA, the generated nicked circular plasmids are used to transform *E. coli* for *in-vivo* repair of the nicked strands (Figure 41).

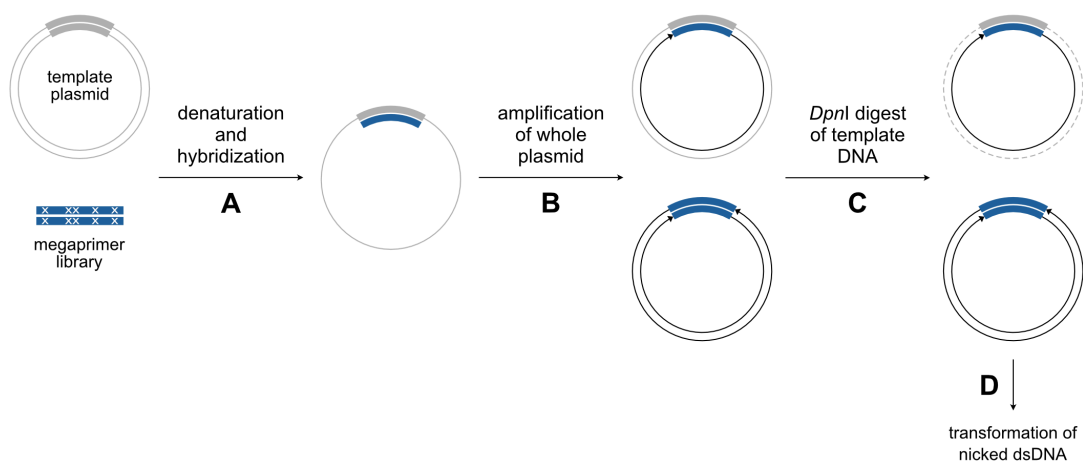


Figure 41. Schematic representation of a MEGAWHOP reaction. In the first step, the mutated megaprimers (blue) hybridize to the template plasmid (gray) containing the target gene without mutations. Then, the whole plasmid is amplified before methylated parental DNA is digested with *DpnI*. The nicked dsDNA generated by amplification can then be used to transform *E. coli* for *in-vivo* nick repair.

It was claimed that this approach is less time-consuming and yields more specific products compared to a classical restriction-ligation approach. This is particularly

important when assembling libraries, as the presence of template DNA or incorrect assemblies containing either no or several inserts is highly undesired.²⁶⁵ As was suggested, 500 ng of the previously prepared megaprimer library and 50 ng of template plasmid were used. Amplification was performed in a two-step fashion with denaturation at 95°C followed by annealing and elongation at 68°C (30 cycles). Subsequently, *DpnI* (10 U) was added to the mixture for template digestion before the reaction mixture was directly used to transform electrocompetent *E. coli* TOP10 cells. As will be outlined in the following Section 5.3.2.3 in more detail, the transformation of the prepared mutated plasmids is not only one of the most important steps on the way to a representative library but also one of the experimentally most challenging. While a certain number of transformants is mandatory to reach a satisfying library coverage, the transformation efficiency of nicked plasmids is usually significantly reduced. To partly account for this low efficiency, the produced amplicon amount can be influenced by increasing the number of reaction cycles or the quality of polymerase used. Both aspects were addressed by eventually increasing the number of cycles from 30 to 35 and by using a fresh aliquot of the polymerase. Yet also an efficient *DpnI* digest proved essential. In several attempts, a high background of wildtype AvLDO was detected after sequencing randomly picked transformants, thus indicating insufficient digest of parental DNA. A failed *DpnI* digest was mostly also reflected in an unusual high number of transformants. To prevent such issues, fresh enzyme was used, the amount of *DpnI* was doubled to 20 U per 50 µL reaction and the digest was run for at least 2 to 2.5 hours. It has yet to be clarified that the high efficiency of this procedure claimed by the author cannot be taken for granted, as it requires every critical step to work undoubtedly. Furthermore, an experiment's success is always an individual measure and closely linked to the individual requirements. Particularly in the present case of a library containing 6300 different amino acid variants, a considerable number of transformants is needed, which can neither be easily achieved when using other cloning or assembly techniques.

5.3.2.3 Transformation of created AvLDO library

The preferable generation of “smart libraries” by choosing degenerate codons only encoding a restricted set of amino acids over fully randomized SSM was

highlighted by Reetz *et al.*³¹ That is, the requirement of a certain degree of “oversampling” to ensure presence of a given variant (with a predefined probability) amongst the screened transformants, thus correlating library size and screening efforts. Particularly when striving for full randomization at several sites by using the NNN codon (covering all 64 codons), combinatorial explosion poses a severe limitation in terms of screening effort that has to be made. These correlations can be summarized in the following Equations (5)-(6), where T is the number of transformants to be screened to find a certain sequence with a specified probability P_i amongst all gene mutants present in the library (V). Consequently, the resulting oversampling factor (O_f) is close to three for reaching 95 % library coverage.³¹

$$T = -V \ln (1 - P_i) \quad (5)$$

$$O_f = T/V = -\ln (1 - P_i) \quad (6)$$

In total, the here created library of AvLDO contains 17 280 unique variants on gene level, resulting in 6 300 different amino acid variants. Following from Equation (5), a minimum of 51 766 transformants has to be screened for 95 % theoretical library coverage.

As briefly outlined in the previous Section 5.3.2.2, efficient transformation of the nicked plasmids obtained after MEGAWHOP was challenging, yet the maximum achievable library coverage was dependent on the diversity of transformed variants during this first and most critical transformation of *E. coli* TOP10. Once the (nicked) plasmid library was transformed and the intact supercoiled plasmid mixture reisolated, obtaining sufficient transformation efficiency with the selection strain *E. coli* BL21(DE3) $3\Delta sucA$ was typically not an issue anymore.

The general transformation procedure required an initial desalting step due to usage of electrocompetent *E. coli* TOP10 cells. A defined amount of this desalted mixture was used for transformation according to standard procedures. Cells were treated with LB-SOC medium after transformation and regenerated for 1.5-2 h. From the regenerated cell suspension, typically 5 μ L were plated on LB-Kan⁴⁰ plates, and the remaining volume was used to inoculate LB-Kan⁴⁰ medium. Colonies on plates were counted, and the total number of CFU present in the

liquid culture was extrapolated thereof. In case sufficient transformants were obtained, several colonies were randomly picked and analyzed by sequencing and the resulting plasmid library was isolated from the liquid culture.

In an initial attempt, only 7 280 transformants were obtained from transforming 5 μL of the desalted MEGAWHOP reaction mixture (Table 25, entry 1). The whole MEGAWHOP reaction was thus repeated, and 10 μL were transformed instead. Although a vast number of transformants was observed this time, only wildtype sequences were found during sequencing, thus indicating a problem during *DpnI* digest (Table 25, entry 2). We also compared whether desalting the reaction mixture *via* commercial PCR clean-up kits would provide us with better yields compared to micro-dialysis, but only an insufficient number of CFU was observed yet again (Table 25, entry 3).

Table 25. Summarized conditions used for MEGAWHOP and transformation of the created plasmid library.

Entry	Cycles	<i>DpnI</i> digest	Desalting	Competent cells	Volume	CFU ^[a]	Result
1	30	10 U, 2 h	dialysis	In-house ^[b] (80 μL)	5 μL	7 280	Insufficient CFU
2	30	10 U, 2 h	dialysis	In-house ^[b] (80 μL)	10 μL	<< 70 000 ^[c]	Only WT
3	30	10 U, 2 h	column	In-house ^[b] (80 μL)	10 μL	2 520	Insufficient CFU
4	35	20 U, 2 h	dialysis	Commercial ^[d] (100 μL)	5 μL	3 960	Insufficient CFU
5	35	20 U, 2 h	dialysis	In-house ^[b] (80 μL)	4 \times 5 μL ^[e]	63 280	Sufficient CFU

[a] The number of colony forming units (CFU) on the plates were extrapolated to the volume used for inoculating the liquid culture. [b] Preparation of electrocompetent *E. coli* TOP10 is described in Section 7.2.2.1. [c] Counting of CFU was discontinued at 500 CFU in 5 μL . [d] TOP10 Electrocomp™ *E. coli* cells were purchased from Invitrogen ($> 10^{10}$ cfu μg^{-1} supercoiled DNA). [e] Four parallel transformation reactions were performed, and the mixed-plasmid preparations isolated from each reaction were pooled.

To tackle issues originating from the MEGAWHOP reaction itself, the number of reaction cycles was raised to 35, and the amount of *DpnI* added increased. To also guarantee a high competency of the used cells, electrocompetent *E. coli* TOP10 cells were purchased. Due to the stated efficiency of $> 10^{10}$ cfu μg^{-1} supercoiled DNA, only 5 μL of reaction mix were used to transform 100 μL

commercial competent cells, yet the resulting number of transformants was underwhelmingly low (Table 25, entry 4).

In view of the many unsatisfying attempts made, the transformation procedure was parallelized, and four reactions were conducted simultaneously. The total number of transformants obtained from all four transformation reactions was 63 280, which was finally sufficing in terms of theoretical library coverage (Table 25, entry 5). The plasmid mixtures isolated from each of the four transformation reactions were pooled and used later on for the transformation of the selection strain *E. coli* BL21(DE3) Δ *sucA* (see 5.3.3.2). Eight clones were picked randomly and analyzed by sequencing (Table 26).

Table 26. Sequence analysis of selected *E. coli* TOP10 clones after transformation of generated plasmid library.

Clone	Res. 67	Res. 70	Res. 78	Res. 158	Res. 206
WT	Y (TAT)	I (ATT)	F (TTT)	E (GAG)	Y (TAC)
1	T (ACT)	F (TTT)	I (ATT)	G (GGT)	L (TTG)
2	Y (TAT)	I (ATT)	V (GTT)	T (ACG)	C (TGT)
3	N (AAT)	S (TCT)	L (CTT)	R (AGG)	L (TTG)
4	T (ACT)	V (GTT)	I (ATT)	T (ACG)	L (TTG)
5	T (ACT)	V (GTT)	V (GTT)	T (ACT)	L (TTG)
6	Y (TAT)	ambiguous	L (CTT)	ambiguous	C (TGT)
7	S (TCT)	I (ATA)		deletion	
8	T (ACT)	T (ACA)		deletion	

While the overall amino acid diversity detected at all targeted positions was satisfying, results obtained for clones 6-8 were unclear. While ambiguous sequencing results were obtained for clone 6 at two mutagenesis sites (residue 70 and 158), clone 7-8 both comprised a deletion of approximately 480 bp ranging from residue 78 to 239. In all three cases, the observed anomalies were confirmed by re-sequencing, and in the case of clone 7-8, the deletion was also observable after colony PCR followed by agarose gel electrophoresis (Appendix E, Figure S 9).

As isolation of correct sized megaprimers after OE-PCR was ensured by agarose gel excision, sources of this rather profound deletion could either be the MEGAWHOP rection itself or undesired *in-vivo* recombination events after transformation. While the used *E. coli* TOP10 strain is deficient in *recA* homologous recombination machinery, multitudinous other DNA repair mechanisms exist in *E. coli*.²⁶⁶ As deletion affected identical sites in both cases, the presence of specific sequence motifs triggering recombination within this section is likely.

5.3.3 Selection of AvLDO variants by enrichment culture

In order to select active dioxygenase variants, the selection host strain *E. coli* BL21(DE3) Δ *sucA* Δ *aceA* Δ *putA* (3 Δ *sucA*) was provided by Prof. Dr. Bruno Bühler (Helmholtz center for environmental research UFZ, Leipzig).⁵⁰ From the pool of created AvLDO variants, only those able to utilize L-isoleucine as a substrate should be enriched due to their ability to complement the interrupted TCA cycle of this knock-out strain by producing succinate as a co-product. The authors emphasized the great potential of the mentioned selection strain to screen large libraries for an extended or alternative substrate spectrum.⁵⁰

5.3.3.1 Verification of the *E. coli* BL21(DE3) Δ *sucA* Δ *aceA* Δ *putA* (3 Δ *sucA*) strain

In order to verify the received knock-out strain *E. coli* BL21(DE3) Δ *sucA* Δ *aceA* Δ *putA* (3 Δ *sucA*),⁵⁰ genomic DNA was isolated from this strain and from *E. coli* BL21(DE3) as a positive control. Suitable primers binding upstream and downstream of the genes encoding α -KG dehydrogenase E1 subunit (*sucA*) and isocitrate lyase (*aceA*) were designed according to the deposited genome of *E. coli* MG1655 (NC_000913.3, *aceA*: 4217109..4218413, *sucA*: 758706..761507). The third knocked-out gene of the selection host strain *E. coli* BL21(DE3) 3 Δ *sucA*, Δ *putA*, is involved in proline catabolism and was thus not of specific interest to this work. Amplification of both genes was attempted in separate reactions, and the resulting products were analyzed by agarose gel electrophoresis (Figure 42).

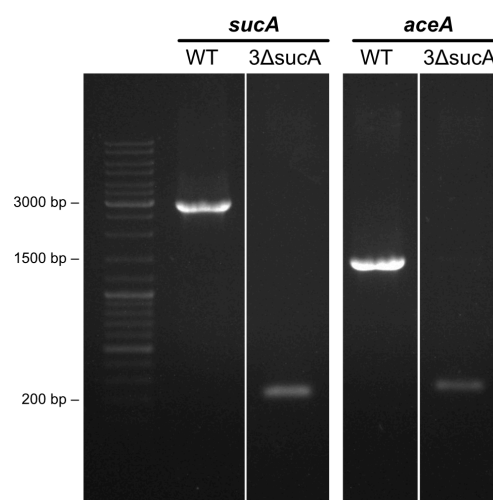


Figure 42. Agarose gel analysis of amplification of the regions containing the *sucA* (2802 bp; left) and *aceA* (1305 bp; right) gene with primers binding up- and downstream of the respective genes. Isolated genomic DNA of *E. coli* BL21(DE3) (WT) or *E. coli* BL21(DE3) Δ *sucA* Δ *aceA* Δ *putA* (3Δ *sucA*) was used as template. For the WT strain, PCR products of 2879 bp (*sucA*) and 1374 bp (*aceA*) are expected. M: GeneRuler DNA Ladder Mix (Thermo Scientific).

For the wildtype strain, amplification of the targeted *sucA* (2802 bp) and *aceA* (1305 bp) genes was successful, and bands of the respective PCR products (2879 and 1374 bp) were clearly visible. In contrast, when the triple knock-out *E. coli* BL21(DE3) 3Δ *sucA* strain was used as a PCR template, the amplicon size was considerably smaller (~200 bp), thus indicating the complete knock-out of both genes in this strain.

5.3.3.2 Preliminary enrichment experiments

In initial attempts to enrich isoleucine-converting variants from the AvLDO Y67X/I70X/F78X/E158X/Y206X library, the protocol previously established and applied for the enrichment of AvLDO I70X/F78X variants was used (Table 27).²⁵⁹ The workflow was initiated by the transformation of electrocompetent *E. coli* BL21(DE3) 3Δ *sucA* cells with the plasmid library or AvLDO WT. A small aliquot of the regenerated cell suspension was spread on LB-Kan⁴⁰ plates to monitor transformation efficiency (2.63×10^7 cfu μ g⁻¹ for pET28a(+) and 8.1×10^5 cfu μ g⁻¹ for pET28a_AvLDO_WT) and a defined amount the remaining volume was used to inoculate an overnight culture (LB-Kan⁴⁰, 37°C). Based on the optical density (OD₆₀₀) of the culture, the volume required to inoculate the M9 cultures to an OD₆₀₀ of 0.1 was calculated. The respective volume was withdrawn from the ONCs and cells were washed twice with M9 medium before suspending them

again in M9 medium to inoculate the first M9-Kan⁴⁰/IPTG/L-Ile culture (t_0). As the cells adapted only very slowly to the minimal growth conditions, the first culture was incubated for 72 h before diluting it 400-fold into fresh medium. All other samples were typically cultivated in 24 h intervals at 30°C. From each culture, small aliquots suitable to obtain single colonies were spread on M9-Kan⁴⁰/IPTG/ α KG/L-Ile plates and incubated at 30°C until colonies were visible (2-3 d). Additionally, at the end of each cultivation interval, the remaining culture volume was harvested, and cells were kept at -20°C for isolation of mixed plasmids from each cultivation step if required.

Table 27. Summarized conditions tested for enrichment of AvLDO Y67X/I70X/F78X/E158X/Y206X library variants in medium containing L-isoleucine as substrate.

Conditions	A	B	C	D
Cultivation vessel	Test tubes (17 mL)	Test tubes (17 mL)	PP tube (50 mL)	Baffled flask (100 mL)
Medium	M9-Kan ⁴⁰ /IPTG/L-Ile			
Culture volume	10 mL	6 mL	10 mL	20 mL
Pre-culture	LB-Kan ⁴⁰ (50 mL), 37°C, 150 rpm, overnight			
Inoculation of M9	OD ₆₀₀ = 0.1 (cells washed with M9 medium)			
Transferred volume	25 μ L (1:400)	15 μ L (1:400)	25 μ L (1:400)	50 μ L (1:400)
Cultivation intervals	24 h; first M9 culture (t_0) for 72 h			
Temperature	30°C			

Initially, test tubes containing 10 mL of medium were used for enrichment (Table 27, A), but the determined cell densities were much lower than previously observed.²⁵⁹ Also, when reducing the culture volume to 6 mL for better mixing and aeration (Table 27, B), growth was only slightly enhanced (Figure 43). Due to complete loss of the reference strain harboring AvLDO WT and overall low growth, this attempt was stopped after five cultivation intervals. Alternatively, in the following enrichment experiment, 50 mL sterile PP tubes (10 mL medium) and 100 mL baffled flasks (20 mL medium) were used as cultivation vessels (Table 27, C-D). Yet as before, only low growth was observed, and mostly *E. coli*

$3\Delta sucA$ [AvLDO_WT] grew to higher optical densities than the respective strain harboring the library (Figure 43).

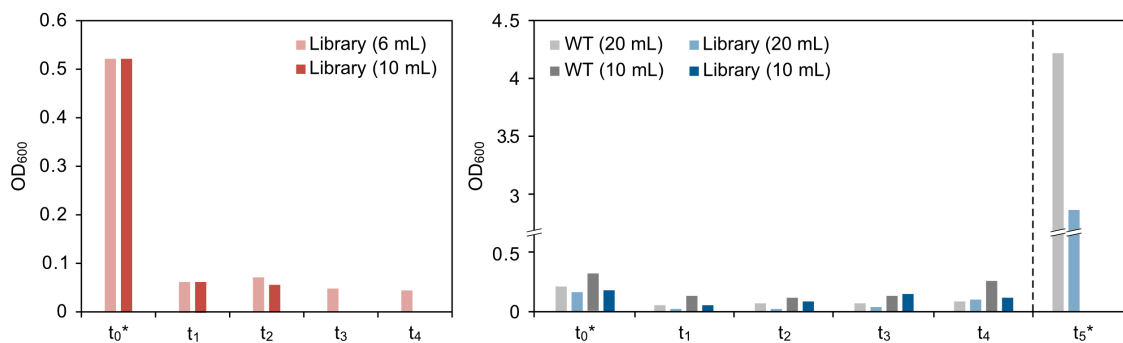


Figure 43. Optical density (OD₆₀₀) determined after each 24 h cultivation interval (72 h if marked with asterisk) of *E. coli* $3\Delta sucA$ harboring either AvLDO WT or AvLDO Y67X/I70X/F78X/E158X/Y206X library. Cultivation followed conditions A and B (left) or C and D (right) as stated in Table 27. Raw data is presented in Appendix H (Table 62).

While this could result from the large diversity within the library, consequently including a certain share of inactive variants, the predominant growth of the strain harboring AvLDO WT did not diminish even after 14 cultivation intervals (Appendix H, Table 62). Instead, after an intermediary 72 h incubation interval (t₅), the determined optical density was many times higher than in the previous samples, thus indicating undesired metabolic adaption to the minimal growth conditions, as the same behavior was then also observed in all following cultures (Figure 43 and Appendix H, Table 62). When further analyzing the colonies obtained from each cultivation step, it became apparent that for both *E. coli* $3\Delta sucA$ harboring AvLDO WT or library variants, colony size was comparable and other than expected, no differently sized colonies were observed within the library samples. These observations corroborated the observed growth pattern of the studied strains, according to which successful enrichment of library variants is unlikely or at least not apparent under the applied conditions.

5.3.3.3 Enrichment of AvLDO Y67X/I70X/F78X/E158X/Y206X variants

In view of the preliminary results obtained from the previous enrichment experiments, several parameters were readjusted. Following the suggestions of Zhang *et al.*, the liquid minimal medium was supplemented with α -ketoglutarate (M9-Kan⁴⁰/IPTG/ α KG/L-Ile) as well, and several aliquots of the regenerated cell suspension were directly spread on M9-Kan⁴⁰/IPTG/ α KG/L-Ile plates for

selection.²⁵⁵ Based on previous experience, baffled flasks were used for cultivation due to superior mixing and aeration compared to other reaction vessels tested (see 5.3.3.2). Further, preparation of an overnight culture in rich medium (LB-Kan⁴⁰) was omitted, and the transformation mixture was directly used to inoculate M9 cultures, which were then cultivated in 24 h intervals and subsequently diluted 100-fold into fresh medium. For practical reasons, cultivation at 37°C was also taken into consideration (Table 28).

Table 28. Summarized conditions applied for enrichment of AvLDO Y67X/I70X/F78X/E158X/Y206X library variants in medium containing L-isoleucine as substrate.

Conditions	E	F
Reaction vessel	Baffled flask (100 mL)	
Medium	M9-Kan ⁴⁰ /IPTG/ α KG/L-Ile	
Culture volume	20 mL	
Pre-culture	none	
Inoculation of M9	600 μ L regenerated transformation mixture	
Transferred volume	200 μ L (1:100)	
Cultivation intervals	24 h	
Temperature	30°C	37°C

This time, growth of the library containing strain surpassed the growth of *E. coli* 3 Δ *sucA* [AvLDO_WT] at each stage (Figure 44). After two cultivation cycles, the culture incubated at 30°C was discontinued, as similar growth patterns were also observed at 37°C and facilities to incubate at this temperature were permanently available. Starting from culture t_2 , differences in optical density observed for wildtype and library harboring strains were less pronounced as in the earlier stages of the experiment. Also, the fact that both strains were growing to higher densities with every cultivation interval indicated advancing adaption to minimal growth conditions, which could consequently hamper the imposed selection pressure at later stages of the experiment (Figure 44).

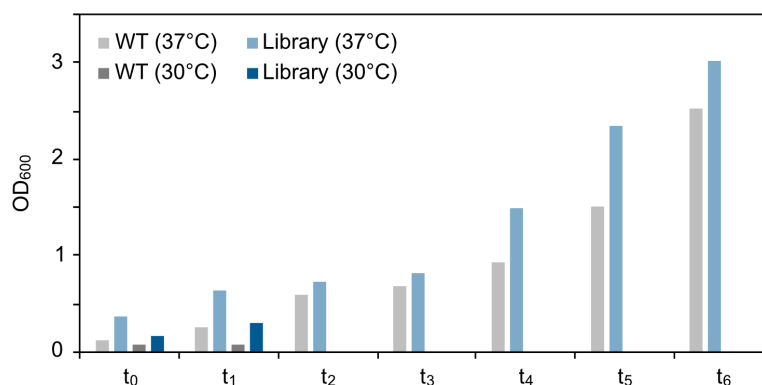


Figure 44. Optical density (OD₆₀₀) determined after each 24 h cultivation interval of *E. coli* 3Δ*sucA* harboring either AvLDO WT or AvLDO Y67X/I70X/F78X/E158X/Y206X library. Cultivation followed conditions E and F as stated in Table 28. Raw data is presented in Appendix H (Table 62).

During the enrichment experiment, small aliquots of each culture were spread again on M9-Kan⁴⁰/IPTG/αKG/L-Ile plates. Unlike in previous experiments, considerable differences in colony size between *E. coli* 3Δ*sucA* [AvLDO_WT] and *E. coli* 3Δ*sucA* [AvLDO_library], but also within the library sample itself were observed, particularly in the earlier stages of the experiment (Figure 45).

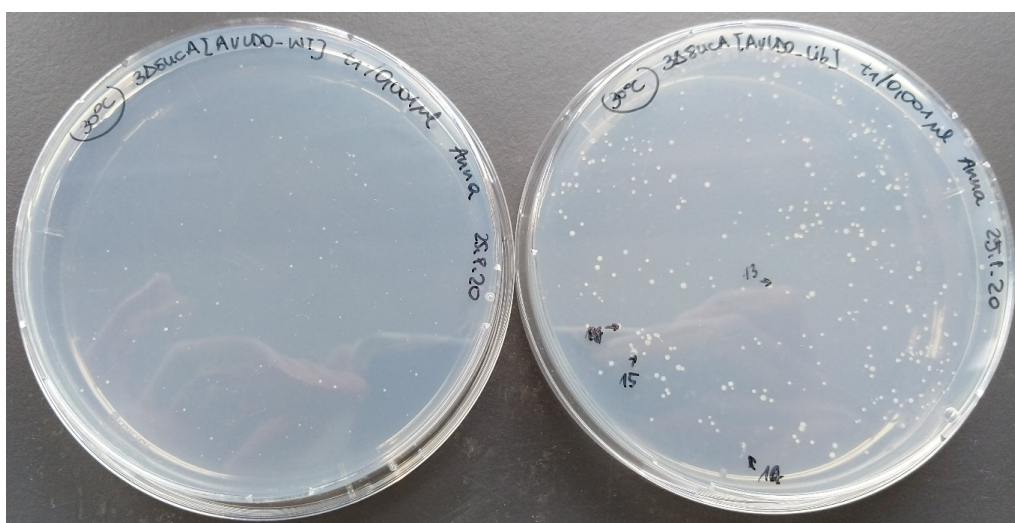


Figure 45. M9-Kan⁴⁰/IPTG/αKG/L-Ile agar plates on which t₁ cultures (30°C) of *E. coli* 3Δ*sucA* [AvLDO_WT] (left) and *E. coli* 3Δ*sucA* [AvLDO_library] (right) were spread (1×10⁻³ μL).

With progressing experiment, differences in colony size were getting less pronounced, similarly to what was observed already for the liquid cultures (Figure 44). This could either indicate advancing enrichment of improved variants or increasing adaption and thus decreasing selection pressure.

5.3.3.4 Selection of enriched AvLDO Y67X/I70X/F78X/E158X/Y206X variants

To test the enrichment procedure, 35 clones were picked for rescreening from the M9-Kan⁴⁰/IPTG/ α KG/L-Ile plates prepared at every stage of the enrichment experiment. Colonies were selected based on apparent differences in colony size (transformation, culture t₀-t₂) or randomly (culture t₃-t₆) (Table 29).

Table 29. List of selected clones 1-35 together with the sample from which the colonies were picked.

Clone	Origin	Clone	Origin	Clone	Origin
1	Transformation	13	t ₁ (30°C)	25	t ₄ (37°C)
2	Transformation	14	t ₁ (30°C)	26	t ₄ (37°C)
3	Transformation	15	t ₁ (30°C)	27	t ₄ (37°C)
4	Transformation	16	t ₂ (37°C)	28	t ₅ (37°C)
5	Transformation	17	t ₂ (37°C)	29	t ₅ (37°C)
6	t ₀ (37°C)	18	t ₂ (37°C)	30	t ₅ (37°C)
7	t ₀ (37°C)	19	t ₂ (37°C)	31	t ₅ (37°C)
8	t ₀ (37°C)	20	t ₃ (37°C)	32	t ₆ (37°C)
9	t ₀ (30°C)	21	t ₃ (37°C)	33	t ₆ (37°C)
10	t ₀ (30°C)	22	t ₃ (37°C)	34	t ₆ (37°C)
11	t ₁ (37°C)	23	t ₃ (37°C)	35	t ₆ (37°C)
12	t ₁ (37°C)	24	t ₄ (37°C)		

Notably, Zhang *et al.* selected variants of *BtlDO* solely by plating the transformation mixture on minimal plates.²⁵⁵ While this approach seemed to be straightforward, relevant parameters like library size and total number of colonies screened remained unstated, thus rendering the significance of this selection procedure rather questionable. While the use of the selection strain *E. coli* Δ sucA allows sorting out of inactive variants, thereby lowering overall screening efforts, an additional enrichment of superior variants would be desirable, particularly in the case of a larger library.

5.3.4 Rescreening of selected AvLDO variants

The previously applied rescreening procedure followed the initially established technique for recording growth curves with the knock-out strain *E. coli* 3 Δ *sucA*. Transformation of the knock-out strain with the respective plasmid followed by continuous cultivation and monitoring of cell density (OD₆₀₀) of the resulting recombinant strain in minimal medium (M9 with antibiotics and substrate of interest) containing IPTG for induction. However, cultures had to be typically maintained over elongated periods of 150 h, thus rendering the procedure extremely prone to contaminations or undesired adaption to the minimal growth conditions. This obstacle was already encountered during earlier experiments.^{259,260}

Due to above-mentioned issues, an alternative rescreening approach described by Zhang *et al.* was assessed instead.²⁵⁵ According to this procedure, the optical density (OD₆₀₀) was determined after two consecutive cultivation intervals of 24 h in minimal medium supplemented with antibiotics, IPTG, substrate and, additionally, α -ketoglutarate.

5.3.4.1 Rescreening of selected AvLDO I70X/F78X variants

To test this alternative rescreening approach, variants selected from the previously created small library of AvLDO,²⁵⁹ namely variant B (I70T), D (I70L) and H (I70T/F78I), were submitted once more to a rescreening following the adapted conditions. *E. coli* 3 Δ *sucA* harboring wildtype AvLDO (Leu⁺, Ile⁻) and BtlDO (Leu⁺, Ile⁺) were used as reference strains. In view of scalability, glass test tubes containing 5 mL of the respective medium were used for cultivation. In contrast to previously used media, the M9 medium was additionally supplemented with α -ketoglutarate. This slight modification helped to considerably shorten the lag-phase of cultures resulting from general adaption to minimal growth conditions combined with the demand of α KG to enable the dioxygenase-catalyzed reaction and thus, growth.

Fresh M9-Kan⁴⁰/IPTG/ α KG medium containing either L-Leu or L-Ile was inoculated with a single colony of the respective *E. coli* 3 Δ *sucA* recombinant strain and incubated for 24 h at 37°C. Then, the optical density was determined (t_1), and this pre-culture was used to inoculate fresh medium (50 μ L inoculation

volume), which was again incubated for 24 h. The optical density (t_2) was determined again, and the culture supernatant was kept for further analysis. The growth pattern observed for each of the studied strains is summarized in Figure 46.

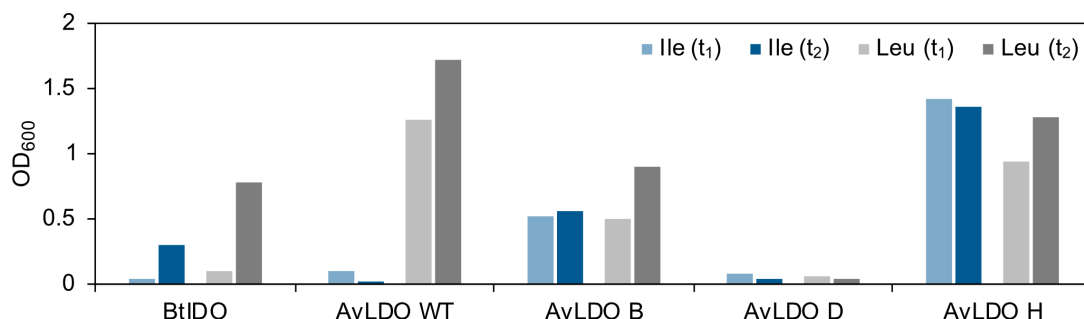


Figure 46. Determined optical density (OD_{600}) of *E. coli* $3\Delta sucA$ harboring *BtlDO*, *AVLDO* WT or variants B, D and H after two consecutive cultivation intervals in M9-Kan⁴⁰/IPTG/ α KG containing either L-Leu or L-Ile (5 mL, 37°C). After 24 h (t_1), 50 μ L of the culture were used to inoculate fresh medium, which was again incubated for 24 h (t_2). Raw data is shown in Appendix H (Table 63).

Prompted by the work of Zhang *et al.*, who monitored substrate conversion in the culture supernatant by HPLC,²⁵⁵ analysis of the here obtained supernatants *via* TLC was attempted. While the feasibility of analyzing *in-vitro* dioxygenase-catalyzed reactions *via* TLC analysis was already demonstrated before,²⁶⁰ it would specifically enable quick assessment of the *in-vivo* rescreening cultures without requiring laborious sample preparation or extensive instrumental set-up, such as needed for HPLC analysis. In general, the observed growth pattern matched with product formation detected after TLC, albeit differences in expression levels and specific activities make a direct comparison of *BtlDO* and *AVLDO* unfeasible. In a medium containing L-leucine, putative product formation was observed for *BtlDO*, *AVLDO* H and, to a lesser extent, for *AVLDO* B, due to appearance of a spot at a similar height compared to 4-hydroxy-L-isoleucine (HIL). Largely, these results were in agreement with the observed growth behavior (Figure 46) with the exception of *AVLDO* WT, where conversion of L-Leu would be also expected. In a medium with L-isoleucine, product formation seemed much weaker but was still visible for variant H. Quite interestingly, the appearance of an additional, prominent spot quite at the bottom ($R_f \sim 0.1$) seemed to be also connected to growth (*i.e.* substrate conversion), as it was visible for *BtlDO*, *AVLDO* WT and variants B and H in M9 with L-Leu and for *BtlDO* and

AvLDO variants B and H in M9 with L-Ile (Figure 47). Yet, identity of this compound remains unclear to date, and further investigations are necessary to unravel both its identity and role in this assay.

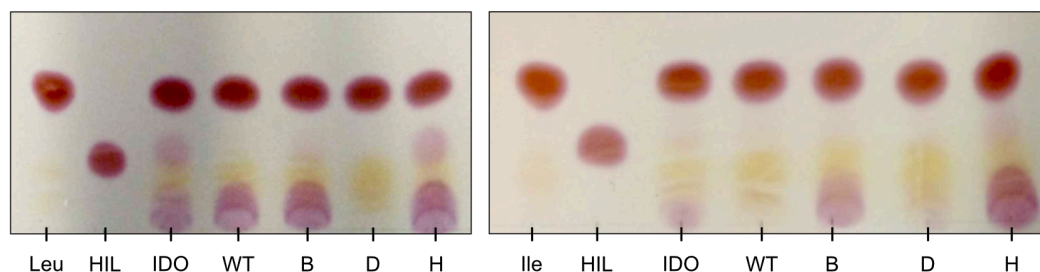


Figure 47. TLC analysis (1-butanol/AcOH/H₂O = 22:5:3, ninhydrin) of culture supernatants obtained from rescreening of *Bt*IDO (IDO), AvLDO WT (WT) and variants B, D and H in M9-Kan⁴⁰/IPTG/αKG/L-Leu (left) or M9-Kan⁴⁰/IPTG/αKG/L-Ile (right). The respective M9 media (Leu or Ile) and 4-Hydroxy-L-isoleucine (HIL) were used as reference.

Quite remarkably, AvLDO variant D seemed to be inactive towards both L-Leu and L-Ile as neither growth nor substrate conversion could be detected for this variant (Figure 46-Figure 47). In the rescreening performed in previous work, this variant showed a similar growth behavior as AvLDO WT with a sharp increase of optical density after 80 h of cultivation.²⁵⁹ This could be, however, also the result of a contamination or undesired adaption, as the strains harboring variants B and H passed through exponential growth already within the first 20 h.

5.3.4.2 Rescreening of selected AvLDO Y67X/I70X/F78X/E158X/Y206X variants

The simple setup of the here established rescreening procedure allowed handling of several samples in parallel. Selected clones (see 5.3.3.4) were rescreened in two batches (clone 1-15 and 16-35), each including *E. coli* 3Δ*sucA* [AvLDO_WT] as a control strain. Again, M9-Kan⁴⁰/IPTG/αKG medium supplemented with L-Ile was used (5 mL in test tubes) and inoculated with cell material from plates containing the re-streaked clones. The resulting cultures were incubated in two consecutive cycles of 24 h at 37°C (Figure 48).

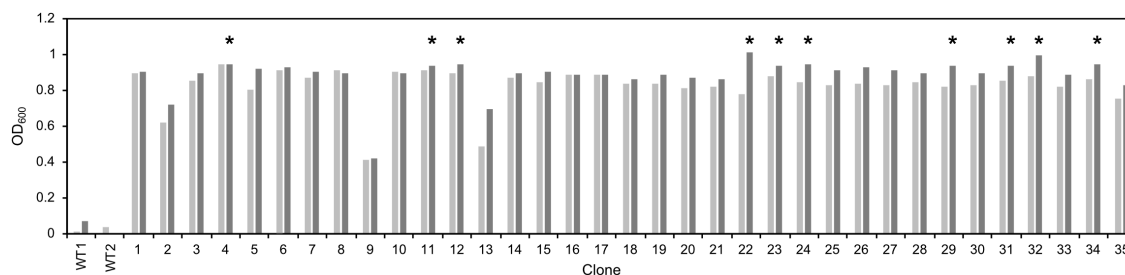


Figure 48. Determined optical density (OD₆₀₀) of *E. coli* 3Δ*sucA* harboring AvLDO WT or variants 1-35 after two consecutive cultivation intervals in M9-Kan⁴⁰/IPTG/αKG/L-Ile (5 mL, 37°C). After 24 h (t₁), 50 μL of the culture were used to inoculate fresh medium, which was again incubated for 24 h (t₂). Asterisks mark the ten clones reaching highest optical density after the second cultivation interval (t₂). Raw data is shown in Appendix H (Table 64).

All screened clones exhibited significantly faster growth on minimal medium containing L-isoleucine (OD₆₀₀ at t₂ of 0.422-1.015) compared to wildtype harboring cultures (OD₆₀₀ at t₂ of 0.003-0.070). This either indicates successful enrichment of L-Ile-converting AvLDO variants or advancing adaption of the studied strains to the minimal growth conditions (see 5.3.4.3). Similarly, as before, culture supernatants were analyzed by TLC. Yet, slightly different conditions were used (acetone/1-butanol/AcOH/H₂O = 7:7:2:4), as they enabled faster separation compared to previous conditions (1-butanol/AcOH/H₂O = 22:5:3) (Figure 49).

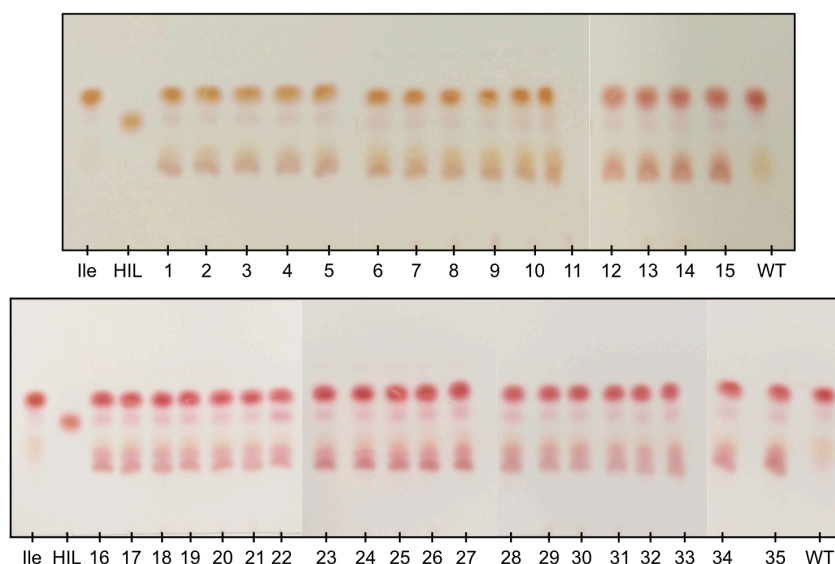


Figure 49. TLC analysis (acetone/1-butanol/AcOH/H₂O = 7:7:2:4, ninhydrin) of culture supernatants obtained from rescreening selected AvLDO library variants 1-35 and wildtype (WT) in M9-Kan⁴⁰/IPTG/αKG/L-Ile. The respective M9 medium (Ile) and 4-Hydroxy-L-isoleucine (HIL) were used as reference.

For all of the 35 screened variants, a spot that could be associated with product formation (*i.e.* similar R_f compared to HIL) was observed. As observed in the previous rescreening (see 5.3.4.1), a prominent spot appeared on the lower end for all cultures showing significant growth. During analysis of clone 16-35 this spot was even further separated into at least two distinct compounds. It should also be noted that an unidentified compound, appearing at a similar height to the assumed product, could be detected in some cases in the reference sample (respective M9 medium) and in the wildtype control as well. This further complicated unambiguous detection of product formation. Furthermore, it remains unclear to which extent *E. coli* can export the putatively formed hydroxy L-isoleucine.²⁵⁵

While ninhydrin is a relatively selective staining reagent, which mostly reacts with primary amines (usually pinkish or violet spots) and secondary amines (usually yellowish),²⁶⁷ a plethora of metabolites is produced during the *in-vivo* selection procedure, thus complicating assignment and identification of the compounds detected *via* TLC. Further, evidence was found in the literature that ninhydrin possibly also reacts with ammonium derivatives. Among the tested compounds, ninhydrin reacted with several organic ammonium salts, including succinate, oxalate (both deep reddish violet), lactate, malate, citrate or acetate (all violet) at 1 % concentration or less.²⁶⁸ It is likely, that *E. coli* produces various of these compounds in the course of the experiment, but it would be of particular interest to our experiment if ammonium succinate is amongst the prominent spots detected, as the appearance of these spots and culture growth were closely connected in the performed rescreening experiments. In this respect, supplementing experiments are currently in preparation.

5.3.4.3 *Critical assessment of selected AvLDO Y67X/I70X/F78X/E158X/Y206X variants*

It has to be taken into account that in contrast to preliminary rescreening attempts (see 5.3.4.1), the strains described in Section 5.3.4.2 were already exposed to the liquid minimal medium during enrichment (6 d) and were plated and stored on solid minimal medium (2-6 d) before selection and rescreening. At the same time, intermediary isolation of plasmid DNA from all clones of interest was deliberately omitted due to economic issues and limited scalability arising thereof.

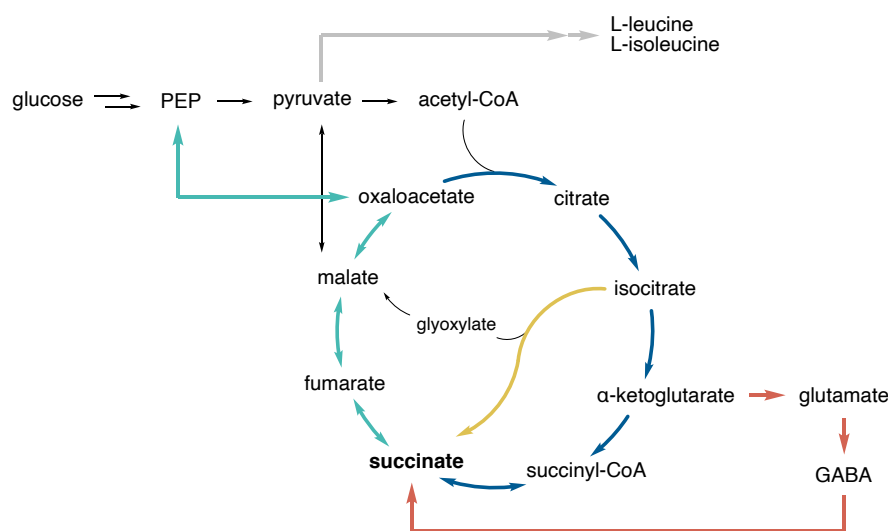
Obviously, exposure to minimal growth conditions cannot be fully prevented due to the need to compare different stages of enrichment in order to be able to select improved variants after all. Furthermore, while the degree of metabolic adaption upon elongated storage on minimal medium plates at 4°C remains uncertain, adaption at cultivation conditions was evident in many cases. For instance, unexpected growth of empty vector strains or growth of AVLDO WT harboring strains in medium containing L-isoleucine has been repeatedly encountered in the present and previous experiments (see, for instance, Figure 43-Figure 44).^{259,260} Particular caution must be also exercised when comparing these results to AVLDO wildtype (Figure 48), as this strain did not undergo any previous cultivation in minimal medium but was freshly prepared for this experiment. On this account, it would be particularly interesting for future investigations to deploy a wildtype reference strain, which was also taken through all cultivation and storage steps before rescreening.

Despite the knock-out of central succinate producing pathways in *E. coli* 3Δ*sucA*, several other pathways are eligible to circumvent the desired selection system. Under aerobic conditions, succinate is typically produced *via* the oxidative TCA cycle or *via* the glyoxylate shunt pathway, both of which are interrupted in *E. coli* 3Δ*sucA* due to deletion of the genes encoding α-KG dehydrogenase E1 subunit (*sucA*) and isocitrate lyase (*aceA*). Alternatively, under anaerobic conditions, succinate can be generated from the reductive branch of the TCA cycle or mixed acid fermentation (fueled by the glycolytic intermediate PEP) *via* the action of fumarate reductase. However, the latter typically produces only insignificant amounts of succinate.^{269,270} Alternatively, succinate can be formed *via* the GABA shunt from α-ketoglutarate. This pathway is, however, mostly used to fine-tune intracellular GABA levels (Scheme 50).²⁷¹

As outlined by Theodosiou *et al.*, the growth of *E. coli* 3Δ*sucA* could be also restored to some extent without external addition of substrate. That is, intracellularly produced proline could act as a substrate of proline-4-hydroxylase (P4H) and resulting endogenous hydroxylation was sufficing to complement the interrupted TCA cycle.⁵⁰ Also, in the present study, endogenous production of L-leucine or L-isoleucine can present a potential risk. Considering that biosynthesis of both amino acids branches off from pyruvate before entering the interrupted TCA cycle and the fact that up to 20 % of the relative protein mass

can be dedicated to amino acid biosynthesis in *E. coli* growing on minimal medium,²⁷² involvement of this pathway is quite likely (Scheme 50). This could particularly hamper the distinguishability and thus selection of variants preferring either L-Leu or L-Ile as substrate. Hence, this issue presents a substantial difference to the original study, where the vital proline-4-hydroxylase and proline biosynthesis both compete for α -ketoglutarate as central intermediate.⁵⁰

Scheme 50. Simplified overview on *E. coli* central carbon metabolism and pathways potentially impairing the applied selection procedure. Succinate is produced *via* the regular, oxidative, TCA cycle (blue) and the glyoxylate shunt pathway (yellow) but can also originate from the reductive branch of the TCA cycle or from mixed acid fermentation under anaerobic conditions (cyan). Alternatively, succinate can be formed *via* the GABA shunt (red). Additionally, pyruvate can be branched off to produce L-leucine or L-isoleucine intracellularly (gray). PEP: phosphoenolpyruvate, GABA: γ -aminobutyric acid. Figure adapted from Theodosiou *et al.*⁵⁰



The potential risk of undesired strain adaption and selecting false positive variants was also reflected in the results obtained from sequence analysis of the ten best performing clones during rescreening (Table 30).

Remarkably, amongst the analyzed variants, clone 4 exhibited a stop codon at position 206, and clone 24 showed a deletion involving target residues 78, 158 and 206, as was also previously found after MEGAWHOP reaction (see 5.3.2.3, Table 26). These truncations would lead to the loss of either one or two of the three highly conserved and functionally crucial iron-complexing residues (H151, D153, H237), which virtually disqualifies these variants from possessing any dioxygenase activity.²⁷³ While these findings are puzzling, important conclusions can be drawn from these observations regarding design of the enrichment

experiment, nonetheless. Both cases highlight the risk of selecting transformants without prior enrichment (e.g. clone 4 was picked from transformation plate) or after undesired adaption and thus in the absence of any selective growth advantage (e.g. clone 24 was picked after five rounds of cultivation). Consequently, this means that the selective procedure is not completely reliable yet and can be only operated under limited framework conditions, which are still not uncovered entirely.

Table 30. Sequence analysis of enriched AVLDO variants selected by rescreening (ten top performing variants). The stated origin denotes the stage of enrichment from which the respective clones were picked starting from the transformation of the library (trafo) followed by consecutive cultivation intervals of 24 h (t_0 - t_6).

Clone	Origin	Res. 67	Res. 70	Res. 78	Res. 158	Res. 206
WT	-	Y (TAT)	I (ATT)	F (TTT)	E (GAG)	Y (TAC)
4	trafo	T (ACT)	S (TCT)	L (CTT)	N (AAT)	* (TAG)
11	t_1	N (AAT)	A (GCA)	L (CTT)	T (ACT)	L (TTG)
12	t_1	N (AAT)	A (GCA)	L (CTT)	T (ACT)	L (TTG)
22	t_3	N (AAT)	A (GCA)	L (CTT)	T (ACT)	L (TTG)
23	t_3	N (AAT)	A (GCA)	L (CTT)	T (ACT)	L (TTG)
24	t_4	S (TCT)	V (GTA)		Deletion	
29	t_5	Y (TAT)	T (ACA)	L (CTT)	D (GAT)	C (TGT)
31	t_5	Y (TAT)	T (ACA)	L (CTT)	D (GAT)	C (TGT)
32	t_6	N (AAT)	A (GCA)	L (CTT)	T (ACT)	L (TTG)
34	t_6	Y (TAT)	T (ACA)	L (CTT)	D (GAT)	C (TGT)

Apart from these caveats, when analyzing the remaining variants, the genotype Y67N/I70A/F78L/E158T/Y206L was clearly predominant in 50 % of all selected clones (clones 11, 12, 22, 23 and 32) and in three cases the genotype Y67Y/I70T/F78L/E158D/Y206C was present (clones 29, 31 and 34). In view of the large total number of sequences theoretically comprised in the created library (17 280 sequences), the accumulation of only two predominant genotypes is yet striking. Quite interestingly, the amino acids present in the wildtype enzyme did not accumulate in any position and only in the case of Y67, the wildtype amino

acid was found at all (3 out of 10 variants). In the case of residue 78, leucine was present in all analyzed variants. While the relatively limited diversity observed after selection and rescreening indicates enrichment of distinct variants at least to some extent, the presence of highly truncated variants exposed limitations of the applied selection approach. In any case, the repetition of the rescreening procedure after retransformation seems indispensable to further clarify the actual contribution of undesired metabolic adaptation along the whole selection and rescreening workflow.

5.3.5 Heterologous production and purification of dioxxygenases

As previously published, a pET28a-based expression construct was used for production of AvLDO.²⁴¹ While utilization of the strong T7 promoter might be beneficial for the production of some enzymes, it can also lead to enhanced formation of inclusion bodies (*i.e.* insoluble protein) (see 3.3.2). As we were facing severe troubles in generating sufficient amounts of AvLDO in a soluble form, an extensive expression study was conducted, including a variation of cultivation and expression temperature, cell density at induction and inducer concentration. Additionally, co-expression of chaperons (GroEL/GroES) and addition of FeSO₄ were investigated.²⁶⁰ Yet, best results were obtained by implementing the initially reported conditions of cultivating *E. coli* BL21(DE3) [pET28a_AvLDO] at 37°C to an optical density of 0.5 (OD₆₀₀) before induction with 0.5 mM IPTG and further incubation at 20°C overnight for enzyme production.^{241,259} However, the successful production of AvLDO was not obvious and large batch-to-batch variations were observed. Preparation of cell lysates presented a further critical step, as many of the prepared batches were completely inactive after sonication,^{259,260} thus severely hampering comparability and reproducibility of AvLDO activity.

In view of the relatively high enzyme amounts used for *in-vitro* reactions (0.5-2 mg mL⁻¹),^{49,241} purification and concentration of AvLDO seems unavoidable. Table 31 summarizes different conditions used previously for the purification of AvLDO. It has to be noted that conditions described in Table 31 (entry 1) were omitted in this work, due to the sensitivity of AvLDO when directly prepared in the reaction buffer (50 mM sodium acetate, pH 4).

Table 31. Summarized conditions for purification of AVLDO via Ni-affinity chromatography as previously established or used in this work.

Entry	Lysis and purification buffer	Buffer exchange	Concentration procedure	Buffer exchanged to	Ref.
1	20 mM Tris-HCl (pH 7.4), 300 mM NaCl, 10 mM imidazole	PD-10 columns	Amicon Ultra Centrifugal Filters (10 kDa MWCO)	50 mM sodium acetate (pH 4), 10 % glycerol	²⁶⁰
2	20 mM Tris-HCl (pH 7.4), 300 mM NaCl, 10 mM imidazole	Combined approach in Amicon Ultra Centrifugal Filters (10 kDa MWCO)		10 mM HEPES (pH 7.0)	²⁴¹
3	20 mM Tris-HCl (pH 7.4), 300 mM NaCl, 10 mM imidazole	PD-10 columns	Amicon Ultra Centrifugal Filters (10 kDa MWCO)	10 mM HEPES (pH 7.0)	This work

In the first attempt, purification and concentration was done according to the reported procedure (Table 31, entry 2).²⁴¹ This procedure suggested to do both buffer exchange and concentration in centrifugal concentrators. This proved, however, highly problematic, as in all cases, protein started to precipitate. In the case of variant B and H, this happened at enzyme concentrations of 1 mg mL⁻¹ or lower to an extent, where the filters were completely blocked which made further concentration impossible. Further, precipitated or inactive protein distorted both analyses *via* BCA assay and SDS-PAGE (Figure S 37-Figure S 40, Appendix F) and led to an overestimation of applied enzyme amount. As can be seen from Figure S 40 (Appendix F), in the case of variant H, the desired protein could not be purified from the CFE. Following from the results of the purification and the observed precipitation, no enzyme activity could be detected in the following *in-vitro* assay, except for a slight conversion of L-Leu observed for the WT enzyme. This is, however, in contrast to previous work, where full conversion was reached after 18 h (0.5 mg mL⁻¹ AVLDO WT, L-Leu).²⁶⁰

Purification was attempted once again with a fresh batch of cells, as considerable batch-to-batch variations were already previously observed in terms of successful enzyme production, cell lysis and detectable enzyme activity.²⁶⁰ However, this time, purified fractions were re-buffered using PD-10 desalting columns before concentrating the combined elution fractions in centrifugal concentrators (Table

31, entry 3). Some precipitation still occurred during the ultrafiltration procedure, but certainly to a lesser extent than during the initial approach.

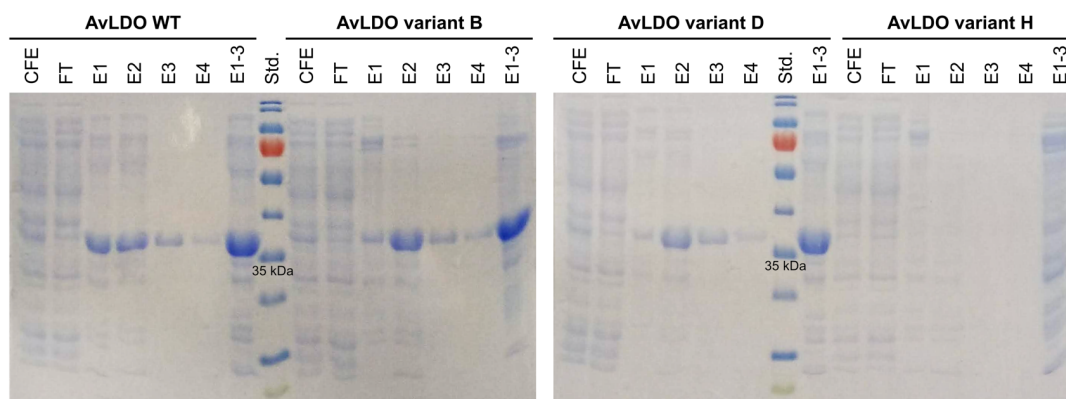


Figure 50. SDS-PAGE analysis of different fractions obtained from Ni-affinity purification of AvLDO WT, variant B, D and H (MW = 32.5 kDa). CFE: cell-free extract, FT: flow through, E: elution fraction. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

While AvLDO WT, variant B and variant D were again successfully purified, accumulation of variant H in the elution fractions was not observed (Figure 50). This repeated observation would suggest issues during binding of this variant to the Ni-sepharose matrix, which could be caused by undesired alterations of the His₆-tag sequence. Yet, the presence and integrity of the *N*-terminal His₆-tag sequence was confirmed after sequencing the respective plasmid once again. In view of the promising results obtained for this variant in several different and independent selection and rescreening experiments within this and previous work,²⁵⁹ further characterization and eventual confirmation of isoleucine-converting activity would be particularly valuable in order to validate the here applied selection procedure.

The successfully purified AvLDO WT and variants B and D were obtained in 2.7–4.1 mg mL⁻¹ after concentrating the combined elution fractions (E1-3).

In view of the many issues and limited reproducibility faced during heterologous production of AvLDO, utilization of an optimized expression system would be desirable. Currently, investigations towards the use of a pASK-based expression vector, which uses the tetracycline (*tet*) promoter/operator to control transcription, are underway. The *tet*-controlled gene expression combines potent induction after addition of anhydrotetracycline (5 000-fold), which is comparable to *lacUV5*, while exhibiting virtually no basal expression (one mRNA molecules estimated per three *E. coli* cells) in the repressed state.^{274,275}

5.3.6 Verification of the hydroxylating activity of AvLDO variants towards L-Ile

Besides a reliable selection and screening procedure, strategies to eventually confirm the desired enzyme activity are also essential. This is particularly important considering that the selection pressure of the applied assay is coupled to the conversion of a cosubstrate and not directly to product formation. This might favor so-called “uncoupling”, that is decarboxylation of the cosubstrate α -ketoglutarate without substrate hydroxylation. While quantification of dioxygenase activity *via* HPLC after suitable derivatization of the substrate and product amino acids is common practice, it is rather laborious and only conditionally suitable for broad-scale confirmation of library hits with yet unknown catalytic abilities. Apart from that, from experience, each single step along the way to detect dioxygenase activity can be a major hurdle: (1) Expression of sufficient enzyme amounts in an active and soluble form, (2) Cell lysis, (3) Purification and concentration, (4) Simple and reliable detection and analysis of reaction products. In this regard, also quite unreliable batch-to-batch variations can accompany each step.

5.3.6.1 *In-vitro* assay

In general, the preparation of cell-free extracts, followed by an *in-vitro* activity assay would be a straightforward approach to assess AvLDO activity. Yet, low soluble expression levels combined with relatively low specific activities render verification of enzyme activity *via* this route impractical. Impedingly, significant batch-to-batch fluctuations in terms of enzyme amount and activity were observed during enzyme production and cell lysis in this and previous studies.^{259,260} In view of the aforementioned issues and the relatively high amounts of AvLDO used in previous *in-vitro* characterization ($0.5\text{-}2\text{ mg mL}^{-1}$),²⁴¹ purification of AvLDO seemed unavoidable, thus severely limiting scalability. As already mentioned previously, AvLDO variants previously enriched and selected from the small library (I70X/F78X), namely variant B (I70T), D (I70L) and H (I70T/F78I), are yet to be characterized *in vitro*. Previous efforts to confirm the L-Ile-converting activity by using cell-free extracts of the respective variants remained without success.²⁵⁹ Hence, the variants of interest were heterologously produced and purified as described in Section 5.3.5. Several attempts were

necessary to obtain sufficient amounts of purified AvLDO variants due to their tendency to precipitate during the process. For characterization, different reactions containing either L-Leu or L-Ile as substrate and 0.5 or 1 mg mL⁻¹ enzyme were prepared similarly to previous reports.^{241,260} If applicable, an additional reaction was prepared containing the maximum possible volume of concentrated enzyme solution. Samples were taken after 0, 2, 16 and 24 h and analyzed *via* TLC.

Following from the observations made during rescreening of AvLDO wildtype and variants (see 5.3.4.1), activity towards conversion of L-leucine was expected for AvLDO WT and variants B and H, whereas variants B and H should be also able to accept L-isoleucine as substrate. Curiously, *E. coli* 3Δ*sucA* harboring variant D did neither grow on medium containing L-Leu nor L-Ile. While variant H was the most promising candidate according to the rescreening, purification of this variant remained unsuccessful at each attempt due to unknown reasons.

Certainly, only low conversion levels could be detected for AvLDO wildtype when L-leucine was used as substrate (Figure 51). Yet, the observed AvLDO activity was underperforming previous experiments, where full conversion was detected after 18 h (0.5 mg mL⁻¹ AvLDO WT, L-Leu).²⁶⁰

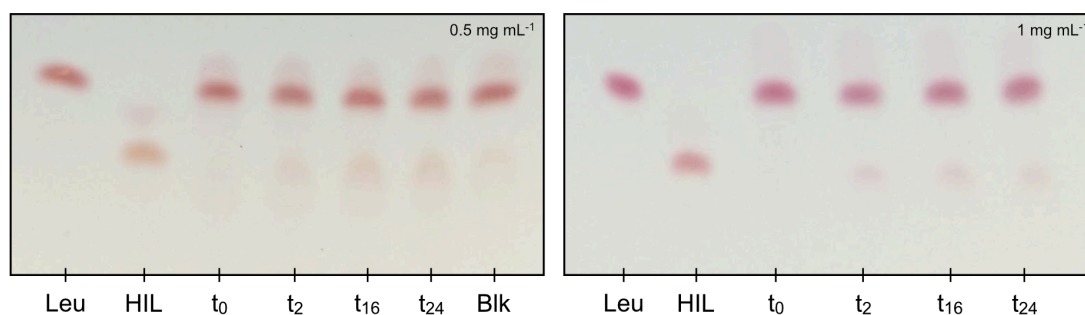


Figure 51. TLC analysis (acetone/1-butanol/AcOH/H₂O = 7:7:2:4, ninhydrin) of *in-vitro* activity assay samples containing 0.5 mg mL⁻¹ (left) or 1 mg mL⁻¹ (right) AvLDO WT and L-Leu as substrate taken after 0, 2, 16 and 24 h. Blank reactions (Blk) did not contain any enzyme. L-Leu and 4-Hydroxy-L-isoleucine (HIL) were used as reference.

5.3.6.2 Resting-cell assay

As reported by Zhang *et al.*, *E. coli* resting cells could be beneficial for the production of hydroxylated amino acids, as the quite unstable dioxygenase is maintained within the cellular environment. They also claimed that prior freezing of the cells would increase cell wall permeability for the hydroxylated product. In

that way, 22.4 g L⁻¹ 4-hydroxyisoleucine (4-HIL) were produced within 12 h by using resting cells of *E. coli* BL21(DE3) harboring a *Bt*IDO variant.²⁵⁵

Following the described procedures, preparation of *E. coli* BL21(DE3) resting cells containing either pET28a, pET28a_ *Bt*IDO, pET28a_ AVLDO or variants B, D, or H was envisaged. Recombinant strains were harvested either 6 h after induction or after incubation overnight. Cells were washed with buffer and frozen at -80°C. Initially, as suggested by Zhang *et al.*, 100 mM Tris-HCl (pH 7) was used to suspend the resting cells for the reaction.²⁵⁵ However, the Tris buffer interfered with detection of our target compounds *via* TLC. Hence, HEPES buffer was employed instead (50 or 100 mM, pH 7) and cells were suspended to 100 mg mL⁻¹ for the reaction. Similar to what was described before, the cell suspension was supplemented with L-Ile (10 mM), α -ketoglutarate (10 mM), FeSO₄ (5 mM) and ascorbate (10 mM). To ensure sufficient headspace for oxygen supply during the reaction, 2 mL of the mixture were incubated in 20 mL vials at 37°C. After 24, 48 and 72 h, samples were withdrawn, cells were separated, and the supernatant was analyzed *via* TLC. While some additional spots were observed after this timespan, no clear decrease of substrate could be detected in any case (Figure 52).

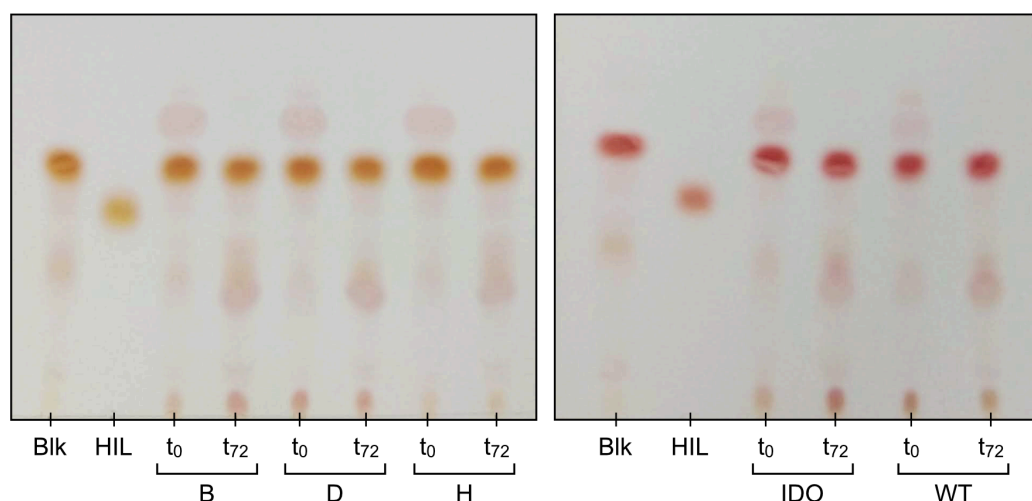


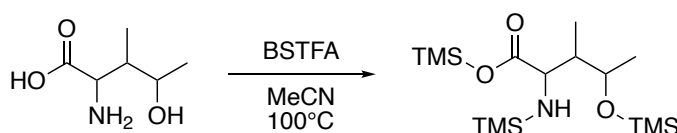
Figure 52. TLC analysis (acetone/1-butanol/AcOH/H₂O = 7:7:2:4, ninhydrin) of resting-cell assay samples taken after 0 and 72 h. Reactions contained 100 mg mL⁻¹ of *E. coli* BL21(DE3) resting cells harboring the respective pET28a plasmids of either *Bt*IDO, AVLDO WT or variants B, D and H in 50 mM HEPES buffer (pH 7). Blank reactions (Blk) contained *E. coli* BL21(DE3) [pET28a] cells. 4-Hydroxy-L-isoleucine (HIL) was used as reference.

Apparently, further optimization would be required to implement a resting cell assay, but in view of the limitations encountered for other methods described in this work, it could still present an appealing methodology to confirm AVLDO activity. Additionally, insufficient documentation and description of methods in the work of Zhang *et al.* raised several questions regarding the actual implementation.

5.3.6.3 GC-MS analysis

As product formation during the performed rescreening of AVLDO variants (see 5.3.4) was assumed based on qualitative analysis of culture supernatants *via* TLC, further characterization of the produced compound would be essential. While *Btl*DO is known to produce 4-hydroxyisoleucine from L-isoleucine,²⁴⁴ the product putatively formed upon conversion of isoleucine by AVLDO is entirely unknown as the wildtype enzyme did not accept L-Ile as substrate. However, in contrast to the C4-selectivity of *Btl*DO, δ -specific (C5) hydroxylation of L-leucine was described for AVLDO,²⁴¹ thus rendering formation of a different hydroxylation product compared to *Btl*DO quite likely. To unravel identity of the formed reaction products of selected AVLDO variants, implementation of GC-MS analysis was envisaged. Due to their low volatility, amino acids *per se* are unsuitable for gas chromatography applications, thus requiring prior derivatization. In this respect, silylation is amongst the most commonly applied derivatization techniques used for GC analysis. Specifically, amongst other reagents, *N,O*-Bis(trimethylsilyl) trifluoroacetamide (BSTFA) was described to react with amino acids in presence of acetonitrile at 100°C,²⁷⁶ thereby replacing all acidic protons (*i.e.* COOH, NH or OH) with a trimethylsilyl (TMS) group (Scheme 51).

Scheme 51. Illustrative derivatization of 4-hydroxyisoleucine with BSTFA. Possibly, the carboxylic acid, amine and hydroxy moiety can react with BSTFA to yield the corresponding triple silylated compound.



As BSTFA is sensitive towards moisture, lyophilization of aqueous samples was required prior to derivatization. Initially, the derivatization procedure was tested

with L-Ile and 4-HIL as reference compounds as well as with lyophilized M9 medium containing L-Ile, HIL and α -KG. While results obtained for pure compounds L-Ile and 4-HIL were unambiguous, derivatization of the complex minimal medium sample resulted in silylation of several compounds present in the mixture (Figure 53). Besides mono- and disilylated isoleucine ($t_R = 6.4$ min (TMS) and 8.1 min (2 TMS)), also silylated phosphate ($t_R = 7.9$ min) and various silylated sugars ($t_R > 12.5$ min) were detected. Quite surprisingly, not only silylated α -ketoglutarate ($t_R = 11.5$ min), but also succinate ($t_R = 8.3$ min) was detected in the medium, the source of which remains unclear. Further, the peak observed for silylated α -ketoglutarate mostly overlapped with another pronounced peak, whereas the peak formerly observed for pure 4-HIL ($t_R = 10.7$ min) was not detected in the mixture. This could indicate undesired reactivities of 4-HIL during derivatization of the complex sample, possibly yielding the unidentified product overlapping with α -KG mentioned before ($t_R = 11.5$ min) (Figure 53).

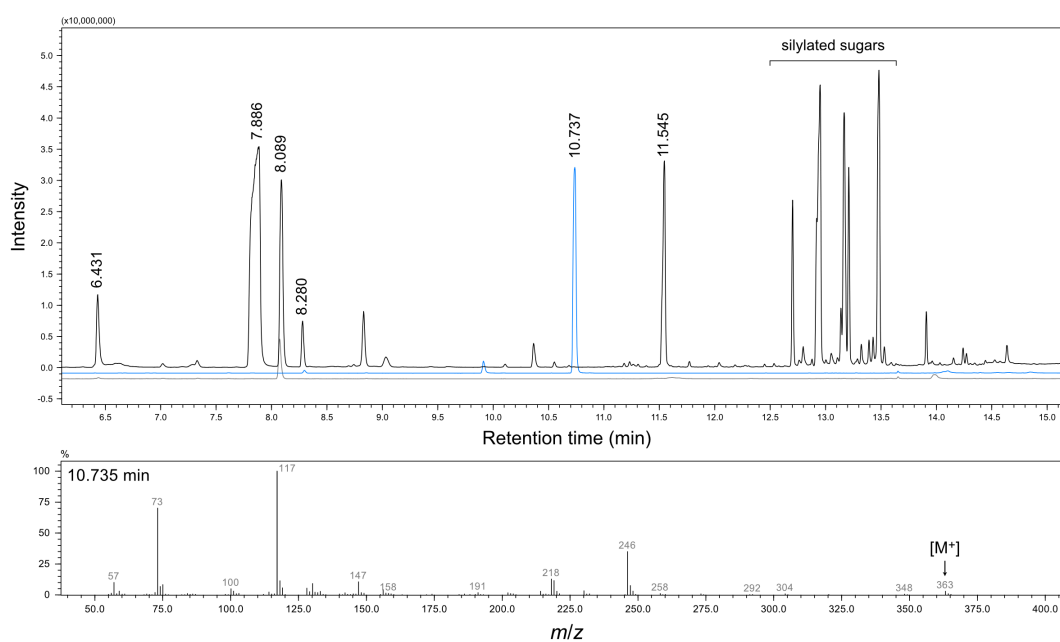


Figure 53. Overlaid TIC obtained from GC-MS analysis of TMS-derivatized M9- α KG/L-Ile/4-HIL (black), L-Ile (gray, $t_R = 8.1$ min) and 4-HIL (blue, $t_R = 10.7$ min) (top) and mass spectrum corresponding to the peak at 10.7 min (bottom). Peaks corresponding to mono ($t_R = 6.4$ min) and disilylated L-Ile ($t_R = 8.1$ min), silylated phosphate ($t_R = 7.9$ min), succinate ($t_R = 8.3$ min) and α -ketoglutarate ($t_R = 11.5$ min) could be assigned based on comparison to the database. At $t_R > 12.5$ min, various silylated sugars were detected.

According to this procedure, culture supernatants obtained from rescreening of AVLDO library variants 1-35 (see 5.3.4.2) were analyzed. Results obtained for all 35 variants were rather comparable, and despite careful examination of the obtained chromatograms, no peak could be clearly assigned to product formation when compared to the sample containing only medium with L-Ile, 4-HIL and α -KG. Particularly, when comparing results obtained for variant 11 (selected as a representative example) with variant 4 and 24 (both truncated, see Table 30), all unidentified and newly emerged peaks were present in each sample (Figure 54).

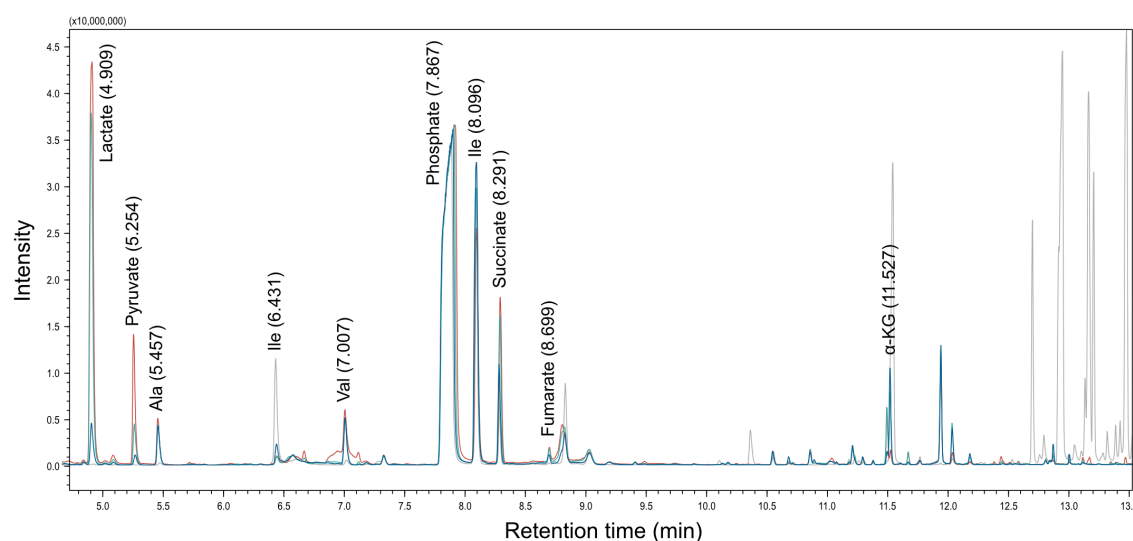


Figure 54. Overlaid TIC obtained from GC-MS analysis of TMS-derivatized M9- α KG/L-Ile/4-HIL (gray) or culture supernatants of variant 11 (blue), variant 4 (cyan) and variant 24 (red). Peaks corresponding to mono ($t_R = 6.4$ min) and disilylated L-Ile ($t_R = 8.1$ min), silylated lactate ($t_R = 4.9$ min), pyruvate ($t_R = 5.3$ min), alanine ($t_R = 5.5$ min), valine ($t_R = 7.0$ min), phosphate ($t_R = 7.9$ min), succinate ($t_R = 8.3$ min), fumarate ($t_R = 8.7$ min) and α -ketoglutarate ($t_R = 11.5$ min) could be assigned based on comparison to the database. At $t_R > 12.5$ min, diverse silylated sugars were detected in the medium sample, but were mostly already metabolized in the analyzed cultures.

Most interestingly, lactate ($t_R = 4.9$ min) was detected in samples of all 35 variants, thus strongly indicating activation of anaerobic mixed acid fermentation during cultivation.²⁷⁷ As already discussed in Section 5.3.4.3, this could potentially result in bypassing of the selection system, although amounts of succinate formed *via* this metabolic route are typically low.^{269,270} Quite unexpectedly, also pyruvate ($t_R = 5.3$ min) and fumarate ($t_R = 8.7$ min) were detected in samples of all screened variants. As both compounds are central intermediates of the TCA cycle, their accumulation is unlikely under regular metabolic conditions,²⁷⁸ thus further underlining the metabolic disbalance caused

by the interrupted TCA cycle. It has to be noted that the activation of anaerobic metabolism, as evidenced by the presence of lactate, can also lead to the formation of succinate *via* the reductive branch of the TCA cycle (*via* PEP, oxaloacetate, malate and fumarate; Scheme 50). The presence of detectable amounts of alanine ($t_R = 5.5$ min) and valine ($t_R = 7.0$ min) further confirm the possibility of endogenous amino acid synthesis under the applied cultivation conditions. Particularly, the presence of valine is concerning, as several biosynthetic steps leading either to valine, leucine or isoleucine are shared.²⁷⁹ The ability of cells to provide themselves with L-leucine would thus largely prevent enrichment towards non-native substrates.

While the analysis revealed a quite detailed insight into the complex minimal medium mixture, several issues remain unresolved, thus requiring supplementary analyses of available reference compounds to further unravel the chemistry and identity of compounds involved. It also became obvious that different reaction matrices, for instance, different conditions used for *in-vivo* rescreening, *in-vitro* activity assays and resting cell assays, led to differing reactivity during derivatization and prevented reliable assignment of compounds and thus comparison of results.

5.4 Summary

In this work, we aimed at extending the limited substrate scope of the Fe(II)/ α KG dependent L-leucine dioxygenase from *Anabaena variabilis* (AvLDO) towards the related substrate L-isoleucine. Even though the crystal structure of AvLDO could be resolved recently, only little is known about general determinants of substrate scope throughout this enzyme family. That is, while binding of the substrate amino acid group and coordination of the cosubstrate α -ketoglutarate are specifically pre-determined by the respective coordination sites, the position of substrate hydroxylation to occur is merely a result of conformation taken by the substrate sidechain, which in turn is determined by the overall architecture of the active site cavity.

Initially, the structure of AvLDO was analyzed separately and in the context of the created 3DM database and target residues for site-saturation mutagenesis were selected thereof. For each of the selected residues Y67, I70, F78, E158 and Y206, restricted degenerate codons were selected with respect to the spatial and chemical nature of the encoded amino acids. A megaprimer library was generated by OE-PCR followed by the integration into suitable vector backbones. In total, the created library theoretically comprised 17 280 individual variants on gene-level encoding 6 300 different amino acid variants. The triple knock-out strain *E. coli* Δ sucA Δ aceA Δ putA (3Δ sucA) was chosen as selection and screening host strain for its viability, which depends on the presence of active α -KG dependent dioxygenases to bypass the interrupted TCA cycle *via* formation of succinate as coproduct (Figure 55).

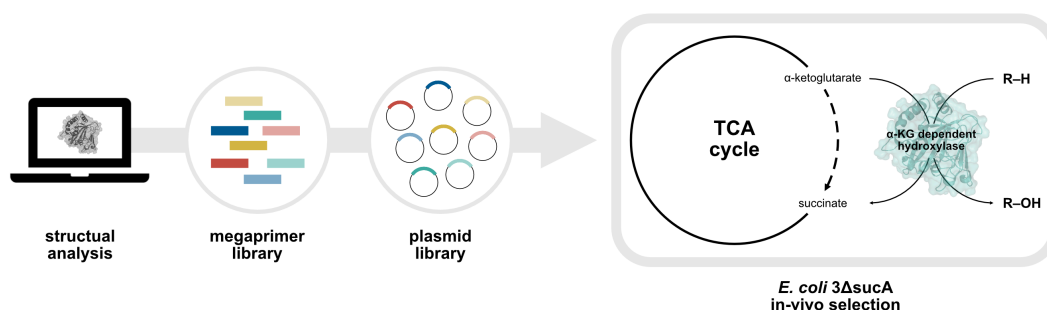


Figure 55. Summarized workflow of AvLDO library preparation and selection. Structural analysis guided the selection of residues for site-saturation mutagenesis (SSM). Based on that, a megaprimer library was created, which was then integrated into a vector backbone. The resulting plasmid library was used to transform *E. coli* 3Δ sucA for *in-vivo* selection of library variants based on their ability to utilize specific substrates supplemented in the growth medium.

By cultivating the resulting recombinant selection strain *E. coli* 3Δ*sucA* in minimal medium containing L-Ile as substrate, those variants unable to utilize isoleucine cannot provide the cell with succinate necessary for survival. In contrast, recombinant strains harboring isoleucine converting AvLDO variants can grow, and the diversity should be further reduced by several cultivation intervals leading to enrichment of superior variants. This selection strain was originally designed and optimized for a proline-converting dioxygenase and might only be conditionally suitable for the selection or differentiation of dioxygenase variants converting either L-Leu or L-Ile. Indeed, despite thorough preliminary examination of growth conditions suitable for our selection purpose, several counterintuitive results were obtained until last. Admittedly, our quite ambitious intention of expanding the scope of AvLDO towards natively non-accepted substrates might have overcharged the chosen selection system. Particularly, the tendency of Fe(II)/αKG dependent dioxygenases to promote uncoupled turnover of α-ketoglutarate in the presence of poorly fitting substrates might have severely hampered selection of truly active variants. Detection of variants possessing a significant deletion of over 400 bp after several intervals of enrichment corroborates this reasoning. Also, the generally vanishing substrate-dependent growth after elongated cultivation in minimal medium presents a considerable risk and the extent to which the selection system is possibly bypassed by alternative sources of succinate or endogenous substrate hydroxylation through intracellular amino acid biosynthesis still remains unclear. Nonetheless, the clear predominance of only two genotypes (Y67N/I70A/F78L/E158T/Y206L and Y67Y/I70T/F78L/E158D/Y206C) detected after sequence analysis of enriched AvLDO variants indicated enrichment and thus reduction of diversity at least to some extent. Reevaluation and further characterization of these variants is pending.

Besides the limitations faced during *in-vivo* selection of library variants, confirmation of AvLDO activity presented a major hurdle. As experienced during the present and previous studies, heterologous production, as well as cell lysis or purification, can render AvLDO completely inactive and inconsistent results were often obtained for different batches. This can be particularly troublesome for proof of non-native reactivities as in the present study. When selecting variants from a library, a quick and reliable verification of the desired reactivity is

inevitable for review of selected clones and to avoid the most laborious steps of enzyme purification and characterization of putatively false-positive hits. In this regard, a whole-cell approach, either *in vivo* or by using resting cells, can be beneficial given the limited stability of AvLDO. Yet, *in-vivo* rescreening performed in this work posed the risk of undesired metabolic adaption of the recombinant *E. coli* $3\Delta sucA$ strains, thus leading to inconclusive results, and a resting cell assay performed according to previous literature reports remained unsuccessful. Furthermore, while TLC analysis presents an easy-to-implement technique to monitor reaction progress, assigning compounds formed during *in-vivo* conversions can be particularly challenging. In this regard, we aimed at implementing GC-MS analysis as complementary method but, similarly to TLC, results require careful evaluation due to the plethora of formed metabolites detectable in *in-vivo* samples. Also, the completely unknown identity of the product putatively formed by AvLDO upon conversion of L-isoleucine complicates unambiguous detection by GC-MS within the complex samples. Detection of several metabolites, including lactate, pyruvate and fumarate, strongly indicated activation of anaerobic metabolism under the applied cultivation conditions during rescreening. This together with the presence of amino acids (alanine and valine) in the analyzed culture supernatants, could both be highly problematic for the here applied selection system, as in either case, alternative routes to evade the selective metabolic step might be established.

5.5 Outlook

To this day, several questions regarding the reliability of the implemented selection procedure still remain unanswered. In view of the advancing adaption to minimal growth conditions after three 24 h cultivation intervals at latest, execution of shorter cultivation periods (e.g. only one to two cycles) followed by isolation of mixed plasmids and retransformation of the selection strain before continuation of enrichment could present a reasonable adaption to current procedures.

Besides the potential risk of selecting false-positive variants through undesired alternative metabolic activities, reliable and reproducible detection of AvLDO activity still represents one of the major challenges. The issue of unambiguously detecting AvLDO activity could be partially tackled by introducing alternative expression systems. The currently used pET28a scaffold relies on the T7 promoter which, besides the generally low amounts of heterologously produced AvLDO, mostly results in a large portion of insoluble or inactive enzyme. Currently, investigations towards implementation of a pASK-based vector which controls transcription *via* the tetracycline (*tet*) promoter/operator are underway. Ideally, optimization of the expression system would result in increased levels of AvLDO and thus detectable activity in cell lysates without prior purification.

Pursuing investigations in our lab further aim at deep mutational scanning of dioxygenases combined with next generation sequencing to potentially unravel global hotspots of enzyme activity and selectivity.

6 Conclusion

In this work, we focused on unraveling correlations of enzyme sequence, structure and function. Manifold aspects of these structure-function relationships were addressed as represented by characterization of the *BsPAD* substrate spectrum in different solvent systems, the rational design of AMDase to study alternative binding modes and semi-rational engineering of *AvLDO* to broaden the substrate scope.

While the substrate scope of enzymes can often be explained in parts, the challenge remains to fully rationalize it, thus allowing predictability. For instance, within the pool of accepted substrates, the impact of single substituents on the reaction rate is often puzzling, as in the case of *BsPAD*, where both electron donating and withdrawing substituents led to significant rate enhancements. While substituent size was generally decisive for acceptance, it seemed that an optimized stabilization resulting from different substituents could enhance the reaction rate even further.

Likewise, substrate acceptance and preference of AMDase were not obvious in all cases. Manifold computational analysis helped to uncover important features of accepted over non-converted substrates such as the capacity of sequestering the active site from water, which was closely connected to the ability of stable substrate accommodation. Substrates with higher flexibility and mobility failed to do so, which left them unconverted by AMDase. We also found clear indication for ground-state destabilization as driver for AMDase activity, as evidenced by the establishment of a hydrophobic cage around the carboxylate to-be-cleaved. The study highlighted the potential of targeted hydrophobic pocket engineering for AMDase once more.

With the computational methodology established for AMDase, we also hope to get a more detailed insight into the conversion of small non-aromatic substrates, as experiments performed to date did not provide clear evidence for the contribution of alternative binding modes. Providing suitable model substrates to prove the hypothesized alternative binding modes experimentally posed a major challenge.

In case of recently discovered enzymes, such as AvLDO, the determinants of substrate scope and selectivity often remain scarce to date, thus requiring random or semi-rational design strategies to broaden their scope. However, the success of these techniques largely relies on the availability of an efficient and reliable selection and screening system. We aimed at extending the substrate scope of AvLDO towards L-isoleucine and selecting active variants by growth complementation of an *E. coli* strain unable to synthesize succinate. While *in-vivo* selection systems relying on growth complementation by active variants offer great potential for the quick sorting of libraries, careful evaluation is necessary due to the considerable metabolic flexibility of the selection host and potential side-reactivities of the enzyme of interest. Particularly, enrichment of uncoupling variants of AvLDO, which provide the vital succinate uncoupled from substrate conversion, posed a severe risk during library selection. Indeed, efforts to eventually confirm the desired reactivity towards L-isoleucine of enriched variants remained without success.

Next to enzymatic studies, reaction engineering complemented the work on BsPAD. The limited solubility of PAD substrates mostly prevented a broader applicability so far. In search of more sustainable alternative reaction systems, deep eutectic solvents were the preferred choice. BsPAD tolerated these non-conventional solvents, but addition of water was necessary to effectively release the formed CO₂, thereby preventing the reverse carboxylation reaction. The established setup allowed to convert 30-fold higher concentrations of substrate compared to aqueous systems and further provided suitable conditions for implementation of a chemo-enzymatic cascade. Different substrate preferences depending on the solvent system highlighted the potential of changing the substrate scope by solvent engineering.

Overall, the necessity of combining various disciplines to gain a thorough understanding of a given enzyme was highlighted in this work. Structure-function relationships of enzymes are often non-obvious and need consideration from various perspectives to shed light on the molecular level determinants of enzyme activity and selectivity. This is inevitable to pave the way for a broader applicability of biocatalysis in sustainable processes.

7 Experimental section

7.1 Microbiological methods

7.1.1 Strains

The strains listed in Table 32 were used within this work, depending on the respective application.

Table 32. List of *E. coli* strains used in this work.

Strain	Genotype	Purpose	Source
<i>E. coli</i> TOP10	F ⁻ <i>mcrA</i> Δ(<i>mrr-hsdRMS-mcrBC</i>) Φ80(<i>lacZ</i>)ΔM15 Δ <i>lacX74</i> <i>recA1</i> <i>araD139</i> Δ(<i>ara-leu</i>)7697 <i>galU</i> <i>galK</i> <i>rpsL</i> (Str ^R) <i>endA1</i> <i>nupG</i>	Cloning experiments	Thermo Scientific
<i>E. coli</i> Mach1	F ⁻ Φ80(<i>lacZ</i>)ΔM15 Δ <i>lacX74</i> <i>hsdR</i> (r _K ⁻ m _K ⁺) Δ <i>recA1398</i> <i>endA1</i> <i>tonA</i>	Cloning experiments	Thermo Scientific
<i>E. coli</i> BL21(DE3)	F ⁻ <i>ompT</i> <i>hsdS_B</i> (r _B ⁻ m _B ⁻) <i>gal</i> <i>dcm</i> (DE3)	Production of recombinant proteins	Thermo Scientific
<i>E. coli</i> BL21(DE3) 3Δ <i>sucA</i>	Δ <i>sucA</i> Δ <i>aceA</i> Δ <i>putA</i>	Selection experiments	Helmholtz center for environmental research (UFZ, Leipzig) ⁵⁰

7.1.2 Over-night cultures (ONCs)

Over-night cultures (ONCs) were routinely prepared for the purpose of plasmid isolation, inoculation of main cultures or glycerol stock preparation. For this, LB medium (5-10 mL) was mixed with the respective antibiotic (Table 33), if applicable, and inoculated with a single colony taken from agar plates, or with glycerol stock (10 μL). The cultures were typically incubated at 37°C and 130 rpm overnight (approx. 16-18 h).

7.1.3 Glycerol stocks

Cultures were maintained in 30 % (v/v) glycerol stocks at -20°C and -80°C. For this purpose, 60 % (v/v) sterile glycerol was thoroughly mixed with the respective overnight culture in a 1:1 ratio in cryogenic tubes. Afterwards, glycerol stocks were frozen and replaced after repeated use.

7.2 Molecular biological methods

7.2.1 List of plasmids

Plasmids listed in Table 33 used within this work either directly or as template for creating further variants.

Table 33. List of vectors and recombinant plasmid constructs used in this work.

Vector	Insert	Resistance marker	6xHis-Tag	Cloning site	Source
pET28a(+)	BsPAD	Kan	N-term	<i>NdeI/XhoI</i>	RUB
pET28a(+)	BsPAD I85A	Kan	N-term	<i>NdeI/XhoI</i>	RUB
pET28a(+)	AMDase	Kan	C-term	<i>NcoI/HindIII</i> ^[a]	RUB
pET28a(+)	AMDase IPLL	Kan	C-term	<i>NcoI/HindIII</i> ^[a]	RUB
pET28a(+)	AMDase CLGIPL	Kan	C-term	<i>NcoI/HindIII</i> ^[a]	RUB
pET28a(+)	AvLDO	Kan	N-term	<i>NdeI/XhoI</i>	RUB
pET28a(+)	BtIDO	Kan	C-term	<i>NdeI/XhoI</i>	²⁵⁹
pET28a(+)	-	Kan	N/C-term	-	In-house collection
pJET2.1	-	Amp	-	blunt ^[b]	Thermo Scientific

[a] The AMDase gene was cut from a previous pBAD-construct together with a C-terminal 6xHis-tag via *NcoI/HindIII* and cloned into pET28a(+) via the same restriction sites. [b] pJET2.1/blunt was purchased in a linearized form as part of the CloneJET PCR Cloning Kit (Thermo Scientific)

Recombinant strains containing the plasmids listed in Table 33 were cultivated in antibiotic containing medium for selection. Therefore, 1000x stocks of the respective antibiotic were prepared in ddH₂O and stored at -20°C (Appendix B,

Table 47). The final antibiotic concentration was $100 \mu\text{g mL}^{-1}$ for ampicillin and $40 \mu\text{g mL}^{-1}$ for kanamycin.

7.2.2 Preparation of competent *E. coli* cells

7.2.2.1 Preparation of electro-competent *E. coli* cells

For the preparation of electro-competent *E. coli* cells, an overnight culture of the respective strain was prepared in 2xTY medium (50 mL) and incubated at 37°C and 200 rpm overnight. In case *E. coli* TOP10 was used, all media were supplemented with streptomycin ($25 \mu\text{g mL}^{-1}$). Sterile ddH₂O and 10 % (v/v) glycerol were pre-cooled at 4°C . At least two main cultures (500 mL 2xTY medium each) were inoculated with ONC (5 mL each) and incubated at 37°C and 200 rpm, until OD₆₀₀ reached 0.5-0.6. The cultures were then incubated on ice for 1 h before cells were harvested ($3\ 000 \times g$, 10 min, 4°C). The obtained cell pellets were each resuspended in a small volume of sterile ddH₂O before two of them were combined and sterile ddH₂O was added (to approx. 100-150 mL). Samples were centrifuged again ($4\ 000 \times g$, 10 min, 4°C) and pellets were suspended in a small volume of sterile ddH₂O before 10 % (v/v) glycerol was added (to approx. 500 mL). After centrifugation ($4\ 500 \times g$, 20 min, 4°C), cell pellets were resuspended in a small volume of sterile 10 % (v/v) glycerol. Glycerol (10 % v/v) was added to approx. 500 mL and after a final centrifugation step ($5\ 000 \times g$, 15 min, 4°C), the obtained cell pellets were resuspended in sterile 10 % (v/v) glycerol (2 ml per liter culture) and 80 μL aliquots were prepared in 1.5 mL tubes and immediately placed on ice. Aliquots were stored at -80°C until use.

7.2.2.2 Preparation of chemo-competent *E. coli* cells

For the preparation of chemo-competent *E. coli* cells, LB-medium (100 mL) was inoculated with an ONC of the desired *E. coli* strain. In case *E. coli* TOP10 was used, all media were supplemented with streptomycin ($25 \mu\text{g mL}^{-1}$). All solutions necessary were pre-cooled at 4°C (Appendix B, Table 51). Cultures were incubated at 37°C and 200 rpm until an OD₆₀₀ of 0.4-0.5 was reached. Cells were harvested ($4\ 000 \times g$, 15 min, 4°C). The obtained cell pellet was resuspended in TfB1 buffer (30 mL) and MgCl₂ solution (1 M, 3.2 mL) and incubated on ice for

15 min. Cells were centrifuged (4 000 x g, 10 min, 4°C) and the supernatant discarded, before the pellet was resuspended in Tfb2 buffer (4 mL). The suspension was incubated on ice for 15 min before 50 µL aliquots were prepared in 1.5 mL tubes. Aliquots were shock-frozen in liquid nitrogen and stored at -80°C until use.

7.2.3 Transformation of competent *E. coli* cells

For transformation, an aliquot of competent cells was thawed on ice and typically 0.5-1 µL of plasmid DNA or 5-10 µL of a whole-plasmid-amplification reaction (e.g. QuikChange, MEGAWHOP) were added to the cell suspension. After the transformation procedure, cell suspensions were regenerated at 37°C and 550 rpm for at least 1 h (Amp resistance), 1.5-2 h (Kan resistance) or 2 h (*E. coli* knock-out strains), respectively. Suitable aliquots were spread on antibiotic-containing LB agar plates as required and incubated at 37°C until colony formation was observed.

*7.2.3.1 Transformation of electro-competent *E. coli* cells*

An aliquot of electro-competent cells (80 µL) was thawed on ice and DNA was added to the cell suspension. If required, the DNA preparation was desalted before transformation for 20-30 min (MF-Millipore Membrane Filter, 0.025 µm pore size, Merck Millipore). The mixture was transferred to electroporation cuvettes (2 mm gap) and an electric pulse was applied (2.5 kV, 5-6 msec) (MicroPulser Electroporator, Bio-Rad). LB-SOC medium (700 µL) was immediately added, and the mixture was regenerated. Suitable aliquots were spread on respective LB-agar plates.

*7.2.3.2 Transformation of chemo-competent *E. coli* cells*

An aliquot of chemo-competent cells (50 µL) was thawed on ice and DNA was added to the cell suspension. The mixture was incubated on ice for 30 min, before a heat-shock at 42°C for 42 s was applied. LB-SOC medium (400 µL) was added, and the mixture was regenerated. Suitable aliquots were spread on respective LB-agar plates.

7.2.3.3 Determination of transformation efficiency

To assess the transformation efficiency, competent *E. coli* cells were transformed with a defined amount (e.g. 10 ng) of a suitable plasmid according to the transformation procedures described in Section 7.2.3. Various dilutions of the transformation mixture were plated on suitable LB agar plates and the colony forming units (cfu) were determined after incubation. The transformation efficiency is expressed in $\text{cfu } \mu\text{g}^{-1}$ and was calculated according to Equation (7).

$$\text{Transformation efficiency } [\text{cfu } \mu\text{g}^{-1}] = \frac{\text{cfu}_{\text{plate}} \times V_{\text{tot}} \times f}{V_{\text{plate}} \times m_{\text{DNA}}} \quad (7)$$

$\text{cfu}_{\text{plate}}$ Colony forming units on plate []

V_{tot} Total volume of transformation mixture [μL]

f Dilution factor of plated dilution []

V_{plate} Plated volume of dilution [μL]

m_{DNA} Amount of plasmid DNA [μg]

7.2.4 Isolation of plasmid DNA

Cell material for plasmid isolation was either obtained from liquid culture (5-10 mL ONC) by centrifugation or from densely grown LB-agar plates. Commercially available GeneJET Plasmid Miniprep Kit (Thermo Scientific) or Wizard Plus SV Minipreps DNA Purification System (Promega) was used to isolate the plasmid DNA according to the manufacturer's instructions. Plasmid DNA was typically eluted in pure ddH₂O.

7.2.5 Gel extraction and PCR clean-up

For isolation of DNA from preparative agarose gels or for clean-up of PCR products, commercially available GeneJET Gel Extraction Kit, GeneJET PCR Purification Kit (Thermo Scientific) or Wizard SV Gel and PCR Clean-Up System (Promega) were used according to the manufacturer's instructions. DNA was typically eluted in pure ddH₂O.

7.2.6 Isolation of genomic DNA

To isolate genomic DNA (gDNA) isolation, overnight cultures of the respective *E. coli* strain were prepared according to Section 7.1.2 and grown to saturation. Cultures (3 mL) were harvested by centrifugation (5 000 rpm, 5 min, 4°C). The obtained cell pellet was resuspended in TNE buffer (1 mL) and the supernatant was discarded after centrifugation (5 000 rpm, 5 min, 4°C) before the cells were resuspended again in TNEX buffer (540 µL). Lysozyme (60 µL, 5 mg mL⁻¹ in ddH₂O) was added and the cell suspension incubated at 37°C for 20 min. Subsequently, Proteinase K (30 µL, 20 mg mL⁻¹, Thermo Scientific) was added and the mixture incubated at 65°C for 60 min. The cleared solution was extracted with PCI (600 µL; phenol/chloroform/isoamyl alcohol = 25:24:1) and phase separation was achieved by centrifugation (13 000 rpm, 10 min, 4°C). The upper, aqueous phase (400 µL) was transferred to a fresh tube before 5 M NaCl (40 µL) and ice-cold absolute EtOH (1 mL) were added. After careful mixing, formation of a cloudy gDNA precipitate was observed, which was pelleted by centrifugation (13 000 rpm, 20 min, 4°C). Optionally, precipitation of gDNA could be extended to overnight at -20°C. After centrifugation, the supernatant was removed and the DNA pellet was washed by addition of 70 % EtOH (700 µL) without mixing and subsequent centrifugation (13 000 rpm, 5 min, 4°C). The washing step was repeated, and the resulting DNA pellet was dried at 60°C. Genomic DNA was dissolved in ddH₂O (100 µL), RNase A (2 µL, 10 mg mL⁻¹, Thermo Scientific) was added and the mixture was incubated for 30 min at 37°C. The resulting DNA preparation was analyzed by agarose gel electrophoresis (3 µL, see 7.2.10) to estimate the concentration and stored at 4°C (short-term) or at -20°C.

7.2.7 Colony PCR (cPCR)

For colony PCR, colonies were picked and restreaked, before the cell material was resuspended in ddH₂O (30 µL). The mixture was heated to 95°C for 10 min followed by centrifugation, which provided the template for cPCR in the supernatant. For amplification, either a Red HS Taq Master Mix (Biozym) or a OneTaq Quick-Load 2X Master Mix (New England Biolabs) was used. Reaction conditions for colony PCR are summarized in Table 34. The obtained PCR products were analyzed by agarose gel electrophoresis (10 µL, see 7.2.10).

Table 34. Sample composition (20 μL) and temperature program for colony PCR using a ready-to-use 2x PCR Master Mix. Annealing temperature and elongation time were adapted to the respective reaction setup.

10 μL	PCR Master Mix 2x	95°C	30/60 sec	Initial denaturation
0.4 μL	Primer fwd [10 μM]	95°C	15 sec	Denaturation
0.4 μL	Primer rev [10 μM]	var.	15 sec	
5 μL	Colony supernatant	72°C	var.	Elongation
4.2 μL	ddH ₂ O	72°C	5 min	
		72°C	5 min	Final elongation
		4°C	∞	

7.2.8 Determination of DNA concentration

The concentration of isolated plasmid DNA or purified PCR products was determined spectrophotometrically with a NanoDrop 2000c spectrophotometer (Peqlab) at 260 nm. Additionally, the absorption at 280 and 230 nm was measured in parallel to assess purity of the DNA sample.

7.2.9 DNA sequencing and sequence analysis

Samples for Sanger sequencing were prepared according to the instructions provided by the company and sent either to GATC Biotech AG or Microsynth Austria GmbH. Respective sequencing primers were either available from the in-house primer collection (Appendix C, Table 55) or provided by the sequencing company. DNA sequences were aligned and analyzed with the online-tool Benchling.²⁸⁰

7.2.10 Agarose gel electrophoresis

Agarose gel electrophoresis was performed to analyze the size of DNA fragments. Depending on the expected fragment size, the prepared gels contained 1-2 % (w/v) agarose (LabQ Agarose LE, LabConsulting) in 1x TAE buffer. The mixture was heated until the agarose was homogeneously dissolved and cooled again to approx. 60°C before DNA-intercalating dye was added (20 000x Green Safe, LabConsulting). The mixture was poured into casts of desired size and suitable combs were added. Gels were left until completely set before used in electrophoresis. Samples meant for agarose gel electrophoresis were combined with 6x DNA Gel Loading Dye (Thermo Scientific) before loading

and GeneRuler DNA Ladder Mix (Thermo Scientific) was used as standard (5 μ L). Analytical gels were routinely run at 120 V and 400 mA for 30-40 min and preparative gels at 100 V and 400 mA for 50-60 min.

7.2.11 Mutagenesis

7.2.11.1 QuikChange

For introduction of one amino acid exchange at a time, an adapted QuikChange protocol was used. Desired base-pair exchanges were encoded by a pair of primers, which were approx. 20-30 bp long, contained the mutation in their 5' complementary region and both had 3' overhangs, which were only complementary to the backbone.

In most cases, a temperature gradient during annealing was chosen according to the calculated melting temperature (T_m) of the respective primers and elongation time was adapted to the total plasmid size used.

A master mix containing all reaction components was prepared and divided into 20 μ L aliquots (Table 35). The protocol shown in Table 35 was used for mutagenesis of *BsPAD* and *AMDase*. Respective primers are listed in Appendix C, Table 56 and Table 57) and the corresponding detailed temperature program conditions in Table 35.

Table 35. Master mix (8 x 20 μ L) and temperature program for QuikChange reaction.

130.4 μ L	ddH ₂ O	95°C	1 min	Initial denaturation
16 μ L	10x <i>Pfu</i> buffer	95°C	30 sec	Denaturation
0.8 μ L	DMSO	Gradient ^[a]	30 sec	
2 μ L	Primer fwd [10 μ M]	72°C	6 min	Elongation
2 μ L	Primer rev [10 μ M]	72°C	10 min	
4 μ L	Template plasmid	4°C	∞	Final elongation
4 μ L	dNTPs [10 mM]	[a] Gradient for mutagenesis of <i>BsPAD</i> : 57, 57.7, 58.6, 59.8, 61.1., 62.5, 63.8, 65.1 °C. Gradient for mutagenesis of <i>AMDase</i> : 63, 63.6, 65.3, 67.7, 69.3, 70°C		
0.8 μ L	<i>Pfu</i> polymerase			

After the reaction, 3 μ L were taken and analyzed on an agarose gel (see 7.2.10). Meanwhile, 0.5 μ L *DpnI* were added to each sample and incubated at 37°C for 2 h, followed by heat-inactivation at 80°C for 20 min. In case of successful

amplification, respective samples were pooled and used for transformation of chemo-competent *E. coli* TOP10 cells (5 μ L, see 7.2.3). Colonies were picked and the plasmids isolated thereof were sent for Sanger sequencing (see 7.2.9). After sequencing confirmed incorporation of the desired mutation, competent cells of *E. coli* BL21(DE3) were transformed with the respective plasmids and transformants were used for heterologous enzyme production (see 7.3.1).

7.2.11.2 Overlap extension PCR (OE-PCR)

For generation of mutant libraries, mutagenic megaprimers containing several mutation sites were generated by overlap extension PCR (OE-PCR). For each mutation site, the selected restricted degenerate codons were encoded by a set of primers, which were approx. 30-60 bp long and were designed as described in Section 7.2.11.1. The complementary region of each primer set provided compatible overlaps between the generated adjacent fragments, which were later used for primerless overlap extension. In addition, two flanking primers F₀ and R₀ were designed, which did not contain any mutational sites. Respective primer sequences and the corresponding melting temperatures (T_m) are listed in Appendix C (Table 58). In the first step, all intended fragments were amplified in separate reactions. The respective primer pairing, expected PCR product size and calculated elongation time for Step 1 and 2 of the OE-PCR protocol are summarized in Table 36.

Table 36. Fragment and primer pairing in the first two steps of OE-PCR. Elongation time was calculated based on the stated extension rate of *Pfu*Plus! DNA polymerase of 1 kb min⁻¹.

Fragment	Primer fwd	Primer rev	PCR product size	Elongation time
1	F ₀	R ₁	195 bp	11.7 sec
2	F ₁	R ₂	298 bp	17.9 sec
3	F ₂	R ₃	184 bp	11.0 sec
4	F ₃	R ₀	181 bp	10.9 sec
1+2	primerless		442 bp	27 sec
2+3	primerless		448 bp	27 sec
3+4	primerless		322 bp	19 sec

For easier handling, a common annealing temperature of 62°C and 20 sec elongation time was selected after verification. Detailed sample composition and temperature program are summarized in Table 37. To separate template DNA, the obtained PCR products were purified *via* preparative agarose gel electrophoresis (see 7.2.10) and extracted from the gel (see 7.2.5).

Table 37. Reaction composition (100 μ L) and temperature program for the amplification of overlapping fragments (OE-PCR, Step 1).

10 μ L	10x <i>Pfu</i> buffer	95°C	1 min	Initial denaturation
2 μ L	Primer fwd [10 μ M]	95°C	30 sec	Denaturation
2 μ L	Primer rev [10 μ M]	62°C	30 sec	
20 ng	Template plasmid	68°C	20 sec	Elongation
2 μ L	dNTPs [10 mM]	68°C	7 min	
5 U	<i>Pfu</i> Plus! Polymerase	4°C	∞	Final elongation
to 100 μ L ddH ₂ O				

In the second step, adjacent fragments were mixed in a pairwise fashion (Table 36) and primerless overlap extension was performed according to Table 38 in three separate reactions. Again, a common annealing temperature and elongation time was chosen.

Table 38. Reaction composition (20 μ L) and temperature program for primerless overlap extension of adjacent fragment pairs (OE-PCR, Step 2).

2 μ L	10x <i>Pfu</i> buffer	94°C	20 sec	Denaturation
2 μ L	dNTPs [2 mM]	56°C	30 sec	
10 ng	DNA Fragment A	68°C	30 sec	Elongation
10 ng	DNA Fragment B			
1 U	<i>Pfu</i> Plus! Polymerase			
to 20 μ L ddH ₂ O				

Subsequently, primerless overlap extension of mixed pairs was executed by mixing all samples of step 2 (3 x 20 μ L each) and pre-extension was performed for 20 cycles of 94°C for 20 sec and 68°C for 30 sec. Eventually, both flanking primers F₀ and R₀ were added to the reaction mixture (40 pmol of each primer

per 100 μ L reaction) to amplify the joined pairs following the temperature program stated in Table 39 in 30 cycles. This procedure was followed by additional 10 cycles of post-extension (Table 39). A suitable annealing temperature was chosen according to the T_m of the flanking primers and elongation was adapted to the expected size of the final fragment (730 bp). The obtained fragment was purified *via* preparative agarose gel electrophoresis (see 7.2.10) and extracted from the gel (see 7.2.5).

Table 39. Temperature program for the final amplification of the entire assembled fragment.

94°C	20 sec		Initial denaturation
94°C	20 sec	} 30x	Denaturation
60°C	30 sec		Annealing
68°C	1 min		Elongation
94°C	30 sec	} 10x	Denaturation
68°C	1 min		Final extension
4°C	∞		

To verify incorporation of the desired mutations, the obtained megaprimer was cloned into the pJET2.1/blunt cloning vector (CloneJET PCR Cloning Kit, Thermo Scientific) according to the manufacturer's protocol. Colonies were randomly picked and screened for correct-sized inserts *via* colony PCR (see 7.2.7) using pJET1.2 forward and reverse primer (Appendix C, Table 55). Plasmids from positive colonies were isolated (see 7.2.4) and sent for Sanger sequencing (see 7.2.9).

7.2.11.3 MEGAWHOP

A megaprimer PCR of whole plasmid (MEGAWHOP) was performed to generate a mutant library based on the mutagenic megaprimer created by OE-PCR (see 7.2.11.2).

Table 40. Reaction composition (50 μL) and temperature program used for MEGAWHOP.

5 μL 10x <i>Pfu</i> buffer	95°C 2 min	Initial denaturation
0.5 μg Megaprimer	95°C 30 sec	} 35x Denaturation
50 ng Template plasmid	68°C 6.5 min	
5 μL dNTPs [2 mM]	68°C 10 min	Final elongation
2.5 U <i>Pfu</i> Plus! Polymerase	4°C ∞	
to 50 μL ddH ₂ O		

A suitable vector construct containing the corresponding non-mutated parental gene was used as template. MEGAWHOP was performed according to the conditions summarized in Table 40. After the reaction, *DpnI* (2 μL) was added and the mixture incubated at 37°C for 2 h, followed by heat-inactivation at 80°C for 20 min. The reaction mixture was desalted before transformation of electrocompetent *E. coli* TOP10 (10 μL , see 7.2.3.1). After an elongated regeneration phase of 2 h at 37°C, aliquots (5 and 10 μL) were plated on selective LB agar plates and the remaining regenerated transformation mixture (700 μL) was used to inoculate antibiotic containing LB medium (10 mL). After incubation at 37°C overnight, colony forming units (cfu) were determined, and mixed plasmids were isolated from the liquid culture.

In order to ensure 95 % library coverage ($P_i = 0.95$), the number of colonies (cfu) projected to the total inoculation volume had to exceed the value T determined from Equation (8).³¹

$$T = -V \ln(1 - P_i) \quad (8)$$

T Number of transformants to be screened []

V Total number of gene mutants within the library []

P_i Probability to find a certain sequence ($P_i = 0.95$)

7.3 Biochemical methods

7.3.1 Heterologous production of recombinant enzymes

For heterologous production of recombinant enzymes, competent cells of a suitable *E. coli* strain were transformed with the respective expression construct

and overnight cultures were prepared (see 7.1.2 and 7.2.3). ONCs were used to inoculate main cultures, which were then cultivated until the desired cell density (OD_{600}) was obtained. After induction, cultures were typically incubated overnight (approx. 16-24 h) before cells were harvested (4 500 rpm, 15 min, 4°C) and washed with a suitable buffer. Cell pellets were directly used or stored at -20°C. Expression conditions for each construct are summarized in Table 41.

Table 41. Summarized conditions for heterologous production of enzymes used in this work. Conditions also apply for variants of the respective enzymes.

Construct	Strain	Medium	Cell growth	Induction	Production phase
pET28a_BsPAD	<i>E. coli</i> BL21(DE3)	TB-Kan ⁴⁰	37°C	$OD_{600} = 0.5-0.8$ 0.1 mM IPTG	20°C
pET28a_AMDase	<i>E. coli</i> BL21(DE3)	TB-Kan ⁴⁰	37°C	$OD_{600} = 0.5-0.8$ 1 mM IPTG	30°C
pET28a_AvLDO	<i>E. coli</i> BL21(DE3)	TB-Kan ⁴⁰	37°C	$OD_{600} = 0.5-0.8$ 0.5 mM IPTG	20°C
pET28a_BtIDO	<i>E. coli</i> BL21(DE3)	TB-Kan ⁴⁰	37°C	$OD_{600} = 0.5-0.8$ 1 mM IPTG	30°C

To study and confirm heterologous enzyme production, normalized samples of the main cultures were regularly taken from induction (t_0) through cell harvest according to Equation (9). The sampled biomass was harvested (4 500 rpm, 5 min, 4°C) and either directly used for SDS-PAGE (see 7.3.5.1) or stored at -20°C until use.

$$V_{sample} [mL] = \frac{7}{OD_{600}} \quad (9)$$

V_{sample} ... Sample volume [mL]

OD_{600} Optical density at 600 nm []

7.3.2 Preparation of cell-free extracts

For preparation of cell-free extracts, the cell wet weight (CWW) was determined, and the cell pellet was resuspended in the respective buffer to a certain cell

density (Table 42). Unless otherwise noted, routine sonication was performed with a Branson sonifier 250 ultrasonic homogenizer (Branson Ultrasonics Corp.) equipped with a 1/2" diameter stepped disruptor horn to typically process 25 mL samples in an ice-cooled beaker. For small-scale sonication, a Sonics Vibra-Cell VC600 ultrasonic homogenizer (Sonics & Materials Inc.) with a tapered 1/8" diameter microtip was used to sonicate 0.5-2 mL samples in 15 mL tubes, which were cooled in an ice-water-bath.

Table 42. Summarized sonication conditions for preparation of cell-free extracts of enzymes used in this work. Conditions also apply for variants of the respective enzymes.

Enzyme	Buffer	Purpose	Cell density ^[a]	Sonication parameters
BsPAD	50 mM KPi (pH 6)	Reaction buffer, freeze-drying	100 mg mL ⁻¹	Duty Cycle 5, Output control 50 %, 5 min
	20 mM Tris-HCl, 300 mM NaCl, 30 mM imidazole (pH 7.4)	His-tag purification binding buffer	100 mg mL ⁻¹	Duty Cycle 5, Output control 50 %, 5 min
AMDase	50 mM Tris-HCl (pH 8)	Reaction buffer	75 mg mL ⁻¹	Duty Cycle 5, Output control 50 %, 5 min
	50 mM Tris-HCl (pH 8)	Small scale sonication	75 mg mL ⁻¹	Duty Cycle 5, Output control 50 %, 2x 15 s
	20 mM Tris-HCl, 300 mM NaCl, 20 mM imidazole (pH 7.4)	His-tag purification binding buffer	75 mg mL ⁻¹	Duty Cycle 5, Output control 50 %, 5 min
AvLDO	20 mM Tris-HCl, 300 mM NaCl, (pH 7.4)	Preparation of CFE	100 mg mL ⁻¹	Duty Cycle 5, Output control 50 %, 3x 2 min ^[b]
	20 mM Tris-HCl, 300 mM NaCl, 10 mM imidazole (pH 7.4)	His-tag purification binding buffer	100 mg mL ⁻¹	Duty Cycle 5, Output control 50 %, 3x 2 min ^[b]
	10 mM HEPES (pH 7)	Reaction buffer	-	- ^[c]

[a] Cell density refers to cell wet weight (CWW) per volume. [b] After each sonication interval, samples were cooled on ice. [c] The purification buffer was exchanged to HEPES buffer (10 mM, pH 7) according to previous reports,²⁴¹ but was not used for sonication in this work.

The obtained crude lysates were cleared by centrifugation (11 000 rpm, 20-30 min, 4°C) (Avanti® J-20 XP, Beckman Coulter®, JA 25.50 rotor) and the supernatants were sterilized by filtration (0.22 µm).

7.3.2.1 Freeze-drying of cell-free extracts

For freeze drying, the obtained cell-free extracts were shock-frozen in liquid nitrogen in a suitable container and frozen samples were left under reduced pressure (approx. 15 µbar) overnight (AdVantage Pro Lyophilizer, SP Scientific). Lyophilized samples were stored at -20°C until use.

7.3.3 Purification of enzymes

Recombinantly produced proteins comprising a 6xHis-tag were purified from the cell-free extract *via* Ni-affinity chromatography. Nickel Sepharose 6 Fast Flow (GE Healthcare) was used as column material, following the manufacturer's gravity-flow purification protocol. Cell-free extracts were either directly prepared in binding buffer (Table 42) or in the corresponding imidazole-free buffer, in case the CFEs were stored at 4°C overnight. Initial binding conditions for purification were obtained by adding the respective amount of the corresponding elution buffer (containing 300 mM imidazole) prior purification. After binding to the column material (flow through fraction) and washing with 10-20 CV binding buffer (wash fraction), the target protein was eluted with elution buffer in 7x 1 CV fractions (elution fractions). The buffer of collected elution fractions was exchanged to the respective reaction buffer (Table 42) with PD-10 desalting columns (GE Healthcare). Purification fractions were analyzed *via* SDS-PAGE (see 7.3.5.2) and protein concentration was determined (see 7.3.4). If necessary, enzyme solutions were concentrated in Vivaspin 20 centrifugal concentrators (MWCO 10 kDa, Sartorius). In case of BsPAD, purified enzyme was stored in 20 % (v/v) glycerol at -20°C.

7.3.4 Determination of protein concentration

For determination of protein concentration either a Pierce BCA Protein Assay Kit (Thermo Scientific) or a Bio-Rad Protein Assay Dye Reagent Concentrate (Bio-Rad) was used according to the manufacturer's instructions. For calibration, dilutions of bovine serum albumin (BSA) in ddH₂O were prepared in the range of

0.1-2 mg mL⁻¹. Samples were prepared in a microtiter plate and measured on a plate reader (e.g. FLUOstar Omega, BMG Labtech) at 562 nm.

7.3.5 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

For SDS-PAGE, respective polyacrylamide gels were either hand-casted (4 % Bis-Tris stacking gel, 12 % Bis-Tris running gel, 10 or 15 wells; Appendix B, Table 52 and Table 53) or purchased from GenScript Biotech (12 % Bis-Tris ExpressPlus PAGE Gels, 12 or 15 wells). Running buffer was chosen according to the gel used. Samples meant for SDS-PAGE were combined with 4x SDS sample buffer and heated at 95°C for 5 min for denaturation. After centrifugation, samples were loaded, and gels were run at typically 200 V and 60 mA per gel for approx. 45-60 min. PageRuler Prestained Protein Ladder (7 µL, Thermo Scientific) was used as a standard. Protein bands were visualized either by UV light-induced fluorescence (gels supplemented with 2,2,2-trichloroethanol) or by Coomassie staining (Appendix B, Table 52).

7.3.5.1 SDS-PAGE of samples from protein expression studies

Samples retrieved for analysis of protein production as described in Section 7.3.1 were either lysed by small-scale ultrasonic homogenization (see 7.3.2; 500 µL buffer, 2x 15 s, output control 50 %, duty cycle 5) or by chemical lysis using BugBuster Master Mix (Merck Millipore) (150 µL). Either total cell lysates or separated insoluble and soluble fractions of the lysate were used for SDS-PAGE. In order to separate fractions, samples were centrifuged after cell lysis (13 200 rpm, 4°C, 10 min). After separating the supernatant (= soluble fraction), the insoluble fraction was resuspended in the same volume of buffer (typically 20 mM KPi, pH 7). Samples (5-10 µL) were mixed with SDS sample buffer (4x) and treated as described above.

7.3.5.2 SDS-PAGE of protein purification samples

After determining the protein content of fractions obtained from protein purification (see 7.3.3 and 7.3.4), aliquots containing approximately 20-30 µg total protein were mixed with SDS sample buffer (4x) and treated as described in Section 7.3.5.

7.4 Specific methods

7.4.1 Specific methods – *BsPAD*

7.4.1.1 Activity assay for *BsPAD*

Determination of specific activity of *BsPAD* and variants was carried out on a 1 mL scale in 1.5 mL tubes. A substrate stock containing **1a-1n** was freshly prepared (200 mM in DMSO). The substrate stock (50 μ L) was mixed with buffer (50 mM KP_i , pH 6) and the mixture was pre-incubated at 30°C (10 mM final substrate concentration, 5 % v/v DMSO). Typically, 5-50 μ g of purified enzyme (stored at -20°C in 20 % v/v glycerol) were added to start the reaction. Samples (100 μ L) were regularly collected (typically after 0, 1, 2, 5, 10, 20, 30 and 60 min) and immediately quenched by addition of acetonitrile containing 400 mM HCl (150 μ L). The mixture was thoroughly mixed and was directly used for HPLC-UV measurements after centrifugation and transfer of the supernatant to glass vials (see 7.5.3).

7.4.1.2 Preparation of 4-vinylphenol derivatives by *BsPAD* catalyzed decarboxylation

In a shake flask (1 L), the respective coumaric acid derivative (5 mmol) was mixed with KP_i buffer (50 mM, pH 6; 500 mL) (final concentration: 10 mM). Lyophilized cell-free extract (100 mg) of the respective *BsPAD* variant was added and the reaction mixture was incubated at 30°C and 130 rpm overnight. When full conversion was confirmed by TLC (CH/EA = 1:1, CAM or UV), the reaction mixture was extracted with MTBE (2 x 300 mL). The combined organic layers were washed with brine (1 x 300 mL) and dried over anhydrous Na_2SO_4 . Parts of the solvent were evaporated under reduced pressure to obtain a product solution in MTBE (50 mM), which ensured enhanced storage stability. The solution was stored at 4°C and the product was only isolated directly before use.

7.4.2 Specific methods – AMDase

7.4.2.1 pH-based microtiter plate activity screening

In order to screen prepared cell-free extracts for AMDase activity, a screening buffer was prepared, containing the pH indicator bromothymol blue (BTB) and AMDase substrate phenylmalonic acid (Appendix B, Table 47).

The screening buffer was thawed completely before use. In a microtiter plate, screening buffer (180 μL) and cell lysate (20 μL) or 50 mM Tris-HCl (pH 8) (20 μL) as negative control were combined and incubated at room temperature (typically overnight). Upon AMDase catalyzed decarboxylation of phenylmalonic acid, the pH increases and a color change from green to blue can be observed.

7.4.2.2 Activity assay for AMDase

For analytical scale conversions of substrates **7a-b** by AMDase variants, cell-free extracts containing the respective variant (75 $\text{mg}_{\text{CWW}} \text{mL}^{-1}$ in 50 mM Tris-HCl, pH 8; see 7.3.2) and a substrate stock (20 mM) in the same buffer were prepared freshly. In a 1.5 mL tube, cell-free extract (50 % (v/v) for AMDase CLGIPL and variants thereof or 5 % (v/v) for AMDase IPLL) was mixed with 50 % (v/v) substrate stock (20 mM). If applicable, buffer was added up to the desired total volume (typically 0.5 or 1 mL). Reactions were incubated at 30°C and 600 rpm until full conversion was detected by TLC (CH/EA/AcOH = 1:1:0.001, KMnO_4). After completion, the reaction mixture was acidified by addition of 2 M HCl (100 μL) and extracted in MTBE (1000 μL). The organic layer was separated and dried over anhydrous MgSO_4 . The obtained samples were directly used for chiral GC-FID analysis (see 7.5.2.2) or purified *via* semi-preparative TLC in case residual substrate was detected in the extracts (see 7.4.2.3). Remaining extracts were stored at -20°C in glass vials.

7.4.2.3 Semi-preparative TLC purification

For separation of remaining malonic acid substrate from chiral reaction products, the extracted reaction mixture (100 μL , see 7.4.2.2) was evenly spread on an analytical TLC plate (see 7.5.4) over the length of 3.5 cm. After developing the plate (CH/EtOAc/AcOH = 1:3:0.001), a narrow strip was cut off along the separation axis, which was stained with KMnO_4 to locate the products. The silica

gel within the corresponding unstained area was scraped off and collected before addition of ethyl acetate (200 μ L). The resulting suspension was thoroughly mixed followed by centrifugation (13 200 rpm, 5 min, 4°C). The supernatant was transferred to a glass vial and directly used for chiral GC-FID analysis (see 7.5.2.2; method GC-FID_E).

7.4.2.4 Determination of enantiomeric excess (*ee*)

In order to verify appropriate chiral separation, control reactions with AMDase IPLL or alternatively CLGIPL and 2-methyl-2-vinylmalonic acid (**7a**) as substrate were performed as described in Section 7.4.2.2. The extracted reaction products (*R/S*)-**8a** were analyzed with chiral GC-FID (see 7.5.2.2; method GC-FID_D). Based on previously reported enantioselectivities of the respective AMDase variants (IPLL: > 99 %*ee* (*R*); CLGIPL: 66 %*ee* (*S*)⁴⁶), peaks corresponding to (*R*)-**8a** ($t_R = 11.8$ min) and (*S*)-**8a** ($t_R = 11.3$ min) were assigned (Appendix I, Figure S 41). Peaks were integrated and the resulting peak areas were used to calculate the *enantiomeric excess* (*ee*) according to Equation (10).

$$\% ee = \frac{|area_{(S)} - area_{(R)}|}{area_{(S)} + area_{(R)}} \times 100 \quad (10)$$

In some instances, the initially applied method GC-FID_D (100°C, isothermal) led to insufficient separation of product enantiomers (*R/S*)-**8a**, so that the oven temperature was reduced to 80°C for further measurements (GC-FID_E, see 7.5.2.2). In case of residual substrate **7a**, reaction mixtures were purified *via* semi-preparative TLC prior chiral analysis to prevent spontaneous decarboxylation of **7a** during injection (see 7.4.2.3).

7.4.3 Specific methods – AvLDO

7.4.3.1 Enrichment of AvLDO variants

For the selection procedure, recombinant strains of *E. coli* BL21(DE3) $3\Delta sucA$ were cultivated in liquid or on solid minimal medium supplemented with antibiotics, IPTG, α -ketoglutarate and substrate of interest (Appendix B, Table 48 and

Table 49). Minimal medium was prepared freshly and typically not used for more than 7 days. The pre-made medium was stored at 4°C and cultivation flasks were filled with medium just before inoculation. IPTG and antibiotics were always added freshly (0.5 mM IPTG, 40 µg mL⁻¹ kanamycin). Filter tips were used when handling minimal medium components or cultures to prevent contaminations.

For enrichment of AvLDO variants, the plasmid library created *via* MEGAWHOP (2 µL of mixed plasmids, see 7.2.11.3) was used to transform electrocompetent *E. coli* BL21(DE3) 3Δ*sucA* (80 µL cells, see 7.2.3.1). In parallel, the same strain was also transformed with pET28a_AvLDO_WT (1 µL). After regenerating the cells in LB-SOC (700 µL) for 2 h at 37°C, aliquots of each transformation reaction (5 µL) were plated on LB-Kan⁴⁰ agar plates for evaluation of transformation efficiency.

From the remaining cell suspension of *E. coli* BL21(DE3) 3Δ*sucA* [AvLDO_library], several aliquots were directly spread on M9-Kan⁴⁰/IPTG/αKG/L-Ile agar plates (10x 5 µL) to test selection directly after transformation.²⁵⁵ The remaining cell suspension (600 µL) was used to inoculate M9-Kan⁴⁰/IPTG/αKG/L-Ile (20 mL in 100 mL baffled flasks; culture t₀). The same was done for *E. coli* BL21(DE3) 3Δ*sucA* [AvLDO_WT]. Liquid cultures were incubated at 37°C and 130 rpm for 24 h and plates were incubated at 37°C until colonies were visible. After 24 h, each culture was diluted into fresh medium (200 µL inoculation volume; culture t₁, 20 mL; *vide infra*), aliquots were spread on M9-Kan⁴⁰/IPTG/αKG/L-Ile agar plates (2.5 × 10⁻⁴ or 1 × 10⁻³ µL *via* preparation of two consecutive 1:100 and one 1:40 or 1:10 dilution in medium, respectively and plating 100 µL of the most diluted sample) and the optical density (OD₆₀₀) was determined. The remaining culture containing the AvLDO library was harvested (4 500 rpm, 5 min, 4°C) and the biomass was stored at -20°C for subsequent mixed plasmid isolation, if required. This workflow was repeated as required. Interesting clones were picked based on formation of larger colonies, or randomly if no differential colony size was observed. Clones were restreaked and stored on LB-Kan⁴⁰ plates at 4°C until conduction of a rescreening (see 7.4.3.2).

7.4.3.2 Rescreening of selected AvLDO variants

For rescreening of selected clones, sterile glass test tubes (straight cut, 17 mL) were filled with M9-Kan⁴⁰/IPTG/αKG/L-Ile (5 mL; Appendix B, Table 48 and

Table 49) and inoculated with cell material from previously prepared LB-Kan⁴⁰ plates containing the restreaked colonies (see 7.4.3.1). Cultures were incubated at 37°C and 130 rpm for 24 h, each culture was diluted into fresh medium (50 µL culture in 5 mL medium) and incubated again under the same conditions for 24 h. After both incubation intervals, optical density (OD₆₀₀) of the cultures was determined. Additionally, after the second interval, the culture supernatants were analyzed *via* TLC (2 µL supernatant; 1-butanol/AcOH/H₂O = 22:5:3 or acetone/1-butanol/ AcOH/H₂O = 7:7:2:4, ninhydrin; see 7.5.4) or GC-MS, which required freeze-drying prior to sample derivatization (see 7.5.2.3). For this purpose, culture supernatant (100 µL) was transferred to screw cap glass vials (2 mL) which were kept at -80°C until completely frozen before lyophilization.

7.4.3.3 *In-vitro* activity assay for AvLDO

In order to detect AvLDO activity *in vitro*, cell-free extracts (20 mM Tris-HCl, 300 mM NaCl, pH 7.4; see 7.3.2) or purified enzyme (in 10 mM HEPES, pH 7; see 7.3.3) were prepared in advance. Additionally, all reaction components were freshly prepared as concentrated stock solutions in ddH₂O (Table 43).

Table 43. Stock solutions and composition of reaction solution used for AvLDO *in-vitro* activity assay. All stock solutions were freshly prepared in ddH₂O.

Component	Stock solution			Reaction solution	
	Molar mass (g mol ⁻¹)	Concentration (mM)	mg mL ⁻¹	Concentration (mM)	Volume (µL)
Substrate (Leu or Ile)	131.18	50	6.6	5	50
α-Ketoglutaric acid	146.10	100	14.6	10	50
Sodium L-ascorbate	198.11	100	19.8	10	50
FeSO ₄ ·7 H ₂ O	278.02	50	13.9	0.5	5
NaOAc·3 H ₂ O ^[a]	136.08	1000	136.1	50	25
Enzyme (CFE/purified)	-	-	-	-	x ^[b]
ddH ₂ O	-	-	-	-	to 500

[a] Buffer stock: dissolve 1 M (136.08 mg mL⁻¹) NaOAc·3 H₂O and adjust pH to 4 with concentrated AcOH. [b] Typically, the max. volume of CFE was used (320 µL) and 0.5-1 mg mL⁻¹ of purified enzyme.

All reaction components were mixed in 2 mL screw cap glass vials and the reaction was initiated by addition of enzyme solution. Typically, 0.5-1 mg mL⁻¹ of purified enzyme or the maximum possible volume of CFE was used for the reaction. In parallel, blank reactions were prepared containing all components except enzyme. The reaction mixtures were incubated at 25°C and 750 rpm and samples (50 µL) were taken regularly and quenched by addition of acetic acid (50 % (v/v), 25 µL). After thorough mixing, samples were centrifuged (13 200 rpm, 3 min, 4°C) and the separated supernatants were stored at -20°C until analyzed by TLC. For TLC analysis, 2 µL of quenched reaction mixture were applied together with reference compounds (substrate and 4-hydroxyisoleucine), which were prepared in the same way as samples (reference compound (5 mM):AcOH (50 % v/v) = 2:1). A mixture of 1-butanol/AcOH/H₂O = 22:5:3 or acetone/1-butanol/AcOH/H₂O = 7:7:2:4 was typically used as mobile phase and 0.2 % ninhydrin in acetone as staining reagent (see 7.5.4).

7.4.3.4 Resting cell assay

For preparation of resting cells, the respective recombinant *E. coli* BL21(DE3) strains were cultivated, and heterologous expression was induced as described in Section 7.3.1. Alternatively, cells were harvested 6 h after induction or after overnight incubation. Cells were washed with the respective buffer used for the resting cell assay (100 mM Tris-HCl or HEPES, pH 7) and frozen at -80°C to enhance cell wall permeability.²⁵⁵

All stock solutions, except buffers, required for the resting cell assay were freshly prepared in ddH₂O (Table 44). A suitable amount of cell material was suspended in the respective buffer used for the resting cell assay (200 mg_{cwd} mL⁻¹ in 100 mM Tris-HCl or HEPES, pH 7). All reaction components were mixed in screw cap glass vials (2 mL reaction volume in 20 mL vials) according to Table 44 and reactions were incubated at 37°C and 130 rpm (typically 24-72 h in total). Samples were taken regularly (100 µL), and cells were separated by centrifugation (13 000 rpm, 1 min, 4°C). The resulting supernatants were kept at -20°C until analyzed either by TLC or GC-MS. For TLC analysis, 2 µL of supernatants were applied on TLC plates together with reference compounds (substrate and 4-hydroxyisoleucine, 10 mM in ddH₂O). A mixture of 1-butanol/AcOH/H₂O = 22:5:3 or acetone/1-butanol/AcOH/H₂O = 7:7:2:4 was typically used

as mobile phase and 0.2 % ninhydrin in acetone as staining reagent (see 7.5.4). For GC-MS analyses, the supernatant was freeze-dried prior to sample derivatization (see 7.5.2.3). Therefore, culture supernatant (100 μ L) was transferred to screw cap glass vials (2 mL), which were kept at -80°C until completely frozen before lyophilization.

Table 44. Stock solutions and composition of reaction solution used for AvLDO resting cell assay. All stock solutions were freshly prepared in ddH₂O.

Component	Stock solution			Reaction solution	
	Molar mass (g mol ⁻¹)	Concentration (mM)	mg mL ⁻¹	Concentration	Volume (μ L)
Substrate (Leu or Ile)	131.18	100	13.1	10 mM	200
α -Ketoglutaric acid	146.10	100	14.6	10 mM	200
Sodium L-ascorbate	198.11	100	19.8	10 mM	200
FeSO ₄ ·7 H ₂ O	278.02	50	13.9	5 mM	200
Buffer ^[a]	-	1000	-	100 mM	100 ^[b]
Resting cells	-	-	200 ^[b]	100 mg mL ⁻¹	1000
ddH ₂ O	-	-	-	-	100

[a] Different buffers were tested including Tris-HCl (pH 7) and HEPES (pH 7). Both buffers were prepared as 1 M stock solution. [b] Resting cells were already suspended in the respective 100 mM diluted buffer to 200 mg mL⁻¹ (2x stock), therefore only 100 μ L of 1 M buffer stock (10x) were added to the final reaction mixture.

7.5 Analytical methods

7.5.1 Nuclear magnetic resonance (NMR)

Standard NMR experiments (¹H, ¹³C, APT, H,H-COSY, HSQC) were performed on a Bruker Avance III 300 MHz FT NMR spectrometer (with 300.36 MHz for ¹H and 75.53 MHz for ¹³C) equipped with an autosampler. Residual solvent signals were used as internal standard for referencing chemical shifts δ (in ppm) (CDCl₃: δ = 7.26 ppm (¹H), 77.3 ppm (¹³C); DMSO-d₆: δ = 2.50 ppm (¹H), 39.5 ppm (¹³C); D₂O: δ = 4.79 ppm (¹H)).²⁸¹ Signal multiplicities were abbreviated as follows: singlet (s), broad singlet (bs), doublet (d), doublet of doublets (dd), triplet (t),

quadruplet (q), multiplet (m). Quaternary carbon atoms were indicated as C_q. Chemical shifts were given in parts per million (ppm) and coupling constants in Hertz (Hz). For sensitivity enhancement and clarity of the recorded spectra, all ¹³C experiments were ¹H decoupled.

7.5.2 Gas chromatography (GC)

7.5.2.1 Gas chromatography with mass spectrometry (GC-MS)

All GC-MS measurements were performed on a Shimadzu GCMS-QP2010 SE device equipped with an AOC-20i/s autosampler and injector unit (Shimadzu) and helium as carrier gas. Typically, a Zebron ZB-5MSi capillary column (30 m × 0.25 mm × 0.25 μm) (Phenomenex) was used for separation. Methods are summarized in Table 45 and detailed method information is stated in Appendix I.

Table 45. Summarized methods used for GC-MS analysis and respective key parameters. Typically, 1 μL of sample was injected.

Method	GC-MS_A	GC-MS_B	GC-MS_C
Flow control mode	Linear velocity (39.5 cm sec ⁻¹)	Linear velocity (39.5 cm sec ⁻¹)	Linear velocity (36.7 cm sec ⁻¹)
Total flow	11.5 mL min ⁻¹	15.0 mL min ⁻¹	13.1 mL min ⁻¹
Column flow	1.13 mL min ⁻¹	1.19 mL min ⁻¹	1.00 mL min ⁻¹
Split ratio	6.5	9.1	9.1
Injection temperature	250°C	250°C	250°C
Temperature program	100°C for 4 min, 20°C min ⁻¹ to 300°C, 300°C for 5 min (Total: 19.0 min)	50°C for 5 min, 40°C min ⁻¹ to 300°C, 300°C for 5 min (Total: 16.25 min)	70°C, 10°C min ⁻¹ to 170°C, 30°C min ⁻¹ to 280°C, 280°C for 3 min (Total: 16.6 min)
MS mode	Scan (0.13 sec), <i>m/z</i> 50-500	Scan (0.13 sec), <i>m/z</i> 50-500	Scan (0.13 sec), <i>m/z</i> 50-500

7.5.2.2 Chiral gas chromatography with flame ionization detection (chiral GC-FID)

All chiral GC-FID measurements were performed on a Shimadzu Nexis GC-2030 device with flame-ionization detector (FID) equipped with an AOC-20i/s

autosampler and injector unit (Shimadzu) and nitrogen as carrier gas. Typically, a HYDRODEX β -6TBDM (25 m \times 0.25 mm, proprietary df) (Macherey Nagel) was used for separation. Methods are summarized in Table 46 and detailed method information is stated in Appendix I.

Table 46. Summarized methods used for chiral GC-FID analysis and respective key parameters. Typically, 1 μ L of sample was injected.

Method	GC-FID_D	GC-FID_E
Flow control mode	Linear velocity (44.8 cm sec ⁻¹)	Linear velocity (44.8 cm sec ⁻¹)
Total flow	35.4 mL min ⁻¹	43.6 mL min ⁻¹
Column flow	1.54 mL min ⁻¹	1.93 mL min ⁻¹
Split ratio	20	20
Injection temperature	250°C	250°C
Temperature program	100°C, isotherm (Total: 60 min)	80°C, isotherm (Total: 60 min)

7.5.2.3 Sample derivatization with BSTFA

For derivatization with BSTFA, supernatants obtained from rescreening of AvLDO variants (see 7.4.3.2) or resting cell assays (see 7.4.3.4) (typically 100 μ L) were freeze-dried as described in the respective sections. The dried sample was mixed with acetonitrile (100 μ L) and BSTFA (100 μ L) directly in the glass vial. The vials were tightly sealed before treatment in an ultrasonic bath for 1 min (floating in a Styrofoam rack) to ensure fine dispersion of the insoluble sample components. Subsequently, after checking again that all screw caps are closed tightly, samples were heated to 100°C for 30-40 min in a thermo block. After cooling to room temperature, samples were transferred into 1.5 mL tubes and solids were separated by centrifugation (13 200 rpm, 5 min, 4°C). The supernatant (100 μ L) was further transferred into GC-vials and directly used for GC-MS analysis (see 7.5.2.1, method GC-MS_C).

7.5.3 High-performance liquid chromatography (HPLC)

All HPLC measurements were conducted on an Agilent Technologies 1100 Series HPLC device (Agilent Technologies) equipped with an autosampler unit and a DAD detector. Typically, a reversed-phase Nucleodur C₁₈ Pyramid column (5 μ m, 250 x 4.6 mm) (Macherey Nagel) was used. Absolute retention times might vary between different batches due to a flexible system setup and hence frequent exchange of capillaries.

7.5.3.1 HPLC method and calibration for determination of specific activity of BsPAD

As eluent, a mixture of acetonitrile and ddH₂O (55 % (v/v) acetonitrile, isocratic, 1 mL min⁻¹) was used and absorption was monitored simultaneously at 254, 280 and 320 nm. Depending on the substrate, 1-2 μ L of each sample were injected (see Appendix J). A wavelength suitable for calibration was chosen for each substrate. Representative retention times for substrates and products and the selected detection wavelengths are summarized in Appendix J (Table 68).

For calibration, dilutions of each substrate were prepared in the range of 0.5-10 mM starting from a 10 mM substrate solution in reaction buffer as described in Section 7.4.1.1. All calibration samples were prepared in triplicates by mixing the respective dilution (100 μ L) with acetonitrile (150 μ L, containing 400 mM HCl). Resulting standard curves are presented in Appendix J (Figure S 43 to Figure S 53).

7.5.4 Thin layer chromatography (TLC)

For thin layer chromatography, silica coated aluminum sheets were purchased from Merck Millipore (silica gel 60 F₂₅₄, 20 x 20 cm). Prior separation, plates were cut into desired size (typically 5 cm length for separation). Depending on the analytes, different solvent mixtures were used for separation (see Appendix G0). For detection, UV light (λ = 254 nm) and different staining reagents were used followed by development in a hot air stream. Preparation of TLC staining reagents is described in Appendix B (Table 54).

7.6 Chemical methods

7.6.1 General experimental aspects, materials and methods

All chemicals and solvents were purchased from abcr, Alfa Aesar, Acros Organics, Carl Roth, TCI Chemicals or Sigma Aldrich and typically used as received, unless stated otherwise.

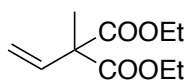
Standard Schlenk techniques were used to perform reactions under inert conditions. Therefore, a suitable reaction vessel was evacuated, flame dried and purged with inert gas (N₂ or Ar) three times consecutively. Sampling or addition of chemicals was carried out under inert gas counter flow to ensure exclusion of moisture and oxygen.

7.6.2 Flash column chromatography

For preparative flash chromatography, silica gel was purchased from Macherey Nagel (Silica gel 60, normal phase, 0.04-0,063 mm particle size, 60 Å pore size). The amount of silica gel was typically adjusted to the 50- to 100-fold (w/w) crude yield. Columns were chosen accordingly to obtain a filling level between 15 and 25 cm. Sticky crude products were adsorbed on silica prior to loading onto the readily packed column by adding the 1.5-fold (w/w) amount of silica and a small amount of volatile solvent (*e.g.* DCM), which was again evaporated under reduced pressure thereafter. Fraction size was typically chosen between 25-30 % of the total column volume. Progress of flash column chromatography was monitored *via* TLC (see 7.5.4).

7.6.3 Experimental procedures

7.6.3.1 Synthesis of diethyl 2-methyl-2-vinylmalonate (**6a**)



6a

In a flame-dried Schlenk flask, absolute THF (20 mL) and absolute diisopropylamine (1.5 mL, 11 mmol, 2 eq) were mixed and the colorless solution was cooled to -78°C . Then, *n*-BuLi (5.2 mL, 2.5 M in hexanes, 12.8 mmol, 2.4 eq) was added dropwise and the resulting yellowish solution was stirred at -78°C . After 10 min, absolute DMPU (10 mL) and diethyl 2-ethylidenemalonate (980 μL , 5.37 mmol, 1 eq) were added and the mixture was stirred at -78°C for 45 min before iodomethane (500 μL , 8.03 mmol, 1.5 eq) was added. The slightly turbid mixture was allowed to warm to room temperature and was stirred overnight. Conversion was checked by TLC (CH/EA = 5:1; KMnO_4) and GC-MS. In case of incomplete conversion, presence of reactive species was detected *via* GC-MS after quenching a small aliquot of the reaction mixture with D_2O . Additional MeI (0.5 eq) was then added, and the mixture was stirred at r.t. overnight. When full conversion was detected, the reaction mixture was quenched by addition of saturated NH_4Cl solution (30 mL) and extracted with ethyl acetate (3x 30 mL). The combined organic layers were washed with brine (1x 100 mL) and dried over Na_2SO_4 before the solvents were evaporated under reduced pressure. The crude brown liquid was purified by column chromatography (500 mL silica, CH/EA = 20:1, 150 mL fractions).

Yield 266.9 mg (1.33 mmol, 25 %), colorless oil

$\text{C}_{10}\text{H}_{16}\text{O}_4$ [200.23]

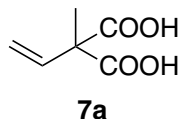
GC-MS t_{R} = 8.46 min (method: GC-MS_B)
 m/z = 155 (5 %) [$\text{M}^+ - \text{C}_2\text{H}_5\text{O}$], 127 (45 %) [$\text{M}^+ - \text{C}_3\text{H}_5\text{O}_2$], 99 (100 %) [$\text{M}^+ - \text{C}_3\text{H}_5\text{O}_2$ and $-\text{C}_2\text{H}_4$], 82 (35 %) [$\text{M}^+ - \text{C}_3\text{H}_5\text{O}_2$ and $-\text{C}_2\text{H}_5\text{O}$].

TLC R_{f} = 0.56 (CH/EA = 5:1; KMnO_4)

$^1\text{H-NMR}$ (300.36 MHz, CDCl_3) δ = 6.29 (dd, $^3J_{\text{HH}}$ = 17.6 Hz, $^3J_{\text{HH}}$ = 10.7 Hz, 1H, $\text{CH}=\text{CH}_2$), 5.22 (dd, $^3J_{\text{HH}}$ = 19.5 Hz, $^3J_{\text{HH}}$ = 14.2 Hz, 2H, $\text{CH}=\text{CH}_2$), 4.19 (q, 4H, OCH_2CH_3), 1.54 (s, 3H, CH_3), 1.25 (t, $^3J_{\text{HH}}$ = 7.1 Hz, 6H, OCH_2CH_3) ppm.

¹³C-NMR (75.53 MHz, CDCl₃) δ = 171.1 (C=O), 136.2 (CH=CH₂), 116.1 (CH=CH₂), 61.7 (CH₂), 56.3 (C_q), 19.7 (CH₃), 14.1 (CH₃) ppm.

7.6.3.2 Synthesis of 2-methyl-2-vinylmalonic acid (**7a**)



Diethyl 2-methyl-2-vinylmalonate (245 mg, 1.22 mmol, 1 eq) was dissolved in EtOH (25 mL) and cooled to 4°C. NaOH (12 mL, 8 M) was added dropwise, and the reaction mixture was allowed to warm to r.t. overnight. Quantitative conversion was detected by TLC (CH/EtOAc/AcOH = 1:1:0.001; KMnO₄). The suspension was acidified to pH 4 with HCl and the solvents were evaporated. The aqueous layer was extracted with EtOAc (3x 30 mL). The combined organic layers were washed with brine (1x 100 mL) and dried over anhydrous Na₂SO₄. Solvents were evaporated under reduced pressure to yield a slightly brownish solid as crude product (150 mg, 85 %). The crude product was suspended in *n*-pentane and sonicated. The supernatant was removed, and the procedure was repeated two more times. The obtained product was dried *in vacuo*.

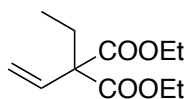
Yield 130.0 mg (0.90 mmol, 74 %), almost colorless solid

C₆H₈O₄ [144.13]

TLC R_f = 0.10 (CH/EtOAc/AcOH = 1:1:0.001; KMnO₄)

¹H-NMR (300.36 MHz, DMSO-d₆) δ = 12.85 (bs, 2H, COOH), 6.22 (dd, ³J_{HH} = 17.6, ³J_{HH} = 10.8 Hz, 1H, CH=CH₂), 5.16 (m, 2H, CH=CH₂), 1.38 (s, 3H, CH₃) ppm.

¹³C-NMR (75.53 MHz, DMSO-d₆) δ = 172.2 (C=O), 137.4 (CH=CH₂), 114.9 (CH=CH₂), 55.6 (C_q), 19.2 (CH₃) ppm.

7.6.3.3 Synthesis of diethyl 2-ethyl-2-vinylmalonate (**6b**)**6b**

In a flame-dried Schlenk flask, absolute THF (20 mL) and absolute diisopropylamine (1.5 mL, 11 mmol, 2 eq) were mixed and the colorless solution was cooled to -78°C . Then, *n*-BuLi (5.2 mL, 2.5 M in hexanes, 12.8 mmol, 2.4 eq) was added dropwise and the resulting yellowish solution was stirred at -78°C . After 10 min, absolute DMPU (10 mL) and diethyl 2-ethylidenemalonate (980 μL , 5.37 mmol, 1 eq) were added and the mixture was stirred at -78°C for 45 min before iodoethane (650 μL , 8.06 mmol, 1.5 eq) were added. The slightly orange mixture was allowed to warm to room temperature and was stirred overnight. Conversion was checked by TLC (CH/EA = 5:1; KMnO_4) and GC-MS. In case of incomplete conversion, additional EtI (0.5 eq) was then added, and the mixture was stirred at r.t. overnight. When full conversion was detected, the reaction mixture was quenched by addition of saturated NH_4Cl solution (30 mL) and extracted with ethyl acetate (3x 30 mL). The combined organic layers were washed with brine (1x 100 mL) and dried over anhydrous Na_2SO_4 before the solvents were evaporated under reduced pressure. The crude yellowish oil was purified by column chromatography (500 mL silica, CH/EA = 20:1, 150 mL fractions).

Yield 431.4 mg (2.01 mmol, 38 %), colorless oil

$\text{C}_{11}\text{H}_{18}\text{O}_4$ [214.26]

GC-MS t_{R} = 8.83 min (method: GC-MS_B)

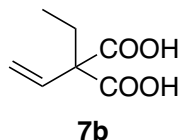
m/z = 169 (3 %) [$\text{M}^+ - \text{C}_2\text{H}_5\text{O}$], 141 (83 %) [$\text{M}^+ - \text{C}_3\text{H}_5\text{O}_2$], 113 (100 %) [$\text{M}^+ - \text{C}_3\text{H}_5\text{O}_2$ and $-\text{C}_2\text{H}_4$].

TLC R_{f} = 0.62 (CH/EA = 5:1; KMnO_4)

$^1\text{H-NMR}$ (300.36 MHz, CDCl_3) δ = 6.32 (dd, $^3J_{\text{HH}}$ = 17.8 Hz, $^3J_{\text{HH}}$ = 10.9 Hz, 1H, $\text{CH}=\text{CH}_2$), 5.23 (dd, $^3J_{\text{HH}}$ = 39.5 Hz, $^3J_{\text{HH}}$ = 14.3 Hz, 2H, $\text{CH}=\text{CH}_2$), 4.20 (q, 4H, OCH_2CH_3), 2.08 (q, 2 H, CH_2CH_3), 1.24 (t, $^3J_{\text{HH}}$ = 7.1 Hz, 6H, OCH_2CH_3), 0.85 (t, $^3J_{\text{HH}}$ = 7.5 Hz, 3H, CH_2CH_3) ppm.

¹³C-NMR (75.53 MHz, CDCl₃) δ = 170.6 (C=O), 134.8 (CH=CH₂), 116.7 (CH=CH₂), 61.5 (OCH₂CH₃), 60.7 (C_q), 28.1 (CH₂CH₃), 14.2 (OCH₂CH₃), 8.8 (CH₂CH₃) ppm.

7.6.3.4 Synthesis of 2-ethyl-2-vinylmalonic acid (**7b**)



Diethyl 2-ethyl-2-vinylmalonate (431.4 mg, 2.01 mmol, 1 eq) was dissolved in EtOH (25 mL) and cooled to 4°C. NaOH (12 mL, 8 M) was added dropwise, and the reaction mixture was allowed to warm to r.t. overnight. The suspension was acidified to pH 4 with conc. HCl and the solvents were evaporated. The aqueous layer was extracted with EtOAc (3x 30 mL) and the combined organic layers were washed with brine (1x 100 mL) and dried over Na₂SO₄. Solvents were evaporated under reduced pressure to yield a partly crystallized yellowish oil as crude product. The crude product was suspended in *n*-pentane and sonicated. The supernatant was removed, and the procedure was repeated two more times. The product could be obtained as slightly yellow solid, after drying *in vacuo*.

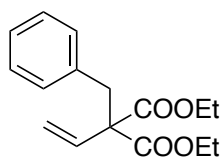
Yield 192.3 mg (1.2 mmol, 60 %), slightly yellowish solid

C₇H₄O₁₀ [158.15]

TLC R_f = 0.18 (CH/Ea = 1:1; KMnO₄)

¹H-NMR (300.36 MHz, DMSO-d₆) δ = 12.86 (s, 2H, COOH), 6.22 (dd, ³J_{HH} = 17.9, ³J_{HH} = 10.9 Hz, 1H, CH=CH₂), 5.26-5.08 (m, 2H, CH=CH₂), 1.92 (q, 2 H, CH₂CH₃), 0.77 (t, ³J_{HH} = 7.4, 3H, CH₂CH₃) ppm.

¹³C-NMR (75.53 MHz, DMSO-d₆) δ = 171.6 (C=O), 135.8 (CH=CH₂), 115.6 (CH=CH₂), 59.9 (C_q), 27.0 (CH₂), 8.7 (CH₃) ppm.

7.6.3.5 Synthesis of diethyl 2-benzyl-2-vinylmalonate (**6c**)**6c**

In a flame-dried Schlenk flask, absolute THF (20 mL) and absolute diisopropylamine (1.5 mL, 11 mmol, 2 eq) were mixed and the colorless solution was cooled to -78°C . Then, *n*-BuLi (5.2 mL, 2.5 M in hexanes, 12.8 mmol, 2.4 eq) was added dropwise and the resulting yellowish solution was stirred at -78°C . After 10 min, absolute DMPU (10 mL) and diethyl 2-ethylidenemalonate (980 μL , 5.37 mmol, 1 eq) were added and the mixture was stirred at -78°C for 45 min before benzyl bromide (900 μL , 7.5 mmol, 1.4 eq) was added. The slightly orange mixture was allowed to gradually warm to room temperature overnight. Conversion was monitored by TLC (CH/EA = 5:1; KMnO_4). In case of incomplete conversion, additional benzyl bromide (0.5 eq) was added and the mixture was again stirred at r.t. overnight. When full conversion was detected, the reaction mixture was quenched by addition of saturated NH_4Cl solution (30 mL) and extracted with ethyl acetate (3x 30 mL). The combined organic layers were washed with brine (1x 100 mL) and dried over Na_2SO_4 before the solvents were evaporated under reduced pressure. The crude orange oil was purified by column chromatography (500 mL silica, CH/EA = 20:1, 150 mL fractions).

Yield 453.4 mg (1.6 mmol, 31 %), colorless oil

$\text{C}_{12}\text{H}_{12}\text{O}_4$ [220.22]

TLC $R_f = 0.50$ (CH/EA = 5:1; KMnO_4 or CAM)

$^1\text{H-NMR}$ (300.36 MHz, CDCl_3) $\delta = 7.19\text{--}7.04$ (m, 5H, H_{Ar}), 6.13 (dd, 1H, $\text{CH}=\text{CH}_2$), 5.20 (dd, 2H, $\text{CH}=\text{CH}_2$), 4.13 (q, $^3J_{\text{HH}} = 7.1$ Hz, 4H, OCH_2CH_3), 3.31 (s, 2H, CH_2Bz), 1.16 (t, $^3J_{\text{HH}} = 7.1$ Hz, 6H, OCH_2CH_3) ppm.

$^{13}\text{C-NMR}$ (75.53 MHz, CDCl_3) $\delta = 170.2$ (C=O), 135.9 ($\text{C}_{\text{q,Ar}}$), 135.5 ($\text{CH}=\text{CH}_2$), 130.3 (2 C_{Ar}), 128.2 (2 C_{Ar}), 127.0 (C_{Ar}), 116.7 ($\text{CH}=\text{CH}_2$), 61.7 (CH_2), 61.5 (C_{q}), 42.1 (CH_2), 14.1 (CH_3) ppm.

8 References

- (1) Gröger, H.; Asano, Y. Introduction - Principles and Historical Landmarks of Enzyme Catalysis in Organic Synthesis. In *Enzyme Catalysis in Organic Synthesis*; Drauz, K., Gröger, H., May, O., Eds.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2012; pp 1–42.
- (2) Faber, K. Introduction and Background Information. In *Biotransformations in Organic Chemistry*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2011; pp 1–30.
- (3) Lichtenthaler, F. W. 100 Years“Schlüssel-Schloss-Prinzip”: What Made Emil Fischer Use This Analogy? *Angew. Chem., Int. Ed.* **1995**, *33*, 2364–2374.
- (4) Agarwal, P. K. Enzymes: An Integrated View of Structure, Dynamics and Function. *Microb. Cell Fact.* **2006**, *5*, 2.
- (5) Hedstrom, L. Enzyme Specificity and Selectivity. In *Encyclopedia of Life Sciences*; John Wiley & Sons, Ltd: Chichester, UK, 2010.
- (6) Hollmann, F.; Otten, L. G. Enantioselectivity of Enzymes. In *Wiley Encyclopedia of Chemical Biology*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2009.
- (7) Buchholz, K.; Collins, J. The Roots—a Short History of Industrial Microbiology and Biotechnology. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 3747–3762.
- (8) Bornscheuer, U. T.; Huisman, G. W.; Kazlauskas, R. J.; Lutz, S.; Moore, J. C.; Robins, K. Engineering the Third Wave of Biocatalysis. *Nature* **2012**, *485*, 185–194.
- (9) Bornscheuer, U. T. The Fourth Wave of Biocatalysis Is Approaching. *Phil. Trans. R. Soc. A* **2018**, *376*, 20170063.
- (10) Sheldon, R. A.; Brady, D. The Limits to Biocatalysis: Pushing the Envelope. *Chem. Commun.* **2018**, *54*, 6088–6104.
- (11) Woodley, J. M. Accelerating the Implementation of Biocatalysis in Industry. *Appl. Microbiol. Biotechnol.* **2019**, *103*, 4733–4739.
- (12) Torrelo, G.; Hanefeld, U.; Hollmann, F. Biocatalysis. *Catal. Lett.* **2015**, *145*, 309–345.
- (13) Woodley, J. M. New Frontiers in Biocatalysis for Sustainable Synthesis. *Curr. Opin. Green Sustain. Chem.* **2020**, *21*, 22–26.
- (14) Adams, J. P.; Brown, M. J. B.; Diaz-Rodriguez, A.; Lloyd, R. C.; Roiban, G. Biocatalysis: A Pharma Perspective. *Adv. Synth. Catal.* **2019**, *361*, 2421–2432.
- (15) Fryszkowska, A.; Devine, P. N. Biocatalysis in Drug Discovery and

- Development. *Curr. Opin. Chem. Biol.* **2020**, *55*, 151–160.
- (16) Panke, S.; Held, M.; Wubbolts, M. Trends and Innovations in Industrial Biocatalysis for the Production of Fine Chemicals. *Curr. Opin. Biotechnol.* **2004**, *15*, 272–279.
- (17) Woodley, J. M. Towards the Sustainable Production of Bulk-Chemicals Using Biotechnology. *New Biotechnol.* **2020**, *59*, 59–64.
- (18) Wu, S.; Snajdrova, R.; Moore, J. C.; Baldenius, K.; Bornscheuer, U. T. Biocatalysis: Enzymatic Synthesis for Industrial Applications. *Angew. Chem., Int. Ed.* **2021**, *60*, 88–119.
- (19) Abdelraheem, E. M. M.; Busch, H.; Hanefeld, U.; Tonin, F. Biocatalysis Explained: From Pharmaceutical to Bulk Chemical Production. *React. Chem. Eng.* **2019**, *4*, 1878–1894.
- (20) Jiao, S.; Li, F.; Yu, H.; Shen, Z. Advances in Acrylamide Bioproduction Catalyzed with *Rhodococcus* Cells Harboring Nitrile Hydratase. *Appl. Microbiol. Biotechnol.* **2020**, *104*, 1001–1012.
- (21) Davids, T.; Schmidt, M.; Böttcher, D.; Bornscheuer, U. T. Strategies for the Discovery and Engineering of Enzymes for Biocatalysis. *Curr. Opin. Chem. Biol.* **2013**, *17*, 215–220.
- (22) Balke, K.; Beier, A.; Bornscheuer, U. T. Hot Spots for the Protein Engineering of Baeyer-Villiger Monooxygenases. *Biotechnol. Adv.* **2018**, *36*, 247–263.
- (23) Denard, C. A.; Ren, H.; Zhao, H. Improving and Repurposing Biocatalysts via Directed Evolution. *Curr. Opin. Chem. Biol.* **2015**, *25*, 55–64.
- (24) Turner, N. J. Directed Evolution of Enzymes for Applied Biocatalysis. *Trends Biotechnol.* **2003**, *21*, 474–478.
- (25) Arnold, F. H.; Volkov, A. A. Directed Evolution of Biocatalysts. *Curr. Opin. Chem. Biol.* **1999**, *3*, 54–59.
- (26) Xiao, H.; Bao, Z.; Zhao, H. High Throughput Screening and Selection Methods for Directed Enzyme Evolution. *Ind. Eng. Chem. Res.* **2015**, *54*, 4011–4020.
- (27) Jaeger, K.-E.; Reetz, M. T. Directed Evolution of Enantioselective Enzymes for Organic Chemistry. *Curr. Opin. Chem. Biol.* **2000**, *4*, 68–73.
- (28) Steiner, K.; Schwab, H. Recent Advances in Rational Approaches for Enzyme Engineering. *Comput. Struct. Biotechnol. J.* **2012**, *2*, e201209010.
- (29) Chica, R. A.; Doucet, N.; Pelletier, J. N. Semi-Rational Approaches to Engineering Enzyme Activity: Combining the Benefits of Directed Evolution and Rational Design. *Curr. Opin. Biotechnol.* **2005**, *16*, 378–384.
- (30) Lutz, S. Beyond Directed Evolution—Semi-Rational Protein Engineering and Design. *Curr. Opin. Biotechnol.* **2010**, *21*, 734–743.

-
- (31) Reetz, M. T.; Kahakeaw, D.; Lohmer, R. Addressing the Numbers Problem in Directed Evolution. *ChemBioChem* **2008**, *9*, 1797–1804.
- (32) Reetz, M. T.; Bocola, M.; Carballeira, J. D.; Zha, D.; Vogel, A. Expanding the Range of Substrate Acceptance of Enzymes: Combinatorial Active-Site Saturation Test. *Angew. Chem., Int. Ed.* **2005**, *44*, 4192–4196.
- (33) Reetz, M. T.; Carballeira, J. D. Iterative Saturation Mutagenesis (ISM) for Rapid Directed Evolution of Functional Enzymes. *Nat. Protoc.* **2007**, *2*, 891–903.
- (34) Reetz, M. T.; Prasad, S.; Carballeira, J. D.; Gumulya, Y.; Bocola, M. Iterative Saturation Mutagenesis Accelerates Laboratory Evolution of Enzyme Stereoselectivity: Rigorous Comparison with Traditional Methods. *J. Am. Chem. Soc.* **2010**, *132*, 9144–9152.
- (35) Reetz, M. T.; Carballeira, J. D.; Peyralans, J.; Höbenreich, H.; Maichele, A.; Vogel, A. Expanding the Substrate Scope of Enzymes: Combining Mutations Obtained by CASTing. *Chem. Eur. J.* **2006**, *12*, 6031–6038.
- (36) Reetz, M. T. Laboratory Evolution of Stereoselective Enzymes: A Prolific Source of Catalysts for Asymmetric Reactions. *Angew. Chem., Int. Ed.* **2011**, *50*, 138–174.
- (37) Wells, A.; Meyer, H.-P. Biocatalysis as a Strategic Green Technology for the Chemical Industry. *ChemCatChem* **2014**, *6*, 918–920.
- (38) Sheldon, R. A. Biocatalysis and Biomass Conversion: Enabling a Circular Economy. *Phil. Trans. R. Soc. A* **2020**, *378*, 20190274.
- (39) van Schijndel, J.; Molendijk, D.; van Beurden, K.; Canalle, L. A.; Noël, T.; Meuldijk, J. Preparation of Bio-Based Styrene Alternatives and Their Free Radical Polymerization. *European Polymer Journal* **2020**, *125*, 109534.
- (40) Zhao, Y. Synthesis Method of *p*-Hydroxystyrene from Benzaldehyde and Malonic Acid via Condensation and Decarboxylation. CN108658731A, 2018.
- (41) Takeshima, H.; Satoh, K.; Kamigaito, M. Bio-based Vinylphenol Family: Synthesis via Decarboxylation of Naturally Occurring Cinnamic Acids and Living Radical Polymerization for Functionalized Polystyrenes. *J. Polym. Sci.* **2020**, *58*, 91–100.
- (42) Gaßmeyer, S. K.; Wetzig, J.; Mügge, C.; Assmann, M.; Enoki, J.; Hilterhaus, L.; Zuhse, R.; Miyamoto, K.; Liese, A.; Kourist, R. Arylmalonate Decarboxylase-Catalyzed Asymmetric Synthesis of Both Enantiomers of Optically Pure Flurbiprofen. *ChemCatChem* **2016**, *8*, 916–921.
- (43) Miyamoto, K.; Kourist, R. Arylmalonate Decarboxylase—a Highly Selective Bacterial Biocatalyst with Unknown Function. *Appl. Microbiol. Biotechnol.* **2016**, *100*, 8621–8631.
-

-
- (44) Enoki, J.; Linhorst, M.; Busch, F.; Baraibar, Á. G.; Miyamoto, K.; Kourist, R.; Mügge, C. Preparation of Optically Pure Flurbiprofen via an Integrated Chemo-Enzymatic Synthesis Pathway. *Mol. Catal.* **2019**, *467*, 135–142.
- (45) Blakemore, C. A.; France, S. P.; Samp, L.; Nason, D. M.; Yang, E.; Howard, R. M.; Coffman, K. J.; Yang, Q.; Smith, A. C.; Evrard, E.; Li, W.; Dai, L.; Yang, L.; Chen, Z.; Zhang, Q.; He, F.; Zhang, J. Scalable, Telescoped Hydrogenolysis–Enzymatic Decarboxylation Process for the Asymmetric Synthesis of (*R*)- α -Heteroaryl Propionic Acids. *Org. Process Res. Dev.* **2020**, in press.
- (46) Enoki, J.; Mügge, C.; Tischler, D.; Miyamoto, K.; Kourist, R. Chemoenzymatic Cascade Synthesis of Optically Pure Alkanoic Acids by Using Engineered Arylmalonate Decarboxylase Variants. *Chem. Eur. J.* **2019**, *25*, 5071–5076.
- (47) Peters, C.; Buller, R. Industrial Application of 2-Oxoglutarate-Dependent Oxygenases. *Catalysts* **2019**, *9*, 221.
- (48) Zwick, C. R.; Renata, H. Harnessing the Biocatalytic Potential of Iron- and α -Ketoglutarate-Dependent Dioxygenases in Natural Product Total Synthesis. *Nat. Prod. Rep.* **2020**, *37*, 1065–1079.
- (49) Enoki, J.; Meisborn, J.; Müller, A. C.; Kourist, R. A Multi-Enzymatic Cascade Reaction for the Stereoselective Production of γ -Oxyfunctionalized Amino Acids. *Front. Microbiol.* **2016**, *7*, 1–8.
- (50) Theodosiou, E.; Breisch, M.; Julsing, M. K.; Falcioni, F.; Bühler, B.; Schmid, A. An Artificial TCA Cycle Selects for Efficient α -Ketoglutarate Dependent Hydroxylase Catalysis in Engineered *Escherichia Coli*. *Biotechnol. Bioeng.* **2017**, *114*, 1511–1520.
- (51) Deng, Y.; Lu, S. Biosynthesis and Regulation of Phenylpropanoids in Plants. *Crit. Rev. Plant Sci.* **2017**, *36*, 257–290.
- (52) Barros, J.; Serrani-Yarce, J. C.; Chen, F.; Baxter, D.; Venables, B. J.; Dixon, R. A. Role of Bifunctional Ammonia-Lyase in Grass Cell Wall Biosynthesis. *Nat. Plants* **2016**, *2*, 16050.
- (53) Chen, H.-C.; Li, Q.; Shuford, C. M.; Liu, J.; Muddiman, D. C.; Sederoff, R. R.; Chiang, V. L. Membrane Protein Complexes Catalyze Both 4- and 3-Hydroxylation of Cinnamic Acid Derivatives in Monolignol Biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 21253–21258.
- (54) Yoon, J.; Choi, H.; An, G. Roles of Lignin Biosynthesis and Regulatory Genes in Plant Development. *J. Integr. Plant Biol.* **2015**, *57*, 902–912.
- (55) Chen, F.; Srinivasa Reddy, M. S.; Temple, S.; Jackson, L.; Shadle, G.; Dixon, R. A. Multi-Site Genetic Modulation of Monolignol Biosynthesis Suggests New Routes for Formation of Syringyl Lignin and Wall-Bound Ferulic Acid in Alfalfa
-

- (*Medicago Sativa* L.). *Plant J.* **2006**, *48*, 113–124.
- (56) Nair, R. B.; Bastress, K. L.; Ruegger, M. O.; Denault, J. W.; Chapple, C. The *Arabidopsis Thaliana* REDUCED EPIDERMAL FLUORESCENCE1 Gene Encodes an Aldehyde Dehydrogenase Involved in Ferulic Acid and Sinapic Acid Biosynthesis. *Plant Cell* **2004**, *16*, 544–554.
- (57) Betts, W. B.; Dart, R. K.; Ball, A. S.; Pedlar, S. L. Biosynthesis and Structure of Lignocellulose. In *Biodegradation*; Betts, W. B., Ed.; Springer: London, 1991; pp 139–155.
- (58) Bugg, T. D. H.; Ahmad, M.; Hardiman, E. M.; Rahmanpour, R. Pathways for Degradation of Lignin in Bacteria and Fungi. *Nat. Prod. Rep.* **2011**, *28*, 1883.
- (59) de Oliveira, D. M.; Finger-Teixeira, A.; Rodrigues Mota, T.; Salvador, V. H.; Moreira-Vilar, F. C.; Correa Molinari, H. B.; Craig Mitchell, R. A.; Marchiosi, R.; Ferrarese-Filho, O.; Dantas dos Santos, W. Ferulic Acid: A Key Component in Grass Lignocellulose Recalcitrance to Hydrolysis. *Plant Biotechnol. J.* **2015**, *13*, 1224–1232.
- (60) Malherbe, S.; Cloete, T. E. Lignocellulose Biodegradation: Fundamentals and Applications. *Rev. Environ. Sci. Bio/Technol.* **2002**, *1*, 105–114.
- (61) Williamson, G.; Kroon, P. A.; Faulds, C. B. Hairy Plant Polysaccharides: A Close Shave with Microbial Esterases. *Microbiology* **1998**, *144*, 2011–2023.
- (62) Besle, J.-M.; Cornu, A.; Jouany, J.-P. Roles of Structural Phenylpropanoids in Forage Cell Wall Digestion. *J. Sci. Food Agric.* **1994**, *64*, 171–190.
- (63) Mnich, E.; Bjarnholt, N.; Eudes, A.; Harholt, J.; Holland, C.; Jørgensen, B.; Larsen, F. H.; Liu, M.; Manat, R.; Meyer, A. S.; Mikkelsen, J. D.; Motawia, M. S.; Muschiol, J.; Møller, B. L.; Møller, S. R.; Perzon, A.; Petersen, B. L.; Ravn, J. L.; Ulvskov, P. Phenolic Cross-Links: Building and de-Constructing the Plant Cell Wall. *Nat. Prod. Rep.* **2020**, *37*, 919–961.
- (64) Mathew, S.; Abraham, T. E. Ferulic Acid: An Antioxidant Found Naturally in Plant Cell Walls and Feruloyl Esterases Involved in Its Release and Their Applications. *Crit. Rev. Biotechnol.* **2004**, *24*, 59–83.
- (65) Clausen, M.; Lamb, C. J.; Megnet, R.; Doerner, P. W. PAD1 Encodes Phenylacrylic Acid Decarboxylase Which Confers Resistance to Cinnamic Acid in *Saccharomyces Cerevisiae*. *Gene* **1994**, *142*, 107–112.
- (66) Baranowski, J. D.; Davidson, P. M.; Nagel, C. W.; Branen, A. L. Inhibition of *Saccharomyces Cerevisiae* by Naturally Occuring Hydroxycinnamates. *J. Food Sci.* **1980**, *45*, 592–594.
- (67) Van Dyk, T. K.; Templeton, L. J.; Cantera, K. A.; Sharpe, P. L.; Sariaslani, F. S. Characterization of the *Escherichia Coli* AaeAB Efflux Pump: A Metabolic Relief

- Valve? *J. Bacteriol.* **2004**, *186*, 7196–7204.
- (68) Larsson, S.; Nilvebrant, N.-O.; Jönsson, L. Effect of Overexpression of *Saccharomyces Cerevisiae* Pad1p on the Resistance to Phenylacrylic Acids and Lignocellulose Hydrolysates under Aerobic and Oxygen-Limited Conditions. *Appl. Microbiol. Biotechnol.* **2001**, *57*, 167–174.
- (69) Kalogeraki, V. S.; Zhu, J.; Eberhard, A.; Madsen, E. L.; Winans, S. C. The Phenolic Vir Gene Inducer Ferulic Acid Is O-Demethylated by the VirH2 Protein of an *Agrobacterium Tumefaciens* Ti Plasmid. *Mol. Microbiol.* **1999**, *34*, 512–522.
- (70) Rosazza, J. P. N.; Huang, Z.; Dostal, L.; Volm, T.; Rousseau, B. Review: Biocatalytic Transformations of Ferulic Acid: An Abundant Aromatic Natural Product. *J. Ind. Microbiol.* **1995**, *15*, 457–471.
- (71) Tinikul, R.; Chenprakhon, P.; Maenpuen, S.; Chaiyen, P. Biotransformation of Plant-Derived Phenolic Acids. *Biotechnol. J.* **2018**, *13*, 1700632.
- (72) Tran, N. P.; Gury, J.; Dartois, V.; Nguyen, T. K. C.; Seraut, H.; Barthelmebs, L.; Gervais, P.; Cavin, J.-F. Phenolic Acid-Mediated Regulation of the PadC Gene, Encoding the Phenolic Acid Decarboxylase of *Bacillus Subtilis*. *J. Bacteriol.* **2008**, *190*, 3213–3224.
- (73) Nguyen, T. K. C.; Tran, N. P.; Cavin, J.-F. Genetic and Biochemical Analysis of PadR-PadC Promoter Interactions during the Phenolic Acid Stress Response in *Bacillus Subtilis* 168. *J. Bacteriol.* **2011**, *193*, 4180–4191.
- (74) Gury, J.; Barthelmebs, L.; Tran, N. P.; Diviès, C.; Cavin, J.-F. Cloning, Deletion, and Characterization of PadR, the Transcriptional Repressor of the Phenolic Acid Decarboxylase-Encoding PadA Gene of *Lactobacillus Plantarum*. *Appl. Environ. Microbiol.* **2004**, *70*, 2146–2153.
- (75) Barthelmebs, L.; Lecomte, B.; Divies, C.; Cavin, J.-F. Inducible Metabolism of Phenolic Acids in *Pediococcus Pentosaceus* Is Encoded by an Autoregulated Operon Which Involves a New Class of Negative Transcriptional Regulator. *J. Bacteriol.* **2000**, *182*, 6724–6731.
- (76) Campos, F. M.; Couto, J. A.; Figueiredo, A. R.; Tóth, I. V.; Rangel, A. O. S. S.; Hogg, T. A. Cell Membrane Damage Induced by Phenolic Acids on Wine Lactic Acid Bacteria. *Int. J. Food Microbiol.* **2009**, *135*, 144–151.
- (77) Heipieper, H. J.; Keweloh, H.; Rehm, H. J. Influence of Phenols on Growth and Membrane Permeability of Free and Immobilized *Escherichia Coli*. *Appl. Environ. Microbiol.* **1991**, *57*, 1213–1217.
- (78) Licandro-Seraut, H.; Roussel, C.; Perpetuini, G.; Gervais, P.; Cavin, J.-F. Sensitivity to Vinyl Phenol Derivatives Produced by Phenolic Acid

- Decarboxylase Activity in *Escherichia Coli* and Several Food-Borne Gram-Negative Species. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 7853–7864.
- (79) Payer, S. E.; Faber, K.; Glueck, S. M. Non-Oxidative Enzymatic (De)Carboxylation of (Hetero)Aromatics and Acrylic Acid Derivatives. *Adv. Synth. Catal.* **2019**, *361*, 2402–2420.
- (80) Lentz, M. The Impact of Simple Phenolic Compounds on Beer Aroma and Flavor. *Fermentation* **2018**, *4*, 20.
- (81) Chatonnet, P.; Dubourdieu, D.; Boidron, J.; Lavigne, V. Synthesis of Volatile Phenols by *Saccharomyces Cerevisiae* in Wines. *J. Sci. Food Agric.* **1993**, *62*, 191–202.
- (82) Hashidoko, Y.; Tahara, S. Stereochemically Specific Proton Transfer in Decarboxylation of 4-Hydroxycinnamic Acids by 4-Hydroxycinnamate Decarboxylase from *Klebsiella Oxytoca*. *Arch. Biochem. Biophys.* **1998**, *359*, 225–230.
- (83) Rodríguez, H.; Angulo, I.; de las Rivas, B.; Campillo, N.; Páez, J. A.; Muñoz, R.; Mancheño, J. M. *P*-Coumaric Acid Decarboxylase from *Lactobacillus Plantarum*: Structural Insights into the Active Site and Decarboxylation Catalytic Mechanism. *Proteins: Struct., Funct., Bioinf.* **2010**, *78*, 1662–1676.
- (84) Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, J. D.; Higgins, D. G. Fast, Scalable Generation of High-quality Protein Multiple Sequence Alignments Using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539.
- (85) Matte, A.; Grosse, S.; Bergeron, H.; Abokitse, K.; Lau, P. C. K. Structural Analysis of *Bacillus Pumilus* Phenolic Acid Decarboxylase, a Lipocalin-Fold Enzyme. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2010**, *66*, 1407–1414.
- (86) Frank, A.; Eborall, W.; Hyde, R.; Hart, S.; Turkenburg, J. P.; Grogan, G. Mutational Analysis of Phenolic Acid Decarboxylase from *Bacillus Subtilis* (BsPAD), Which Converts Bio-Derived Phenolic Acids to Styrene Derivatives. *Catal. Sci. Technol.* **2012**, *2*, 1568–1574.
- (87) Sheng, X.; Lind, M. E. S.; Himo, F. Theoretical Study of the Reaction Mechanism of Phenolic Acid Decarboxylase. *FEBS J.* **2015**, *282*, 4703–4713.
- (88) The PyMOL Molecular Graphics System, Version 2.2. Schrödinger LLC.
- (89) Wuensch, C.; Pavkov-Keller, T.; Steinkellner, G.; Gross, J.; Fuchs, M.; Hromic, A.; Lyskowski, A.; Fauland, K.; Gruber, K.; Glueck, S. M.; Faber, K. Regioselective Enzymatic β -Carboxylation of *Para*-Hydroxy-Styrene Derivatives Catalyzed by Phenolic Acid Decarboxylases. *Adv. Synth. Catal.* **2015**, *357*,

- 1909–1918.
- (90) Wuensch, C.; Gross, J.; Steinkellner, G.; Gruber, K.; Glueck, S. M.; Faber, K. Asymmetric Enzymatic Hydration of Hydroxystyrene Derivatives. *Angew. Chem., Int. Ed.* **2013**, *52*, 2293–2297.
- (91) Sheng, X.; Himo, F. Theoretical Study of Enzyme Promiscuity: Mechanisms of Hydration and Carboxylation Activities of Phenolic Acid Decarboxylase. *ACS Catal.* **2017**, *7*, 1733–1741.
- (92) Fraaije, M. W.; van Berkel, W. J. H. Catalytic Mechanism of the Oxidative Demethylation of 4-(Methoxymethyl)Phenol by Vanillyl-Alcohol Oxidase. *J. Biol. Chem.* **1997**, *272*, 18111–18116.
- (93) Bennett, J. P.; Bertin, L.; Moulton, B.; Fairlamb, I. J. S.; Brzozowski, A. M.; Walton, N. J.; Grogan, G. A Ternary Complex of Hydroxycinnamoyl-CoA Hydratase–Lyase (HCHL) with Acetyl-CoA and Vanillin Gives Insights into Substrate Specificity and Mechanism. *Biochem. J.* **2008**, *414*, 281–289.
- (94) Wuensch, C.; Glueck, S. M.; Gross, J.; Koszelewski, D.; Schober, M.; Faber, K. Regioselective Enzymatic Carboxylation of Phenols and Hydroxystyrene Derivatives. *Org. Lett.* **2012**, *14*, 1974–1977.
- (95) Kourist, R.; Schweiger, A.; Büchenschütz, H. Enzymatic Decarboxylation as a Tool for the Enzymatic Defunctionalization Organic Acids. In *Lipid Modification by Enzymes and Engineered Microbes*; Bornscheuer, U. T., Ed.; Elsevier, 2018; pp 89–118.
- (96) Maeda, H.; Dudareva, N. The Shikimate Pathway and Aromatic Amino Acid Biosynthesis in Plants. *Annu. Rev. Plant Biol.* **2012**, *63*, 73–105.
- (97) Qi, W.; Vannellil, T.; Breinig, S.; Benbassat, A.; Gatenby, A.; Haynie, S.; Sariaslani, F. Functional Expression of Prokaryotic and Eukaryotic Genes in *Escherichia Coli* for Conversion of Glucose to *p*-Hydroxystyrene. *Metab. Eng.* **2007**, *9*, 268–276.
- (98) Ben-Bassat, A.; Lowe, D. J. A Method for Producing *Para*-Hydroxystyrene and Other Multifunctional Aromatic Compounds Usind Two-Phase Extractive Fermentation. WO 2004/092392, 2004.
- (99) Verhoef, S.; Wierckx, N.; Westerhof, R. G. M.; de Winde, J. H.; Ruijssenaars, H. J. Bioproduction of *p*-Hydroxystyrene from Glucose by the Solvent-Tolerant Bacterium *Pseudomonas Putida* S12 in a Two-Phase Water-Decanol Fermentation. *Appl. Environ. Microbiol.* **2009**, *75*, 931–936.
- (100) Busto, E.; Simon, R. C.; Kroutil, W. Vinylation of Unprotected Phenols Using a Biocatalytic System. *Angew. Chem., Int. Ed.* **2015**, *54*, 10899–10902.
- (101) Gómez Baraibar, Á.; Reichert, D.; Mügge, C.; Seger, S.; Gröger, H.; Kourist, R.

- A One-Pot Cascade Reaction Combining an Encapsulated Decarboxylase with a Metathesis Catalyst for the Synthesis of Bio-Based Antioxidants. *Angew. Chem., Int. Ed.* **2016**, *55*, 14823–14827.
- (102) Schweiger, A. K.; Ríos-Lombardía, N.; Winkler, C. K.; Schmidt, S.; Morís, F.; Kroutil, W.; González-Sabín, J.; Kourist, R. Using Deep Eutectic Solvents to Overcome Limited Substrate Solubility in the Enzymatic Decarboxylation of Bio-Based Phenolic Acids. *ACS Sustainable Chem. Eng.* **2019**, *7*, 16364–16370.
- (103) Ríos-Lombardía, N.; Rodríguez-Álvarez, M. J.; Morís, F.; Kourist, R.; Comino, N.; López-Gallego, F.; González-Sabín, J.; García-Álvarez, J. DESign of Sustainable One-Pot Chemoenzymatic Organic Transformations in Deep Eutectic Solvents for the Synthesis of 1,2-Disubstituted Aromatic Olefins. *Front. Chem.* **2020**, *8*, 1–11.
- (104) Pesci, L.; Baydar, M.; Glueck, S.; Faber, K.; Liese, A.; Kara, S. Development and Scaling-Up of the Fragrance Compound 4-Ethylguaiacol Synthesis via a Two-Step Chemo-Enzymatic Reaction Sequence. *Org. Process Res. Dev.* **2017**, *21*, 85–93.
- (105) Morley, K. L.; Grosse, S.; Leisch, H.; Lau, P. C. K. Antioxidant Canolol Production from a Renewable Feedstock via an Engineered Decarboxylase. *Green Chem.* **2013**, *15*, 3312.
- (106) Rosano, G. L.; Morales, E. S.; Ceccarelli, E. A. New Tools for Recombinant Protein Production in *Escherichia Coli*: A 5-year Update. *Protein Sci.* **2019**, *28*, 1412–1422.
- (107) Shilling, P. J.; Mirzadeh, K.; Cumming, A. J.; Widesheim, M.; Köck, Z.; Daley, D. O. Improved Designs for PET Expression Plasmids Increase Protein Production Yield in *Escherichia Coli*. *Commun. Biol.* **2020**, *3*, 214.
- (108) Haddadin, F. T.; Harcum, S. W. Transcriptome Profiles for High-Cell-Density Recombinant and Wild-Type *Escherichia Coli*. *Biotechnol. Bioeng.* **2005**, *90*, 127–153.
- (109) Wurm, D. J.; Quehenberger, J.; Mildner, J.; Eggenreich, B.; Slouka, C.; Schwaighofer, A.; Wieland, K.; Lendl, B.; Rajamanickam, V.; Herwig, C.; Spadiut, O. Teaching an Old PET New Tricks: Tuning of Inclusion Body Formation and Properties by a Mixed Feed System in *E. Coli*. *Appl. Microbiol. Biotechnol.* **2018**, *102*, 667–676.
- (110) Cavin, J.-F.; Dartois, V.; Diviès, C. Gene Cloning, Transcriptional Analysis, Purification, and Characterization of Phenolic Acid Decarboxylase from *Bacillus Subtilis*. *Appl. Environ. Microbiol.* **1998**, *64*, 1466–1471.
- (111) Baneyx, F. Recombinant Protein Expression in *Escherichia Coli*. *Curr. Opin.*

-
- Biotechnol.* **1999**, *10*, 411–421.
- (112) Morley, K. L.; Grosse, S.; Leisch, H.; Lau, P. C. K. Antioxidant Canolol Production from a Renewable Feedstock via an Engineered Decarboxylase. *Green Chem.* **2013**, *15*, 3312–3317.
- (113) Leisch, H.; Grosse, S.; Morley, K.; Abokitse, K.; Perrin, F.; Denault, J.; Lau, P. C. K. Chemicals from Agricultural Biomass: Chemoenzymatic Approach for Production of Vinylphenols and Polyvinylphenols from Phenolic Acids. *Green Process. Synth.* **2013**, *2*, 7–17.
- (114) Grabner, B.; Schweiger, A. K.; Gavric, K.; Kourist, R.; Gruber-Woelfler, H. A Chemo-Enzymatic Tandem Reaction in a Mixture of Deep Eutectic Solvent and Water in Continuous Flow. *React. Chem. Eng.* **2020**, *5*, 263–269.
- (115) EnginZyme. EziG™ Product Data Sheet <http://enginzyme.com/wp-content/uploads/2017/03/EziG™-Product-Data-Sheet.pdf> (accessed Nov 27, 2020).
- (116) Engelmark Cassimjee, K.; Kadow, M.; Wikmark, Y.; Svedendahl Humble, M.; Rothstein, M. L.; Rothstein, D. M.; Bäckvall, J.-E. A General Protein Purification and Immobilization Method on Controlled Porosity Glass: Biocatalytic Applications. *Chem. Commun.* **2014**, *50*, 9134–9137.
- (117) Thompson, M. P.; Derrington, S. R.; Heath, R. S.; Porter, J. L.; Mangas-Sanchez, J.; Devine, P. N.; Truppo, M. D.; Turner, N. J. A Generic Platform for the Immobilisation of Engineered Biocatalysts. *Tetrahedron* **2019**, *75*, 327–334.
- (118) Zhang, Q.; De Oliveira Vigier, K.; Royer, S.; Jérôme, F. Deep Eutectic Solvents: Syntheses, Properties and Applications. *Chem. Soc. Rev.* **2012**, *41*, 7108–7146.
- (119) Gotor-Fernández, V.; Paul, C. E. Deep Eutectic Solvents for Redox Biocatalysis. *J. Biotechnol.* **2019**, *293*, 24–35.
- (120) Sarmad, S.; Mikkola, J.-P.; Ji, X. Carbon Dioxide Capture with Ionic Liquids and Deep Eutectic Solvents: A New Generation of Sorbents. *ChemSusChem* **2017**, *10*, 324–352.
- (121) Hammond, O. S.; Bowron, D. T.; Edler, K. J. The Effect of Water upon Deep Eutectic Solvent Nanostructure: An Unusual Transition from Ionic Mixture to Aqueous Solution. *Angew. Chem., Int. Ed.* **2017**, *56*, 9782–9785.
- (122) Ferrario, V.; Hansen, N.; Pleiss, J. Interpretation of Cytochrome P450 Monooxygenase Kinetics by Modeling of Thermodynamic Activity. *J. Inorg. Biochem.* **2018**, *183*, 172–178.
- (123) Pleiss, J. Thermodynamic Activity-Based Interpretation of Enzyme Kinetics. *Trends Biotechnol.* **2017**, *35*, 379–382.
- (124) Pleiss, J. Thermodynamic Activity-Based Progress Curve Analysis in Enzyme
-

- Kinetics. *Trends Biotechnol.* **2018**, *36*, 234–238.
- (125) Ruß, C.; König, B. Low Melting Mixtures in Organic Synthesis – an Alternative to Ionic Liquids? *Green Chem.* **2012**, *14*, 2969–2982.
- (126) Rudroff, F.; Mihovilovic, M. D.; Gröger, H.; Snajdrova, R.; Iding, H.; Bornscheuer, U. T. Opportunities and Challenges for Combining Chemo- and Biocatalysis. *Nat. Catal.* **2018**, *1*, 12–22.
- (127) Nawaz, W.; Zhou, Z.; Deng, S.; Ma, X.; Ma, X.; Li, C.; Shu, X. Therapeutic Versatility of Resveratrol Derivatives. *Nutrients* **2017**, *9*, 1188.
- (128) Tsai, H.-Y.; Ho, C.-T.; Chen, Y.-K. Biological Actions and Molecular Effects of Resveratrol, Pterostilbene, and 3'-Hydroxypterostilbene. *J. Food Drug Anal.* **2017**, *25*, 134–147.
- (129) Lackner, F.; Hiebler, K.; Grabner, B.; Gruber-Woelfler, H. Optimization of a Catalytic Chemoenzymatic Tandem Reaction for the Synthesis of Natural Stilbenes in Continuous Flow. *Catalysts* **2020**, *10*, 1404.
- (130) Miyamoto, K.; Ohta, H. Enzyme-Mediated Asymmetric Decarboxylation of Disubstituted Malonic Acids. *J. Am. Chem. Soc.* **1990**, *112*, 4077–4078.
- (131) Miyamoto, K.; Ohta, H. Purification and Properties of a Novel Arylmalonate Decarboxylase from *Alcaligenes Bronchisepticus* KU 1201. *Eur. J. Biochem.* **1992**, *210*, 475–481.
- (132) Miyamoto, K.; Tsuchiya, S.; Ohta, H. Stereochemistry of Enzyme-Catalyzed Decarboxylation of Alpha-Methyl-Alpha-Phenylmalonic Acid. *J. Am. Chem. Soc.* **1992**, *114*, 6256–6257.
- (133) Okrasa, K.; Levy, C.; Hauer, B.; Baudendistel, N.; Leys, D.; Micklefield, J. Structure and Mechanism of an Unusual Malonate Decarboxylase and Related Racemases. *Chem. Eur. J.* **2008**, *14*, 6609–6613.
- (134) Kim, Y. S.; Kolattukudy, P. E. Stereospecificity of Malonyl-CoA Decarboxylase, Acetyl-CoA Carboxylase, and Fatty Acid Synthetase from the Uropygial Gland of Goose. *J. Biol. Chem.* **1980**, *255*, 686–689.
- (135) Miyamoto, K.; Ohta, H. Cloning and Heterologous Expression of a Novel Arylmalonate Decarboxylase Gene from *Alcaligenes Bronchisepticus* KU 1201. *Appl. Microbiol. Biotechnol.* **1992**, *38*, 234–238.
- (136) Miyazaki, M.; Kakidani, H.; Hanzawa, S.; Ohta, H. Cysteine188 Revealed as Being Critical for the Enzyme Activity of Arylmalonate Decarboxylase by Site-Directed Mutagenesis. *Bull. Chem. Soc. Jpn.* **1997**, *70*, 2765–2769.
- (137) Matoishi, K.; Ueda, M.; Miyamoto, K.; Ohta, H. Mechanism of Asymmetric Decarboxylation of α -Aryl- α -Methylmalonate Catalyzed by Arylmalonate Decarboxylase Originated from *Alcaligenes Bronchisepticus*. *J. Mol. Catal. B:*

-
- Enzym.* **2004**, *27*, 161–168.
- (138) Glavas, S.; Tanner, M. E. Catalytic Acid/Base Residues of Glutamate Racemase. *Biochemistry* **1999**, *38*, 4106–4113.
- (139) Ijima, Y.; Matoishi, K.; Terao, Y.; Doi, N.; Yanagawa, H.; Ohta, H. Inversion of Enantioselectivity of Asymmetric Biocatalytic Decarboxylation by Site-Directed Mutagenesis Based on the Reaction Mechanism. *Chem. Commun.* **2005**, 877–879.
- (140) Terao, Y.; Miyamoto, K.; Ohta, H. Improvement of the Activity of Arylmalonate Decarboxylase by Random Mutagenesis. *Appl. Microbiol. Biotechnol.* **2006**, *73*, 647–653.
- (141) Miyamoto, K.; Tsutsumi, T.; Terao, Y.; Ohta, H. Stereochemistry of Decarboxylation of Arylmalonate Catalyzed by Mutant Enzymes. *Chem. Lett.* **2007**, *36*, 656–657.
- (142) Terao, Y.; Miyamoto, K.; Ohta, H. Introduction of Single Mutation Changes Arylmalonate Decarboxylase to Racemase. *Chem. Commun.* **2006**, 3600–3602.
- (143) Terao, Y.; Ijima, Y.; Miyamoto, K.; Ohta, H. Inversion of Enantioselectivity of Arylmalonate Decarboxylase via Site-Directed Mutation Based on the Proposed Reaction Mechanism. *J. Mol. Catal. B: Enzym.* **2007**, *45*, 15–20.
- (144) Gerlt, J. A.; Babbitt, P. C.; Rayment, I. Divergent Evolution in the Enolase Superfamily: The Interplay of Mechanism and Specificity. *Arch. Biochem. Biophys.* **2005**, *433*, 59–70.
- (145) Okrasa, K.; Levy, C.; Wilding, M.; Goodall, M.; Baudendistel, N.; Hauer, B.; Leys, D.; Micklefield, J. Structure-Guided Directed Evolution of Alkenyl and Arylmalonate Decarboxylases. *Angew. Chem., Int. Ed.* **2009**, *48*, 7691–7694.
- (146) Obata, R.; Nakasako, M. Structural Basis for Inverting the Enantioselectivity of Arylmalonate Decarboxylase Revealed by the Structural Analysis of the Gly74Cys/Cys188Ser Mutant in the Liganded Form. *Biochemistry* **2010**, *49*, 1963–1969.
- (147) Puig, E.; Mixcoha, E.; Garcia-Viloca, M.; González-Lafont, A.; Lluch, J. M. How the Substrate D-Glutamate Drives the Catalytic Action of *Bacillus Subtilis* Glutamate Racemase. *J. Am. Chem. Soc.* **2009**, *131*, 3509–3521.
- (148) Miyauchi, Y.; Kourist, R.; Uemura, D.; Miyamoto, K. Dramatically Improved Catalytic Activity of an Artificial (S)-Selective Arylmalonate Decarboxylase by Structure-Guided Directed Evolution. *Chem. Commun.* **2011**, *47*, 7503.
- (149) Lind, M. E. S.; Himo, F. Theoretical Study of Reaction Mechanism and Stereoselectivity of Arylmalonate Decarboxylase. *ACS Catal.* **2014**, *4*, 4153–4160.
-

-
- (150) Dasgupta, S.; Herbert, J. M. Using Atomic Confining Potentials for Geometry Optimization and Vibrational Frequency Calculations in Quantum-Chemical Models of Enzyme Active Sites. *J. Phys. Chem. B* **2020**, *124*, 1137–1147.
- (151) Busch, F.; Enoki, J.; Hülsemann, N.; Miyamoto, K.; Bocola, M.; Kourist, R. Semiempirical QM/MM Calculations Reveal a Step-Wise Proton Transfer and an Unusual Thiolate Pocket in the Mechanism of the Unique Arylpropionate Racemase AMDase G74C. *Catal. Sci. Technol.* **2016**, *6*, 4937–4944.
- (152) Glavas, S.; Tanner, M. E. Active Site Residues of Glutamate Racemase. *Biochemistry* **2001**, *40*, 6199–6204.
- (153) Karmakar, T.; Balasubramanian, S. Molecular Dynamics and Free Energy Simulations of Phenylacetate and CO₂ Release from AMDase and Its G74C/C188S Mutant: A Possible Rationale for the Reduced Activity of the Latter. *J. Phys. Chem. B* **2016**, *120*, 11644–11653.
- (154) Miyamoto, K.; Ohta, H.; Osamura, Y. Effect of Conformation of the Substrate on Enzymatic Decarboxylation of α -Arylmalonic Acid. *Bioorg. Med. Chem.* **1994**, *2*, 469–475.
- (155) Tamura, K.; Terao, Y.; Miyamoto, K.; Ohta, H. Asymmetric Decarboxylation of α -Hydroxy- and α -Amino- α -Phenylmalonate Catalyzed by Arylmalonate Decarboxylase from *Alcaligenes Bronchisepticus*. *Biocatal. Biotransform.* **2008**, *26*, 253–257.
- (156) Yoshida, S.; Enoki, J.; Kourist, R.; Miyamoto, K. Engineered Hydrophobic Pocket of (S)-Selective Arylmalonate Decarboxylase Variant by Simultaneous Saturation Mutagenesis to Improve Catalytic Performance. *Biosci., Biotechnol., Biochem.* **2015**, *79*, 1965–1971.
- (157) Kourist, R.; Miyauchi, Y.; Uemura, D.; Miyamoto, K. Engineering the Promiscuous Racemase Activity of an Arylmalonate Decarboxylase. *Chem. Eur. J.* **2011**, *17*, 557–563.
- (158) Miyamoto, K.; Tsuchiya, S.; Ohta, H. Microbial Asymmetric Decarboxylation of Fluorine-Containing Arylmalonic Acid Derivatives. *J. Fluor. Chem.* **1992**, *59*, 225–232.
- (159) Lewin, R.; Goodall, M.; Thompson, M. L.; Leigh, J.; Breuer, M.; Baldenius, K.; Micklefield, J. Enzymatic Enantioselective Decarboxylative Protonation of Heteroaryl Malonates. *Chem. Eur. J.* **2015**, *21*, 6557–6563.
- (160) Fukuyama, Y.; Matoishi, K.; Iwasaki, M.; Takizawa, E.; Miyazaki, M.; Ohta, H.; Hanzawa, S.; Kakidani, H.; Sugai, T. Preparative-Scale Enzyme-Catalyzed Synthesis of (R)- α -Fluorophenylacetic Acid. *Biosci., Biotechnol., Biochem.* **1999**, *63*, 1664–1666.
-

-
- (161) Matoishi, K.; Kakidani, H.; Suzuki, M.; Sugai, T.; Ohta, H.; Hanzawa, S. The First Synthesis of Both Enantiomers of [α - 2 H]Phenylacetic Acid in High Enantiomeric Excess. *Chem. Commun.* **2000**, 1519–1520.
- (162) Kawasaki, T.; Horimai, E.; Ohta, H. On the Conformation of the Substrate Binding to the Active Site during the Course of Enzymatic Decarboxylation. *Bull. Chem. Soc. Jpn.* **1996**, *69*, 3591–3594.
- (163) Nishio, M.; Umezawa, Y.; Hirota, M.; Takeuchi, Y. The CH/ π Interaction: Significance in Molecular Recognition. *Tetrahedron* **1995**, *51*, 8665–8701.
- (164) Kourist, R.; Domínguez de María, P.; Miyamoto, K. Biocatalytic Strategies for the Asymmetric Synthesis of Profens – Recent Trends and Developments. *Green Chem.* **2011**, *13*, 2607–2618.
- (165) Gaßmeyer, S. K.; Yoshikawa, H.; Enoki, J.; Hülsemann, N.; Stoll, R.; Miyamoto, K.; Kourist, R. STD-NMR-Based Protein Engineering of the Unique Arylpropionate-Racemase AMDase G74C. *ChemBioChem* **2015**, *16*, 1943–1949.
- (166) Brogden, R. N. Non-Steroidal Anti-Inflammatory Analgesics Other than Salicylates. *Drugs* **1986**, *32*, 27–45.
- (167) Brooks, P. Use and Benefits of Nonsteroidal Anti-Inflammatory Drugs. *Am. J. Med.* **1998**, *104*, 9S-13S.
- (168) Tracy, T. S.; Hall, S. D. Metabolic Inversion of (*R*)-Ibuprofen. Epimerization and Hydrolysis of Ibuprofenyl-Coenzyme A. *Drug Metab. Dispos.* **1992**, *20*, 322–327.
- (169) Abdel-Aziz, A. A.-M.; Al-Badr, A. A.; Hafez, G. A. Flurbiprofen. In *Profiles of Drug Substances, Excipients and Related Methodology*; Brittain, H. G., Ed.; Elsevier, 2012; pp 113–181.
- (170) Liu, J. K.; Patel, S. K.; Gillespie, D. L.; Whang, K.; Couldwell, W. T. *R*-Flurbiprofen, a Novel Nonsteroidal Anti-Inflammatory Drug, Decreases Cell Proliferation and Induces Apoptosis in Pituitary Adenoma Cells in Vitro. *J. Neuro-Oncol.* **2012**, *106*, 561–569.
- (171) Jin, H.; Wang, Z.; Liu, L.; Gao, L.; Sun, L.; Li, X.; Zhao, H.; Pan, Y.; Shi, H.; Liu, N.; Hong, L.; Liang, J.; Wu, Q.; Yang, Z.; Wu, K.; Fan, D. *R*-Flurbiprofen Reverses Multidrug Resistance, Proliferation and Metastasis in Gastric Cancer Cells by P75 NTR Induction. *Mol. Pharmaceutics* **2010**, *7*, 156–168.
- (172) Geerts, H. Drug Evaluation: (*R*)-Flurbiprofen-an Enantiomer of Flurbiprofen for the Treatment of Alzheimer's Disease. *IDrugs* **2007**, *10*, 121–133.
- (173) Terao, Y.; Ijima, Y.; Kakidani, H.; Ohta, H. Enzymatic Synthesis of (*R*)-Flurbiprofen. *Bull. Chem. Soc. Jpn.* **2003**, *76*, 2395–2397.
- (174) Hylton, T. A.; Walker, J. A. Process for Preparing Arylmethylmaonate Esters,
-

- Novel Products Thereof, and Processes for Converting the Products to Therapeutic 2-Arylpropionic Acids and Esters. EP0032620A1, 1981.
- (175) Mizushima, T.; Otsuka, M.; Okamoto, Y.; Yamakawa, N. 2-Fluorophenylpropionic Acid Derivative. EP2799424A1, 2014.
- (176) Lu, G.; Franzén, R.; Yu, X. J.; Xu, Y. J. Synthesis of Flurbiprofen via Suzuki Reaction Catalyzed by Palladium Charcoal in Water. *Chin. Chem. Lett.* **2006**, *17*, 461–464.
- (177) Aßmann, M.; Stöbener, A.; Mügge, C.; Gaßmeyer, S. K.; Hilterhaus, L.; Kourist, R.; Liese, A.; Kara, S. Reaction Engineering of Biocatalytic (S)-Naproxen Synthesis Integrating in-Line Process Monitoring by Raman Spectroscopy. *React. Chem. Eng.* **2017**, *2*, 531–540.
- (178) Aßmann, M.; Mügge, C.; Gaßmeyer, S. K.; Enoki, J.; Hilterhaus, L.; Kourist, R.; Liese, A.; Kara, S. Improvement of the Process Stability of Arylmalonate Decarboxylase by Immobilization for Biocatalytic Profen Synthesis. *Front. Microbiol.* **2017**, *8*, 1–8.
- (179) Wong, L. S.; Okrasa, K.; Micklefield, J. Site-Selective Immobilisation of Functional Enzymes on to Polystyrene Nanoparticles. *Org. Biomol. Chem.* **2010**, *8*, 782–787.
- (180) Markošová, K.; Husarčíková, J.; Halášová, M.; Kourist, R.; Rosenberg, M.; Stloukal, R.; Zajoncová, L.; Rebroš, M. Immobilization of Arylmalonate Decarboxylase. *Catalysts* **2018**, *8*, 603.
- (181) Cantone, S.; Ferrario, V.; Corici, L.; Ebert, C.; Fattor, D.; Spizzo, P.; Gardossi, L. Efficient Immobilisation of Industrial Biocatalysts: Criteria and Constraints for the Selection of Organic Polymeric Carriers and Immobilisation Methods. *Chem. Soc. Rev.* **2013**, *42*, 6262–6276.
- (182) Miyamoto, K.; Yatake, Y.; Tamura, K.; Terao, Y.; Ohta, H. Purification and Characterization of Arylmalonate Decarboxylase from *Achromobacter* Sp. KU1311. *J. Biosci. Bioeng.* **2007**, *104*, 263–267.
- (183) Yatake, Y.; Miyamoto, K.; Ohta, H. Screening, Cloning, Expression, and Purification of an Acidic Arylmalonate Decarboxylase from *Enterobacter Cloacae* KU1313. *Appl. Microbiol. Biotechnol.* **2008**, *78*, 793–799.
- (184) Maimanakos, J.; Chow, J.; Gaßmeyer, S. K.; Güllert, S.; Busch, F.; Kourist, R.; Streit, W. R. Sequence-Based Screening for Rare Enzymes: New Insights into the World of AMDases Reveal a Conserved Motif and 58 Novel Enzymes Clustering in Eight Distinct Families. *Front. Microbiol.* **2016**, *7*, 1332.
- (185) Fraser, R. R.; Mansour, T. S. The Effect of Additives on Organolithiums. *Tetrahedron Lett.* **1986**, *27*, 331–334.

-
- (186) Collum, D. B.; McNeil, A. J.; Ramirez, A. Lithium Diisopropylamide: Solution Kinetics and Implications for Organic Synthesis. *Angew. Chem., Int. Ed.* **2007**, *46*, 3002–3017.
- (187) Arnold, R. T.; Campos, M. de M.; Lindsay, K. L. Participation of a Neighboring Carboxyl Group in Addition Reactions. I. The Mechanism of the Reaction of Bromine with γ,δ -Unsaturated Acids and Esters. *J. Am. Chem. Soc.* **1953**, *75*, 1044–1047.
- (188) Ingold, C. K.; Rogers, M. A. T. The Modes of Addition to Conjugated Unsaturated Systems. Part VIII. Reduction of α -Vinylcinnamic Acid. *J. Chem. Soc.* **1935**, 717–721.
- (189) Pollack, R. M. Decarboxylations of β -Keto Acids and Related Compounds. In *Transition States of Biochemical Processes*; Grandour, R. D., Schowen, R. L., Eds.; Springer US: Boston, MA, 1978; pp 467–492.
- (190) Hamlin, T. A.; Swart, M.; Bickelhaupt, F. M. Nucleophilic Substitution (S_N2): Dependence on Nucleophile, Leaving Group, Central Atom, Substituents, and Solvent. *ChemPhysChem* **2018**, *19*, 1315–1330.
- (191) Dostrovsky, I.; Hughes, E. D. Mechanism of Substitution at a Saturated Carbon Atom. Part XXVI. The Role of Steric Hindrance. (Section A) Introductory Remarks, and a Kinetic Study of the Reactions of Methyl, Ethyl, *n*-Propyl, Isobutyl, and Neopentyl Bromides with Sodium Ethoxide. *J. Chem. Soc.* **1946**, 157–161.
- (192) Arisawa, M.; Akamatsu, K.; Yamaguchi, M. GaCl_3 -Promoted Ethenylation of Silylated β -Dicarbonyl Compound with Silylethyne. Synthesis of Ethenylmalonate. *Org. Lett.* **2001**, *3*, 789–790.
- (193) Phillips, R. S. Temperature Effects on Stereochemistry of Enzymatic Reactions. *Enzyme Microb. Technol.* **1992**, *14*, 417–419.
- (194) Phillips, R. S. Temperature Modulation of the Stereochemistry of Enzymatic Catalysis: Prospects for Exploitation. *Trends Biotechnol.* **1996**, *14*, 13–16.
- (195) Straathof, A. J. J.; Jongejan, J. A. The Enantiomeric Ratio: Origin, Determination and Prediction. *Enzyme Microb. Technol.* **1997**, *21*, 559–571.
- (196) Warshel, A.; Weiss, R. M. An Empirical Valence Bond Approach for Comparing Reactions in Solutions and in Enzymes. *J. Am. Chem. Soc.* **1980**, *102*, 6218–6226.
- (197) Barducci, A.; Bussi, G.; Parrinello, M. Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Phys. Rev. Lett.* **2008**, *100*, 020603.
- (198) Biler, M.; Crean, R. M.; Schweiger, A. K.; Kourist, R.; Kamerlin, S. C. L. Ground-
-

- State Destabilization by Active-Site Hydrophobicity Controls the Selectivity of a Cofactor-Free Decarboxylase. *J. Am. Chem. Soc.* **2020**, *142*, 20216–20231.
- (199) Kulkarni, Y. S.; Liao, Q.; Petrović, D.; Krüger, D. M.; Strodel, B.; Amyes, T. L.; Richard, J. P.; Kamerlin, S. C. L. Enzyme Architecture: Modeling the Operation of a Hydrophobic Clamp in Catalysis by Triosephosphate Isomerase. *J. Am. Chem. Soc.* **2017**, *139*, 10514–10525.
- (200) Richard, J. P.; Amyes, T. L.; Goryanova, B.; Zhai, X. Enzyme Architecture: On the Importance of Being in a Protein Cage. *Curr. Opin. Chem. Biol.* **2014**, *21*, 1–10.
- (201) Blaha-Nelson, D.; Krüger, D. M.; Szeler, K.; Ben-David, M.; Kamerlin, S. C. L. Active Site Hydrophobicity and the Convergent Evolution of Paraoxonase Activity in Structurally Divergent Enzymes: The Case of Serum Paraoxonase 1. *J. Am. Chem. Soc.* **2017**, *139*, 1155–1167.
- (202) Li, C.-H.; Tu, S.-C. Active Site Hydrophobicity Is Critical to the Bioluminescence Activity of *Vibrio Harveyi* Luciferase. *Biochemistry* **2005**, *44*, 12970–12977.
- (203) Nguyen, C. N.; Kurtzman Young, T.; Gilson, M. K. Grid Inhomogeneous Solvation Theory: Hydration Structure and Thermodynamics of the Miniature Receptor Cucurbit[7]Uril. *J. Chem. Phys.* **2012**, *137*, 044101.
- (204) Nguyen, C. N.; Cruz, A.; Gilson, M. K.; Kurtzman, T. Thermodynamics of Water in an Enzyme Active Site: Grid-Based Hydration Analysis of Coagulation Factor Xa. *J. Chem. Theory Comput.* **2014**, *10*, 2769–2780.
- (205) Kraml, J.; Kamenik, A. S.; Waibl, F.; Schauerl, M.; Liedl, K. R. Solvation Free Energy as a Measure of Hydrophobicity: Application to Serine Protease Binding Interfaces. *J. Chem. Theory Comput.* **2019**, *15*, 5872–5882.
- (206) Hayaishi, O. Oxygenases. In *Encyclopedia of Biological Chemistry*; Lennarz, W. J., Lane, M. D., Eds.; Elsevier, 2013; pp 371–374.
- (207) Hayaishi, O.; Katagiri, M.; Rothberg, S. Mechanism of the Pyrocatechase Reaction. *J. Am. Chem. Soc.* **1955**, *77*, 5450–5451.
- (208) Mason, H. S.; Fowlks, W. L.; Peterson, E. Oxygen Transfer and Electron Transport by the Phenolase Complex 1. *J. Am. Chem. Soc.* **1955**, *77*, 2914–2915.
- (209) Nozaki, M.; Ishimura, Y. Oxygenases. In *Microbial Iron Metabolism*; Neilands, J. B., Ed.; Elsevier, 1974; pp 417–444.
- (210) Faber, K. Biocatalytic Applications. In *Biotransformations in Organic Chemistry*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2011; pp 31–313.
- (211) Aik, W.; McDonough, M. A.; Thalhammer, A.; Chowdhury, R.; Schofield, C. J. Role of the Jelly-Roll Fold in Substrate Binding by 2-Oxoglutarate Oxygenases.

-
- Curr. Opin. Struct. Biol.* **2012**, *22*, 691–700.
- (212) Islam, M. S.; Leissing, T. M.; Chowdhury, R.; Hopkinson, R. J.; Schofield, C. J. 2-Oxoglutarate-Dependent Oxygenases. *Annu. Rev. Biochem.* **2018**, *87*, 585–620.
- (213) Martinez, S.; Hausinger, R. P. Catalytic Mechanisms of Fe(II)- and 2-Oxoglutarate-Dependent Oxygenases. *J. Biol. Chem.* **2015**, *290*, 20702–20711.
- (214) Hausinger, R. P. Fe(II)/ α -Ketoglutarate-Dependent Hydroxylases and Related Enzymes. *Crit. Rev. Biochem. Mol. Biol.* **2004**, *39*, 21–68.
- (215) Dann, C. E.; Bruick, R. K.; Deisenhofer, J. Nonlinear Partial Differential Equations and Applications: Structure of Factor-Inhibiting Hypoxia-Inducible Factor 1: An Asparaginyl Hydroxylase Involved in the Hypoxic Response Pathway. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 15351–15356.
- (216) Strijbis, K.; Vaz, F. M.; Distel, B. Enzymology of the Carnitine Biosynthesis Pathway. *IUBMB Life* **2010**, *65*, 357–362.
- (217) Shoulders, M. D.; Raines, R. T. Collagen Structure and Stability. *Annu. Rev. Biochem.* **2009**, *78*, 929–958.
- (218) Trewick, S. C.; Henshaw, T. F.; Hausinger, R. P.; Lindahl, T.; Sedgwick, B. Oxidative Demethylation by *Escherichia Coli* AlkB Directly Reverts DNA Base Damage. *Nature* **2002**, *419*, 174–178.
- (219) Rose, N. R.; McDonough, M. A.; King, O. N. F.; Kawamura, A.; Schofield, C. J. Inhibition of 2-Oxoglutarate Dependent Oxygenases. *Chem. Soc. Rev.* **2011**, *40*, 4364.
- (220) Hibi, M.; Ogawa, J. Characteristics and Biotechnology Applications of Aliphatic Amino Acid Hydroxylases Belonging to the Fe(II)/ α -Ketoglutarate-Dependent Dioxygenase Superfamily. *Appl. Microbiol. Biotechnol.* **2014**, *98*, 3869–3876.
- (221) Blasiak, L. C.; Vaillancourt, F. H.; Walsh, C. T.; Drennan, C. L. Crystal Structure of the Non-Haem Iron Halogenase SyrB2 in Syringomycin Biosynthesis. *Nature* **2006**, *440*, 368–371.
- (222) Clifton, I. J.; McDonough, M. A.; Ehrismann, D.; Kershaw, N. J.; Granatino, N.; Schofield, C. J. Structural Studies on 2-Oxoglutarate Oxygenases and Related Double-Stranded β -Helix Fold Proteins. *J. Inorg. Biochem.* **2006**, *100*, 644–669.
- (223) Bleijlevens, B.; Shivarattan, T.; van den Boom, K. S.; de Haan, A.; van der Zwan, G.; Simpson, P. J.; Matthews, S. J. Changes in Protein Dynamics of the DNA Repair Dioxygenase AlkB upon Binding of Fe²⁺ and 2-Oxoglutarate. *Biochemistry* **2012**, *51*, 3334–3341.
- (224) Grzyska, P. K.; Appelman, E. H.; Hausinger, R. P.; Proshlyakov, D. A. Insight into the Mechanism of an Iron Dioxygenase by Resolution of Steps Following the
-

- Fe^{IV}=O Species. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 3982–3987.
- (225) Mantri, M.; Zhang, Z.; McDonough, M. A.; Schofield, C. J. Autocatalysed Oxidative Modifications to 2-Oxoglutarate Dependent Oxygenases. *FEBS J.* **2012**, *279*, 1563–1575.
- (226) Counts, D. F.; Cardinale, G. J.; Udenfriend, S. Prolyl Hydroxylase Half Reaction: Peptidyl Prolyl-Independent Decarboxylation of Alpha-Ketoglutarate. *Proc. Natl. Acad. Sci. U. S. A.* **1978**, *75*, 2145–2149.
- (227) Myllylä, R.; Majamaa, K.; Günzler, V.; Hanauske-Abel, H. M.; Kivirikko, K. I. Ascorbate Is Consumed Stoichiometrically in the Uncoupled Reactions Catalyzed by Prolyl 4-Hydroxylase and Lysyl Hydroxylase. *J. Biol. Chem.* **1984**, *259*, 5403–5405.
- (228) Liu, A.; Ho, R. Y. N.; Que, L.; Ryle, M. J.; Phinney, B. S.; Hausinger, R. P. Alternative Reactivity of an α -Ketoglutarate-Dependent Iron(II) Oxygenase: Enzyme Self-Hydroxylation. *J. Am. Chem. Soc.* **2001**, *123*, 5126–5127.
- (229) Ryle, M. J.; Liu, A.; Muthukumar, R. B.; Ho, R. Y. N.; Koehntop, K. D.; McCracken, J.; Que, L.; Hausinger, R. P. O₂- and α -Ketoglutarate-Dependent Tyrosyl Radical Formation in TauD, an α -Keto Acid-Dependent Non-Heme Iron Dioxygenase. *Biochemistry* **2003**, *42*, 1854–1862.
- (230) Sundheim, O.; Vågbø, C. B.; Bjørås, M.; Sousa, M. M. L.; Talstad, V.; Aas, P. A.; Drabløs, F.; Krokan, H. E.; Tainer, J. A.; Slupphaug, G. Human ABH3 Structure and Key Residues for Oxidative Demethylation to Reverse DNA/RNA Damage. *EMBO J.* **2006**, *25*, 3389–3397.
- (231) Zhang, Z.; Barlow, J. N.; Baldwin, J. E.; Schofield, C. J. Metal-Catalyzed Oxidation and Mutagenesis Studies on the Iron(II) Binding Site of 1-Aminocyclopropane-1-Carboxylate Oxidase. *Biochemistry* **1997**, *36*, 15999–16007.
- (232) Remuzon, P. *Trans*-4-Hydroxy-L-Proline, a Useful and Versatile Chiral Starting Block. *Tetrahedron* **1996**, *52*, 13803–13835.
- (233) Hara, R.; Kino, K. Characterization of Novel 2-Oxoglutarate Dependent Dioxygenases Converting L-Proline to *Cis*-4-Hydroxy-L-Proline. *Biochem. Biophys. Res. Commun.* **2009**, *379*, 882–886.
- (234) Shibasaki, T.; Mori, H.; Ozaki, A. Enzymatic Production of *Trans*-4-Hydroxy-L-Proline by Regio- and Stereospecific Hydroxylation of L-Proline. *Biosci., Biotechnol., Biochem.* **2000**, *64*, 746–750.
- (235) Mori, H.; Shibasaki, T.; Yano, K.; Ozaki, A. Purification and Cloning of a Proline 3-Hydroxylase, a Novel Enzyme Which Hydroxylates Free L-Proline to *Cis*-3-Hydroxy-L-Proline. *J. Bacteriol.* **1997**, *179*, 5677–5683.

-
- (236) Strieker, M.; Kopp, F.; Mahlert, C.; Essen, L.-O.; Marahiel, M. A. Mechanistic and Structural Basis of Stereospecific C β -Hydroxylation in Calcium-Dependent Antibiotic, a Daptomycin-Type Lipopeptide. *ACS Chem. Biol.* **2007**, *2*, 187–196.
- (237) Yin, X.; Zabriskie, T. M. VioC Is a Non-Heme Iron, α -Ketoglutarate-Dependent Oxygenase That Catalyzes the Formation of 3S-Hydroxy-L-Arginine during Viomycin Biosynthesis. *ChemBioChem* **2004**, *5*, 1274–1277.
- (238) Kodera, T.; Smirnov, S. V.; Samsonova, N. N.; Kozlov, Y. I.; Koyama, R.; Hibi, M.; Ogawa, J.; Yokozeki, K.; Shimizu, S. A Novel L-Isoleucine Hydroxylating Enzyme, L-Isoleucine Dioxygenase from *Bacillus Thuringiensis*, Produces (2S,3R,4S)-4-Hydroxyisoleucine. *Biochem. Biophys. Res. Commun.* **2009**, *390*, 506–510.
- (239) Hibi, M.; Kawashima, T.; Kasahara, T.; Sokolov, P. M.; Smirnov, S. V.; Kodera, T.; Sugiyama, M.; Shimizu, S.; Yokozeki, K.; Ogawa, J. A Novel Fe(II)/ α -Ketoglutarate-Dependent Dioxygenase from *Burkholderia Ambifaria* Has β -Hydroxylating Activity of *N*-Succinyl L-Leucine. *Lett. Appl. Microbiol.* **2012**, *55*, 414–419.
- (240) Hibi, M.; Kawashima, T.; Sokolov, P. M.; Smirnov, S. V.; Kodera, T.; Sugiyama, M.; Shimizu, S.; Yokozeki, K.; Ogawa, J. L-Leucine 5-Hydroxylase of *Nostoc Punctiforme* Is a Novel Type of Fe(II)/ α -Ketoglutarate-Dependent Dioxygenase That Is Useful as a Biocatalyst. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 2467–2472.
- (241) Correia Cordeiro, R. S.; Enoki, J.; Busch, F.; Mügge, C.; Kourist, R. Cloning and Characterization of a New Delta-Specific L-Leucine Dioxygenase from *Anabaena Variabilis*. *J. Biotechnol.* **2018**, *284*, 68–74.
- (242) Luesch, H.; Hoffmann, D.; Hevel, J. M.; Becker, J. E.; Golakoti, T.; Moore, R. E. Biosynthesis of 4-Methylproline in Cyanobacteria: Cloning of *NosE* and *NosF* Genes and Biochemical Characterization of the Encoded Dehydrogenase and Reductase Activities. *J. Org. Chem.* **2003**, *68*, 83–91.
- (243) Smirnov, S. V.; Sokolov, P. M.; Kodera, T.; Sugiyama, M.; Hibi, M.; Shimizu, S.; Yokozeki, K.; Ogawa, J. A Novel Family of Bacterial Dioxygenases That Catalyse the Hydroxylation of Free L-Amino Acids. *FEMS Microbiology Letters* **2012**, *331*, 97–104.
- (244) Hibi, M.; Kawashima, T.; Kodera, T.; Smirnov, S. V.; Sokolov, P. M.; Sugiyama, M.; Shimizu, S.; Yokozeki, K.; Ogawa, J. Characterization of *Bacillus Thuringiensis* L-Isoleucine Dioxygenase for Production of Useful Amino Acids. *Appl. Environ. Microbiol.* **2011**, *77*, 6926–6930.
- (245) Broca, C.; Gross, R.; Petit, P.; Sauvaire, Y.; Manteghetti, M.; Tournier, M.;
-

- Masiello, P.; Gomis, R.; Ribes, G. 4-Hydroxyisoleucine: Experimental Evidence of Its Insulinotropic and Antidiabetic Properties. *Am. J. Physiol.: Endocrinol. Metab.* **1999**, *277*, E617–E623.
- (246) Haefel , C.; Bonfils, C.; Sauvaire, Y. Characterization of a Dioxygenase from *Trigonella Foenum-Graecum* Involved in 4-Hydroxyisoleucine Biosynthesis. *Phytochemistry* **1997**, *44*, 563–566.
- (247) Alcock, N. W.; Crout, D. H. G.; Gregorio, M. V. M.; Lee, E.; Pike, G.; Samuel, C. J. Stereochemistry of the 4-Hydroxyisoleucine from *Trigonella Foenum-Graecum*. *Phytochemistry* **1989**, *28*, 1835–1841.
- (248) Hibi, M.; Kawashima, T.; Yajima, H.; Smirnov, S. V.; Kodera, T.; Sugiyama, M.; Shimizu, S.; Yokozeki, K.; Ogawa, J. Enzymatic Synthesis of Chiral Amino Acid Sulfoxides by Fe(II)/ α -Ketoglutarate-Dependent Dioxygenase. *Tetrahedron: Asymmetry* **2013**, *24*, 990–994.
- (249) Ogawa, J.; Kodera, T.; Smirnov, S. V.; Hibi, M.; Samsonova, N. N.; Koyama, R.; Yamanaka, H.; Mano, J.; Kawashima, T.; Yokozeki, K.; Shimizu, S. A Novel L-Isoleucine Metabolism in *Bacillus Thuringiensis* Generating (2S,3R,4S)-4-Hydroxyisoleucine, a Potential Insulinotropic and Anti-Obesity Amino Acid. *Appl. Microbiol. Biotechnol.* **2011**, *89*, 1929–1938.
- (250) Smirnov, S. V.; Kodera, T.; Samsonova, N. N.; Kotlyarova, V. A.; Rushkevich, N. Y.; Kivero, A. D.; Sokolov, P. M.; Hibi, M.; Ogawa, J.; Shimizu, S. Metabolic Engineering of *Escherichia Coli* to Produce (2S, 3R, 4S)-4-Hydroxyisoleucine. *Appl. Microbiol. Biotechnol.* **2010**, *88*, 719–726.
- (251) Theodosiou, E.; Frick, O.; B hler, B.; Schmid, A. Metabolic Network Capacity of *Escherichia Coli* for Krebs Cycle-Dependent Proline Hydroxylation. *Microb. Cell Fact.* **2015**, *14*, 108.
- (252) Zhang, H.-L.; Zhang, C.; Pei, C.-H.; Han, M.-N.; Xu, Z.-D.; Li, C.-H.; Li, W. Efficient Production of *Trans*-4-Hydroxy-L-Proline from Glucose by Metabolic Engineering of Recombinant *Escherichia Coli*. *Lett. Appl. Microbiol.* **2018**, *66*, 400–408.
- (253) Zhang, Y.; Zhang, Y.; Shang, X.; Wang, B.; Hu, Q.; Liu, S.; Wen, T. Reconstruction of Tricarboxylic Acid Cycle in *Corynebacterium Glutamicum* with a Genome-Scale Metabolic Network Model for *Trans*-4-Hydroxyproline Production. *Biotechnol. Bioeng.* **2019**, *116*, 99–109.
- (254) Sun, D.; Gao, D.; Liu, X.; Zhu, M.; Li, C.; Chen, Y.; Zhu, Z.; Lu, F.; Qin, H.-M. Redesign and Engineering of a Dioxygenase Targeting Biocatalytic Synthesis of 5-Hydroxyl Leucine. *Catal. Sci. Technol.* **2019**, *9*, 1825–1834.
- (255) Zhang, C.; Ma, J.; Li, Z.; Liang, Y.; Xu, Q.; Xie, X.; Chen, N. A Strategy for L-

- Isoleucine Dioxygenase Screening and 4-Hydroxyisoleucine Production by Resting Cells. *Bioengineered* **2018**, *9*, 72–79.
- (256) Busch, F.; Brummund, J.; Calderini, E.; Schürmann, M.; Kourist, R. Cofactor Generation Cascade for α -Ketoglutarate and Fe(II)-Dependent Dioxygenases. *ACS Sustainable Chem. Eng.* **2020**, *8*, 8604–8612.
- (257) Sun, D.; Liu, X.; Zhu, M.; Chen, Y.; Li, C.; Cheng, X.; Zhu, Z.; Lu, F.; Qin, H.-M. Efficient Biosynthesis of High-Value Succinic Acid and 5-Hydroxyisoleucine Using a Multienzyme Cascade and Whole-Cell Catalysis. *J. Agric. Food Chem.* **2019**, *67*, 12502–12510.
- (258) Kuipers, R. K.; Joosten, H.-J.; van Berkel, W. J. H.; Leferink, N. G. H.; Rooijen, E.; Ittmann, E.; van Zimmeren, F.; Jochens, H.; Bornscheuer, U.; Vriend, G.; Martins dos Santos, V. A. P.; Schaap, P. J. 3DM: Systematic Analysis of Heterogeneous Superfamily Data to Discover Protein Functionalities. *Proteins: Struct., Funct., Bioinf.* **2010**, *78*, 2101–2113.
- (259) Pieber, S. Engineering of Dioxygenases. Master's Thesis, Graz University of Technology, 2020.
- (260) Hanreich, S. Identification and Selection of Active α -Ketoglutarate Dependent Dioxygenases in Engineered *Escherichia Coli*. Master's Thesis, Graz University of Technology, 2019.
- (261) Frey, R.; Hayashi, T.; Buller, R. M. Directed Evolution of Carbon–Hydrogen Bond Activating Enzymes. *Curr. Opin. Biotechnol.* **2019**, *60*, 29–38.
- (262) Lukat, P.; Katsuyama, Y.; Wenzel, S.; Binz, T.; König, C.; Blankenfeldt, W.; Brönstrup, M.; Müller, R. Biosynthesis of Methyl-Proline Containing Griselimycins, Natural Products with Anti-Tuberculosis Activity. *Chem. Sci.* **2017**, *8*, 7521–7527.
- (263) Zwick, C. R.; Renata, H. Remote C–H Hydroxylation by an α -Ketoglutarate-Dependent Dioxygenase Enables Efficient Chemoenzymatic Synthesis of Manzacidin C and Proline Analogs. *J. Am. Chem. Soc.* **2018**, *140*, 1165–1169.
- (264) An, Y.; Ji, J.; Wu, W.; Lv, A.; Huang, R.; Wei, Y. A Rapid and Efficient Method for Multiple-Site Mutagenesis with a Modified Overlap Extension PCR. *Appl. Microbiol. Biotechnol.* **2005**, *68*, 774–778.
- (265) Miyazaki, K. MEGAWHOP Cloning. In *Methods in Enzymology*; Elsevier Inc., 2011; Vol. 498, pp 399–406.
- (266) Persky, N. S.; Lovett, S. T. Mechanisms of Recombination: Lessons from *E. Coli*. *Crit. Rev. Biochem. Mol. Biol.* **2008**, *43*, 347–370.
- (267) Friedman, M. Applications of the Ninhydrin Reaction for Analysis of Amino Acids, Peptides, and Proteins to Agricultural and Biomedical Sciences. *J. Agric.*

-
- Food Chem.* **2004**, *52*, 385–406.
- (268) Harding, V. J.; Warneford, F. H. S. The Ninhydrin Reaction With Amino-Acids and Ammonium Salts. *J. Biol. Chem.* **1916**, *25*, 319–335.
- (269) Meng, J.; Wang, B.; Liu, D.; Chen, T.; Wang, Z.; Zhao, X. High-Yield Anaerobic Succinate Production by Strategically Regulating Multiple Metabolic Pathways Based on Stoichiometric Maximum in *Escherichia Coli*. *Microb. Cell Fact.* **2016**, *15*, 141.
- (270) Zhu, L.-W.; Tang, Y.-J. Current Advances of Succinate Biosynthesis in Metabolically Engineered *Escherichia Coli*. *Biotechnol. Adv.* **2017**, *35*, 1040–1048.
- (271) Pham, V. D.; Somasundaram, S.; Lee, S. H.; Park, S. J.; Hong, S. H. Engineering the Intracellular Metabolism of *Escherichia Coli* to Produce Gamma-Aminobutyric Acid by Co-Localization of GABA Shunt Enzymes. *Biotechnol. Lett.* **2016**, *38*, 321–327.
- (272) Li, Z.; Nimtz, M.; Rinas, U. The Metabolic Potential of *Escherichia Coli* BL21 in Defined and Rich Medium. *Microb. Cell Fact.* **2014**, *13*, 45.
- (273) Hewitson, K. S.; Holmes, S. L.; Ehrismann, D.; Hardy, A. P.; Chowdhury, R.; Schofield, C. J.; McDonough, M. A. Evidence That Two Enzyme-Derived Histidine Ligands Are Sufficient for Iron Binding and Catalysis by Factor Inhibiting HIF (FIH). *J. Biol. Chem.* **2008**, *283*, 25971–25978.
- (274) Bertram, R.; Hillen, W. The Application of Tet Repressor in Prokaryotic Gene Regulation and Expression. *Microb. Biotechnol.* **2008**, *1*, 2–16.
- (275) Skerra, A. Use of the Tetracycline Promoter for the Tightly Regulated Production of a Murine Antibody Fragment in *Escherichia Coli*. *Gene* **1994**, *151*, 131–135.
- (276) Rashaid, A. H. B.; Harrington, P. de B.; Jackson, G. P. Profiling Amino Acids of Jordanian Scalp Hair as a Tool for Diabetes Mellitus Diagnosis: A Pilot Study. *Anal. Chem.* **2015**, *87*, 7078–7084.
- (277) Clark, D. P. The Fermentation Pathways of *Escherichia Coli*. *FEMS Microbiology Letters* **1989**, *63*, 223–234.
- (278) Martinez, I.; Gao, H.; Bennett, G. N.; San, K.-Y. High Yield Production of Four-Carbon Dicarboxylic Acids by Metabolically Engineered *Escherichia Coli*. *J. Ind. Microbiol. Biotechnol.* **2018**, *45*, 53–60.
- (279) Freundlich, M.; Burns, R. O.; Umbarger, H. E. Control of Isoleucine, Valine, and Leucine Biosynthesis, I. Multi-Valent Repression. *Proc. Natl. Acad. Sci. U. S. A.* **1962**, *48*, 1804–1808.
- (280) Benchling [Biology Software] <https://benchling.com>.
- (281) Fulmer, G. R.; Miller, A. J. M.; Sherden, N. H.; Gottlieb, H. E.; Nudelman, A.;
-

- Stoltz, B. M.; Bercaw, J. E.; Goldberg, K. I. NMR Chemical Shifts of Trace Impurities: Common Laboratory Solvents, Organics, and Gases in Deuterated Solvents Relevant to the Organometallic Chemist. *Organometallics* **2010**, *29*, 2176–2179.
- (282) Nomenclature for Incompletely Specified Bases in Nucleic Acid Sequences. Recommendations 1984. Nomenclature Committee of the International Union of Biochemistry (NC-IUB). *Proc. Natl. Acad. Sci. U. S. A.* **1986**, *83*, 4–8.

Appendix A Abbreviations

(v/v)	Volume per volume	COMT	Caffeic acid <i>O</i> -methyl transferase
(w/v)	Weight per volume		
(w/w)	Weight per weight	COSY	Correlated spectroscopy NMR experiment
°C	Degrees Celsius		
4CL	4-coumaroyl CoA ligase	cPCR	Colony polymerase chain reaction
AcOH	Acetic acid		
AMDase	Arylmalonate decarboxylase	CV	Column volume
AMKP	(2 <i>S</i> ,3 <i>R</i>)-2-amino-3-methyl-4-ketopentanoate	CWW	Cell wet weight
		DCM	Dichloromethane
APT	Attached proton test NMR experiment	ddH₂O	Double-distilled water
		DES	Deep eutectic solvent
BLAST	Basic Local Alignment Search Tool	DFT	Density functional theory
		dH₂O	Distilled water
bp	Base pair	DIPA	Diisopropylamine
BSA	Bovine serum albumin	DMPU	<i>N,N'</i> -Dimethylpropyleneurea
BSTFA	<i>N,O</i> -Bis(trimethylsilyl) trifluoroacetamide	dNTPs	Deoxynucleotide Triphosphates
		DSBH	Double-stranded β -helix fold
BTB	Bromothymol blue	<i>E</i>	Enantiomeric ratio
Bu	Butyl	EA	Ethyl acetate
BuLi	Butyl lithium	ee	Enantiomeric excess
C3H	<i>p</i> -coumarate 3-hydroxylase	eq	Equivalent
C4H	Cinnamate 4-hydroxylase	Et	Ethyl
CAM	Cerium ammonium molybdate stain	EtOAc	Ethyl acetate
		EVB	Empirical valence bond
CAST	Combinatorial active-site saturation test	EWG	Electron withdrawing group
		F5H	Ferulate 5-hydroxylase
CCoAOMT	Caffeoyl-CoA <i>O</i> -methyl transferase	FAD	Ferulic acid decarboxylase
		FDC	Ferulate decarboxylase
CCR	Cinnamoyl-CoA reductase	FerA	Feruloyl CoA synthetase
CFE	Cell-free extract	FerB	Feruloyl CoA hydratase/lyase
cfu	Colony forming units	fwd	Forward
CH	Cyclohexane	<i>g</i>	Earth's gravitational field
ChCl	Choline chloride	GABA	γ -Aminobutyric acid
CHES	<i>N</i> -Cyclohexyl-2-aminoethanesulfonic acid	GC	Gas chromatography
		GC-FID	Gas chromatography with flame ionization detection

GC-MS	Gas chromatography with mass spectrometry	MHz	Megahertz (10^6 Hz)
gDNA	Genomic DNA	min	Minute
GIST	Grid inhomogeneous solvation theory	mL	Milliliter (10^{-3} L)
GPP	General phenylpropanoid pathway	mm	Millimeter (10^{-3} m)
h	Hours	mM	Millimolar (mmol L ⁻¹)
HCHL	Hydroxycinnamoyl-CoA hydratase lyase	msec	Milliseconds
HFCS	High-fructose corn syrup	MTBE	Methyl <i>tert</i> -butyl ether
HIL	Hydroxyisoleucine	MWCO	Molecular weight cut-off
HPLC	High performance liquid chromatography	n.a.	Not applicable
HSQC	Heteronuclear single quantum coherence NMR experiment	n.c.	Not converted
<i>i</i>	<i>Iso</i>	n.d.	Not determined
IDO	Isoleucine dioxygenase	ng	Nanogram (10^{-9} g)
IPTG	Isopropyl β -D-1-thiogalactopyranoside	nm	Nanometer (10^{-9} m)
ISM	Iterative saturation mutagenesis	NMR	Nuclear magnetic resonance
Kan	Kanamycin	NSAID	Non-steroidal anti-inflammatory drug
kb	Kilobase (10^3 bases)	o	<i>Ortho</i>
<i>k_{cat}</i>	Turnover number (sec ⁻¹)	o/n	Over night
kDa	Kilodalton (10^3 Da)	OD₆₀₀	Optical density at 600 nm
<i>K_M</i>	Michaelis constant (in M)	OE-PCR	Overlap extension polymerase chain reaction
KP_i	Potassium phosphate	ONC	Overnight culture
kV	Kilovolt (103 V)	<i>p</i>	<i>Para</i>
LDA	Lithium diisopropylamide	P3H	Proline-3-hydroxylase
LDO	Leucine dioxygenase	P4H	Proline-4-hydroxylase
LGOX	L-Glutamate oxidase	PAD	Phenolic acid decarboxylase
M	Molar (mol L ⁻¹)	PAL	Phenylalanine ammonia lyase
<i>m</i>	<i>Meta</i>	PCR	Polymerase chain reaction
mA	Milliampere (10^{-3} A)	Pd/C	Palladium on charcoal
Me	Methyl	PDB	Protein data bank
MeCN	Acetonitrile	PDC	Phenolic acid decarboxylase
MEGAWHOP	Megaprimer PCR of whole plasmid	PEG	Polyethylene glycol
		PEP	Phosphoenolpyruvate
		Ph	Phenyl
		pmol	Picomole (10^{-12} mol)
		ppm	Parts per million
		Pr	Propyl
		PSAR	Phenolic acid stress-response

PSI-BLAST	Position-specific iterative basic local alignment search tool	αKG	Alpha-ketoglutarate (2-oxoglutarate)
PTAL	Bifunctional ammonia lyase	μg	Microgram (10 ⁻⁶ g)
PVA	Polyvinyl alcohol	μL	Microliter (10 ⁻⁶ L)
r.t.	Room temperature	μM	Micromolar (μmol L ⁻¹)
REF1	Reduced epidermal fluorescence 1		
rev	Reverse		
R_f	Retention factor		
rpm	Revolutions per minute		
SB-MSA	Structure-based multiple sequence alignment		
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis		
SN	Supernatant		
SSM	Site-saturation mutagenesis		
STD-NMR	Saturation-transfer-difference NMR		
T	Temperature (in °C)		
t	Time		
TAL	Tyrosine ammonia lyase		
TCA cycle	Tricarboxylic acid cycle		
TLC	Thin layer chromatography		
T_m	Melting temperature (in °C)		
TMS	Trimethylsilyl		
TPL	Tyrosine phenol lyase		
t_R	Retention time (in min)		
T_r	Racemic temperature (in °C)		
TTN	Total turnover number		
U	Enzyme unit		
UV	Ultraviolet light		
V	Volt		
VAO	Vanillyl alcohol oxidase		
var.	Variable		
WT	Wildtype		
WT-MetaD	Well-tempered metadynamics simulations		
Å	Angstrom (10 ⁻¹⁰ m)		

Appendix B Media, buffers and solutions

General media and buffers

Table 47. Composition of general media, buffers and solutions.

LB medium	
Tryptone	10 g
Yeast extract	5 g
NaCl	10 g
<i>Alternative:</i>	
Ready-to-use mixture	20 g
dH ₂ O	To 1 L and autoclave.
2xTY medium	
Tryptone	16 g
Yeast extract	10 g
NaCl	5 g
dH ₂ O	To 1 L and autoclave.
TB medium	
Tryptone	12 g
Yeast extract	24 g
Glycerol	5 g
dH ₂ O	To 900 mL and autoclave. Combine with 100 mL 10x TB salts before use.
10x TB salts	
KH ₂ PO ₄	2.31 g
K ₂ HPO ₄	12.54 g
dH ₂ O	To 100 mL and autoclave.
10x SOC	
KCl	2 g L ⁻¹
MgCl ₂ (anhyd.)	20 g L ⁻¹
MgSO ₄ (anhyd.)	20 g L ⁻¹
Glucose monohydrate	40 g L ⁻¹
Dissolve in dH ₂ O and sterilize by filtration (0.22 µm) or autoclaving.	
LB-SOC	
Dilute sterile 10x SOC in LB medium before use.	
Kanamycin 1000x (Kan)	
Kanamycin sulphate	40 mg mL ⁻¹
Dissolve in ddH ₂ O and sterilize by filtration (0.22 µm). Store aliquots at -20°C.	
Ampicillin 1000x (Amp)	
Ampicillin sodium salt	100 mg mL ⁻¹
Dissolve in ddH ₂ O and sterilize by filtration (0.22 µm). Store aliquots at -20°C.	
IPTG (1 M)	
Isopropyl-β-D-thiogalactopyranoside (IPTG)	238.3 mg mL ⁻¹
Dissolve in ddH ₂ O and sterilize by filtration (0.22 µm). Store aliquots at -20°C.	

Continuation of Table 47.

50x TAE (pH 8)	
Tris	242.3 g
EDTA	14.6 g
Acetic acid (conc.)	57 mL
Add ddH ₂ O to 1 L and adjust pH to 8.	
TNE buffer	
Tris	1.21 g
NaCl	0.58 g
EDTA (disodium dihydrate)	3.72 g
Add ddH ₂ O to 1 L and adjust pH to 8.	
TNEX buffer	
Triton X-100 (final: 1 % (v/v))	10 µL
TNE buffer	1 mL
BTB screening buffer	
Bromothymol blue (CAS: 76-59-5)	1 mg
Tris	121.1 mg (final: 10 mM)
Phenylmalonic acid	360.3 mg (final: 20 mM)
Add 100 mL ddH ₂ O, adjust pH to 7 and store aliquots at -20°C. Color of indicator changes upon freezing. Let thaw completely before use.	

Minimal media

Table 48. Solutions for preparation of M9 minimal medium.

M9 Stock (5x)	
Na ₂ HPO ₄ ·2 H ₂ O	42.5 g
KH ₂ PO ₄	15 g
NaCl	2.5 g
NH ₄ Cl	5 g
ddH ₂ O	To 1 L
Adjust pH to 7.0 with 4 M NaOH. Autoclavable at 121°C for 20 min. Store at r.t.	
MgSO₄ Stock (1 M)	
MgSO ₄ ·7 H ₂ O	24.6 g
ddH ₂ O	To 100 mL
Prepare solution and sterilize by autoclaving. Store at r.t.	
Thiamine Stock (1000x)	
Thiamine·HCl	50 mg
ddH ₂ O	To 50 mL
Dissolve and sterilize by filtration (0.22 µm). Store aliquots at -20°C.	
Biotin Stock (20 000x)	
Biotin	5 mg
ddH ₂ O	45 mL
NaOH, 1 N	Add until dissolved (1 eq)
ddH ₂ O	To 50 mL
Sterilize by filtration (0.22 µm). Store aliquots at -20°C.	

Continuation of Table 48.

Glucose Stock (20 %)	
Glucose Monohydrate	220 g
ddH ₂ O	To 1 L
Prepare solution and sterilize by autoclaving. Store at 4°C.	
US^{Fe} Trace Element Solution (1000x)	
FeSO ₄ ·7 H ₂ O	8.87 g
CaCl ₂ ·2 H ₂ O	4.12 g
MnCl ₂ ·2 H ₂ O	1.23 g
ZnSO ₄ ·7 H ₂ O	1.87 g
H ₃ BO ₃	0.30 g
Na ₂ MoO ₄ ·2 H ₂ O	0.25 g
CuCl ₂ ·2 H ₂ O	0.15 g
Disodium EDTA·2 H ₂ O	0.84 g
Dissolve in 1 M HCl, sterilize by filtration (0.22 µm) and store aliquots at -20°C.	
α-Ketoglutarate Stock (25x)	
α-Ketoglutaric acid	25 mg
ddH ₂ O	1 mL
Prepare freshly and sterilize by filtration (0.22 µm)	
Substrate Stock (50 mM) (10x)	
Freshly prepare a 50 mM solution (e.g.: 6.56 g/L Leu or Ile) of desired substrate in ddH ₂ O and sterilize by filtration (0.22 µm).	
Substrate Stock (25 g L⁻¹) (25x)	
Freshly prepare a 25 g L ⁻¹ solution (e.g.: Leu or Ile) of desired substrate in ddH ₂ O and sterilize by filtration (0.22 µm).	

Table 49. Composition of liquid M9 minimal medium.

M9 medium with substrate	M9	M9-Substrate	M9-αKG	M9-Substrate/αKG
M9 stock (5x)	200 mL	200 mL	200 mL	200 mL
MgSO ₄ stock (1 M)	2 mL	2 mL	2 mL	2 mL
Thiamine stock (1000x)	1 mL	1 mL	1 mL	1 mL
Biotin stock (20 000x)	50 µL	50 µL	50 µL	50 µL
US ^{Fe} trace elements (1000x)	1 mL	1 mL	1 mL	1 mL
Glucose (20 %)	25 mL	25 mL	25 mL	25 mL
Substrate stock (10x)	-	100 mL	-	-
Substrate stock (25x)	-	-	-	40 mL
α-Ketoglutarate Stock (25x)	-	-	50 mL	50 mL
ddH ₂ O, sterile	To 1 L	To 1 L	To 1 L	To 1 L
Add all components to sterile ddH ₂ O and store at 4°C. Always freshly prepare medium.				

Table 50. Composition of solid M9 minimal medium.

M9 agar plates	
Agar agar	7.5 g
ddH ₂ O	345 mL
Autoclave mixture and let cool to ~60°C before adding all other sterile components. Avoid bubble formation during ventilation of autoclave.	
M9 stock (5x)	100 mL
MgSO ₄ stock (1 M)	1 mL
Thiamine stock (1000x)	0.5 mL
Biotin stock (20 000x)	25 µL
US ^{Fe} trace elements (1000x)	0.5 µL
Glucose (20 %)	12.5 mL
Substrate stock (25x)	20 mL
α-Ketoglutarate Stock (25x)	20 mL
Kan (1000x)	500 µL
IPTG (1 M)	250 µL

Solutions for preparation of competent *E. coli* cells

Table 51. Buffers and solutions for preparation of chemo-competent *E. coli* cells.

TfB1 buffer	
CH ₃ COOK	1.47 g (0.03 M)
KCl	3.73 g (0.1 M)
CaCl ₂	0.55 g (0.01 M)
Glycerol	65 mL (13 %)
dH ₂ O	To 500 mL and autoclave.
TfB2 buffer	
KCl	0.37 g (0.01 M)
CaCl ₂	5.55 g (0.1 M)
MOPS	1.05 g (0.1 M)
Glycerol	65 mL (13 %)
dH ₂ O	To 500 mL and autoclave.
MgCl₂ solution	
MgCl ₂	0.0952 g/mL (1 M)
Prepare solution in dH ₂ O and sterilize by filtration.	

Buffers and solutions for SDS-PAGE

Table 52. Buffers and solutions for SDS-PAGE.

Tris buffer A	
Tris	45.4 g
SDS	1 g
dH ₂ O	To 250 mL, adjust to pH 8.8 with HCl

Continuation of Table 52.

Tris buffer B	
Tris	15.1 g
SDS	0.25 g
dH ₂ O	To 250 mL, adjust to pH 6.8 with HCl
10x Running buffer	
Tris	30 g
Glycin	144 g
SDS	10 g
dH ₂ O	To 1 L, adjust to pH 8.3 with HCl
4x Sample buffer	
Tris-HCl 0.5 M, pH 6.8	1.25 mL
10 % (w/v) SDS in dH ₂ O	2 mL (can be stored at r.t.)
0.5 % (w/v) bromphenol blue in dH ₂ O	0.2 mL (can be stored at -20°C)
Glycerol	2.5 mL
dH ₂ O	3.55 mL
Prepare 950 µL aliquots and store at -20°C. Add 50 µL β-mercaptoethanol per aliquot before first use.	
Staining solution	
Coomassie Brilliant Blue R-250	1 g
Acetic acid, conc.	100 mL
Ethanol, 100 %	300 mL
dH ₂ O	600 mL
Dissolve Coomassie in ethanol first, add water and acetic acid last. Filtrate before first use.	
Destaining solution	
Acetic acid, conc.	100 mL
Ethanol, 100 %	300 mL
dH ₂ O	600 mL
Solution can be regenerated over charcoal after use.	

Table 53. Components of hand-casted SDS-gels (86x67 mm, 1 mm gap).

Running gel (12 % Bis-Tris)	<i>Amount per Gel</i>
Tris buffer A	2 mL
Acrylamide/Bis (30 %)	3.33 mL
dH ₂ O	2.76 mL
TEMED	4 µL
APS (10 %)	40 µL
Optional: Trichloro ethanol (TCE)	75 µL
Stacking gel (4 % Bis-Tris)	<i>Amount per Gel</i>
Tris buffer B	1 mL
Acrylamide/Bis (30 %)	0.53 mL
dH ₂ O	2.47 mL
TEMED	4 µL
APS (10 %)	40 µL

TLC staining solutions

Table 54. Staining solutions used for TLC.

KMnO₄ staining solution	
KMnO ₄	1.5 g
K ₂ CO ₃	10 g
ddH ₂ O	200 mL
NaOH, 10 % (w/v)	1.25 mL
Stir until dissolved and store solution away from light.	
CAM staining solution	
(NH ₄) ₆ Mo ₇ O ₂₄ ·4 H ₂ O	50 g
Ce(SO ₄) ₂	2 g
ddH ₂ O	400 mL
H ₂ SO ₄ , conc.	50 mL
Stir until dissolved.	
Ninhydrin staining solution	
Ninhydrin (final: 0.2 % w/v)	200 mg
Acetone	100 mL
Dissolve and store solution away from light at 4°C.	

Appendix C Primers

Table 55. List of various primers used in this work.

Name	Sequence (5'-3')	Purpose
T7_pET_mod	CCCGCGAAATTAATACGACTCAC	Sequencing of plasmids with T7 promoter
T7_term	CTAGTTATTGCTCAGCGGT	Sequencing of plasmids with T7 terminator
pJET1.2 forward	CGACTCACTATAGGGAGAGCGGC	cPCR or sequencing of pJET1.2 constructs
pJET1.2 reverse	AAGAACATCGATTTTCCATGGCAG	cPCR or sequencing of pJET1.2 constructs
sucA_upstr_fwd	GGCGAAGTAAGCATAAAAAAGATGCTTAAGGG	Detection of <i>sucA</i> gene
sucA_dwnstr_rev	GGACCAGAATATCTACGCTACTCATTGTGTATC	Detection of <i>sucA</i> gene
aceA_upstr_fwd	CTGTTAGCGTAAACCACCACATAACTATGGAG	Detection of <i>aceA</i> gene
aceA_dwnstr_rev	GCCTACAGTCAGCAACGGTTGTTG	Detection of <i>aceA</i> gene

Table 56. List of primers used for *BsPAD* mutagenesis. Mutated bases are shown in bold letters.

Name	Sequence (5'-3')	T _m [°C]
bsPAD_I85L_fwd	CATGGCATT CTG TTCTTCCCGAAATGGGTG	63.4
bsPAD_I85L_rev	GAAGAAC AGA ATGCCATGCATGCGTTTTTCATTG	63.3
bsPAD_I85V_fwd	CATGGCATT GTG TTCTTCCCGAAATGGGTG	63.8
bsPAD_I85V_rev	GAAGAAC ACA ATGCCATGCATGCGTTTTTCATTG	63.6

Table 57. List of primers used for *AMDase* mutagenesis. Mutated bases are shown in bold letters.

Name	Sequence (5'-3')	T _m [°C]
AMD_CLGIPL_I43A_fwd	CCTGGGATCC GCC ACTCCCGAAGGCTATG	68.3
AMD_CLGIPL_I43A_rev	GGGAGT GGC GGATCCCAGGCCCAAAC	68.4
AMD_CLGIPL_I43F_fwd	CCTGGGATC CTTCA CTCCCGAAGGCTATG	64.8
AMD_CLGIPL_I43F_rev	GGGAGT GA AGGATCCCAGGCCCAAAC	64.4
AMD_CLGIPL_I43L_fwd	CCTGGGATCC CTG ACTCCCGAAGGCTAT	65.6
AMD_CLGIPL_I43L_rev	GGAGT CAG GGATCCCAGGCCCAAACC	65.9
AMD_CLGIPL_I43M_fwd	CCTGGGATC CATG ACTCCCGAAGGCTATG	64.8
AMD_CLGIPL_I43M_rev	GGGAGT CAT GGATCCCAGGCCCAAAC	64.4
AMD_CLGIPL_I43V_fwd	CCTGGGATCC GTG ACTCCCGAAGGCTATG	66.4
AMD_CLGIPL_I43V_rev	GGGAGT CAC GGATCCCAGGCCCAAAC	66.3
AMD_CLGIPL_L156A_fwd	CCGGAG CCG GAAGCGCTGGCTC	66.7
AMD_CLGIPL_L156A_rev	CGCTT CGG CTCCGGTAATGCCTAAGCT	66.1
AMD_CLGIPL_L159M_fwd	GGAAGCG ATG GCTCGAGTGGATACCGCCAC	68.7
AMD_CLGIPL_L159M_rev	CTCGAG CCAT CGCTTCCAGTCCGGTAATGCC	68.5
AMD_CLGIPL_G188A_fwd	GTTGTCG GGC GGGCGGTCTGCTTACCTTG	68.5
AMD_CLGIPL_G188A_rev	GACCGCC GGC GACAACAGAATCCCATC	68.2
AMD_CLGIPL_G188S_fwd	GTTGTCG AGC GGGCGGTCTGCTTACCTTG	66.8
AMD_CLGIPL_G188S_rev	GACCGCC GCT CGACAACAGAATCCCATC	66.5
AMD_CLGIPL_G188A/G190A_fwd	GTTGTCG GGC GGGCGGTCTGCTTACCTTG	70.9
AMD_CLGIPL_G188A/G190A_rev	GCGCGCC GGC GACAACAGAATCCCATC	70.6
AMD_CLGIPL_G188S/G190A_fwd	GTTGTCG AGC GGGCGGTCTGCTTACCTTG	69.2
AMD_CLGIPL_G188S/G190A_rev	GCGCGCC GCT CGACAACAGAATCCCATC	68.9

Table 58. List of primers used for *AvLDO* mutagenesis. Mutated bases are shown in bold letters.

Name	Sequence (5'-3')	T _m [°C]
F ₀	CAAGAATCTTCCCTTAACAACAGAATCTTTAAAAATGCTCTGGGAAAACCAAATTCC	65.5
R _{1.1}	GTTATATT CAAN CACTGTGATGCCAATTCGCTC WRH TTTAG AKW CACATCTTGATAAC	61.8
R _{1.2}	GTTATATT CATA CACTGTGATGCCAATTCGCTC WRH TTTAG AKW CACATCTTGATAAC	61.9
F _{1.1}	GATGT GWMT CTCTAA ADY WGAGCGAATTGGCATCACAGT GN TTGAATATAACCAAATTAG	61.7
F _{1.2}	GATGT GWMT CTCTAA ADY WGAGCGAATTGGCATCACAGT G TATGAATATAACCAAATTAG	62.5
R ₂	CTTCCCATCCAGACT GMBY GTATGGAGCATAATCAAC	60.0
F ₂	GATTATGCTCCATAC RVK CAGTCTGGATGGGAAGTTTGTAC	60.4
R ₃	CAATGACTGTATCACTATAGCC MH AGGAATCGAGCTTATATTGGTCATCTC	63.9
F ₃	CAATATAAGCTCGATT CCTDK GGCTATAGTGATACAGTCATTGCAGATGC	64.4
R ₀	CACCGTTAGGAAGTAATCCCATCGCCGAAGTGAAA	65.8

Table 59. List of nucleotides and their complements.²⁸²

Symbol	Meaning	Description	Complement
G	G	Guanine	C
A	A	Adenine	T
T	T	Thymine	A
C	C	Cytosine	G
R	G or A	puRine	Y
Y	T or C	pYrimidine	R
M	A or C	aMino	K
K	G or T	Keto	M
S	G or C	Strong	S
W	A or T	Weak	W
H	A or C or T	not-G	D
B	G or T or C	not-A	V
V	G or C or A	not-T	B
D	G or A or T	not-C	H
N	G or A or T or C	aNy	N

Appendix D DNA and protein sequences

Underlined parts of the sequences shown were added through the cloning process.

DNA and protein sequence of *BsPAD*

ATGGGCAGCAGCCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGAAAAC
TTATCGGAAGCCACATGATTTATACGTATGAAAACGGATGGGAATACGAGATTTATATTA
TACAATTGATTATAGAATTCATAGCGGAATGGTTGCCGACGCTGGGTTTCGAGATCAGGAAGTGAATATT
GTCAAACGACAGAAGGCGTATATAAAGTGTCTTGACAGAGCCGACTGGCACGGATGTTTCATTA
TTATGCCAAATGAAAACGCATGCATGGCATTaTTTTCTTCCCGAAATGGGTGCATGAACATCCTGAAAT
TACGGTTTGCTACCAAATGACCACATTGATTTGATGAAAGAATCCCGCGAAAAATATGAAACGTATCCA
AAATACGTTGTACCTGAATTTGCGGAAATTACATTTCTGAAAAATGAAGGAGTCGACAACGAAGAAGTGA
TTTCGAAGGCTCCTTATGAGGGAATGACAGACGATATTCGCGCGGGAAGATTATAA

MGSSHHHHHSSGLVPRGSHMENFIGSHMIYTYENGWEYEIYIKNDHTIDYRIHSGMVAGRWRDQEVNI
VKLTGEGVYKVSWTEPTGTDVSLNFMFPNEKRMHGIIFPKVWEHPEITVCYQNDHIDLKESREKYETYP
KYVVPFAEITFLKNEGVDNEEVIKAPYEGMTDDIRAGRL*

DNA and protein sequence of *AMDase*

ATGGGCCAAATGCAACAGGCAAGCACCCCGACCATTGGTATGATCGTACCTCCCGCAGCTGGACTAGTGC
CAGCCGACGCGCTCGTCTGTATCCGGATTTGCCGTTTATCGCGTCGGGTTTGGGCCTGGGATCCGTGAC
TCCCGAAGGCTATGACGCGTTATAGAAAGCGTGGTTGATCATGCTCGTCGCTGCAAAAACAGGGGGCA
GCCGTTGTGTCTCTCATGGGGACCTCCTTGTCTGTTCTACCGCGCGCCGCTTTTAAACGGCGCTGACTG
TCGCCATGCGTGAGGCTACTGGCCTGCCGTGACTACCATGAGTACCGCCGTGCTAAACGGCCTGCGTGC
ACTGGGCGTCCGGCGTGTGGCACTGGCGACCGCCTATATCGATGATGTTAATGAACGGCTTGCAGCGTTT
CTGGCGGAAGAATCCCTGGTACCGACAGGTTGTCTGACTTAGGCATTACCGGAGTAGAAGCGATGGCTC
GAGTGGATACCGCCACTCTGGTGCATCTGTGCGTCCGCGCCTTTGAAGCAGCACCCAGATAGCGATGGGAT
TCTGTTGTGCTGTGGCGGTCTGCTTACCTTGGACGCAATCCCGGAAGTCGAGCGTCGCTGGGTGTGCCA
GTCGTCTCAAGCAGTCCGGCGGGCTTTTGGGATGCGGTCCGTTTGGCTGGCGGAGGCGCCAAAGCACGCC
CGGGTTACGGCCGGCTTTTTGATGAGTCCGGTGGCAGCCACCATCACCATCACCATTAA

MGQMQQASTPTIGMIVPPAAGLVPADGARLYPDLPIASGLGLGSVTPEGYDAVIESVVDHARRLQKQGA
AVVSLMGTSLSFYRGAAFNAALTVMREATGLPCTTMSTAVLNGLRALGVRVALATAYIDDVNERLAAF
LAEESLVPTGCRSLGITGVEAMARVDTATLVDL CVRAFEAAPDSGILLSCGGLLTLDAIPEVERRLGVP
VVSSSPAGFWDAVRLAGGGAKARPGYGRLFDESGGSHHHHHH*

DNA and protein sequence of AvLDO

ATGGGCAGCAGCCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGACTGCTA
CTTCTCAACAGCTTAAGTCCAAAGTATGGAAAAACAATCAAGAATCTTCCTTAACAACAGAATCTTTAAA
AATGCTCTGGGAAAACCAAATTCCTTTAATTCGACTCCAGGAATTTGCTACACCCCAAGAGTGTGAAAAC
TTAGTAGCCAGGCTCAATTATTTAACTTTGATAGTTATCAAGATGTGTATCCTAAAATTGAGCGAATTG
GCATCACAGTGTGTTGAATATAACCAAATTAGCAAGGCGGGTTATTTTCAAGCAGTAGAACAAGCAACTGA
ATTAAGAGACTTGATTTTGGCAGCCTCTTTTAAATCCATTAGAACGCCTAATGGTGAAAATTAGAGAGTGT
ACAGGCGCGACTGTGCGTATTGCTTCGGAGCTATTATATGGTGTTATTATGCAGGACTGATCAGAAAAA
TAGAGCAGGGTACTCAACTTCATGTTGATTATGCTCCATACGAGCAGTCTGGATGGGAAGTTTGTACAAT
TACTGCTCAACTTTCTTGAATTTGTATTTGAAGAAAACGCCAGTAGTCATGGTAAAACGCGTATTTAC
GATCGCCAATGGCAACCAGGAGATGACCAATATAAGCTCGATTCCCTACGGCTATAGTGATACAGTCATTG
CAGATGCAGACATGATTACTTTCCAACCTCATGTGGGAGATGTGTTTCAATTTTAAATACGCGCAACTATCA
CATTGTTGAGCCGATGGATGGACAACGTATAACTTTCACTTCGGCGATGGGATTACTTCCTAACGGTGAA
ATTATTTTGTGGTCTTAA

MGSSHHHHHSSGLVPRGSHMTATSQQLKSKVWKNQESSLTTESLKMLWENQIPLIRLQEFATPQECEN
LVAQAQLFNFDYQDVYPKIERIGITVFEYNQISKAGYFQAVEQATELRDLILAASFNPLERLMVKIREC
TGATVRIASELLYGGYYAGLIRKIEQGTQLHVDYAPYEQSGWEVCTITAQLSWNLYLKKTASSHGKTRIY
DRQWQPGDDQYKLDYGYSDTVIADADMITFQPHVGDVFIENRNYHIVEPMDGQRITFTSAMGLLPNGE
IILWS*

Appendix E Agarose gels

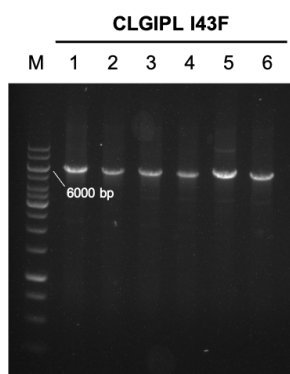


Figure S 1. Agarose gel analysis of QuikChange reaction for generation of AMDase CLGIPL I43F with temperature gradient ($T_1 = 63^\circ\text{C}$, $T_2 = 63.6^\circ\text{C}$, $T_3 = 65.3^\circ\text{C}$, $T_4 = 67.7^\circ\text{C}$, $T_5 = 69.3^\circ\text{C}$, $T_6 = 70^\circ\text{C}$). pET28a[AMDase_CLGIPL] (6005 bp) was used as template. M: GeneRuler 1 kb DNA Ladder (Thermo Scientific).

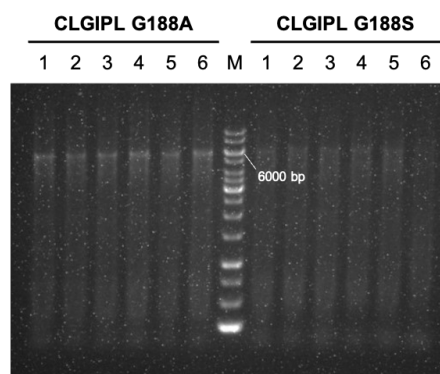


Figure S 2. Agarose gel analysis of QuikChange reaction for generation of AMDase CLGIPL G188A (left) and G188S (right) with temperature gradient ($T_1 = 63^\circ\text{C}$, $T_2 = 63.6^\circ\text{C}$, $T_3 = 65.3^\circ\text{C}$, $T_4 = 67.7^\circ\text{C}$, $T_5 = 69.3^\circ\text{C}$, $T_6 = 70^\circ\text{C}$). pET28a[AMDase_CLGIPL] (6005 bp) was used as template. M: GeneRuler 1 kb DNA Ladder (Thermo Scientific).

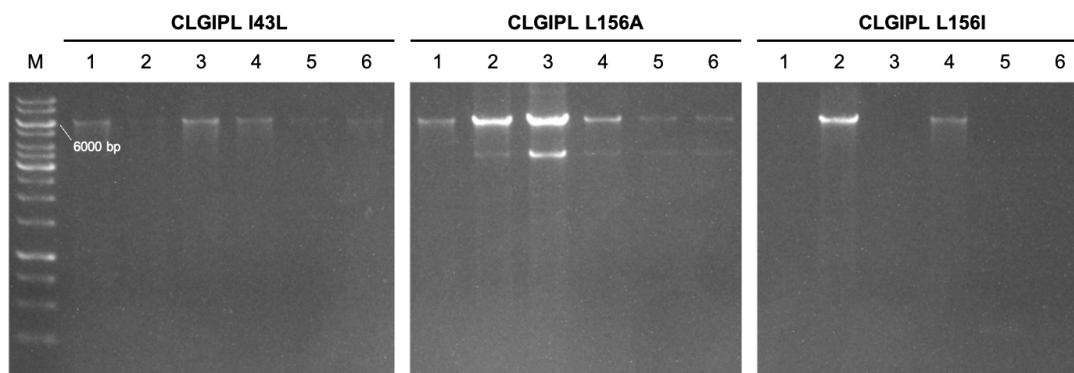


Figure S 3. Agarose gel analysis of QuikChange reaction for generation of AMDase CLGIPL I43L (left), L156A (center) and L156I (right) with temperature gradient ($T_1 = 63^\circ\text{C}$, $T_2 = 63.6^\circ\text{C}$, $T_3 = 65.3^\circ\text{C}$, $T_4 = 67.7^\circ\text{C}$, $T_5 = 69.3^\circ\text{C}$, $T_6 = 70^\circ\text{C}$). pET28a[AMDase_CLGIPL] (6005 bp) was used as template. M: GeneRuler 1 kb DNA Ladder (Thermo Scientific).

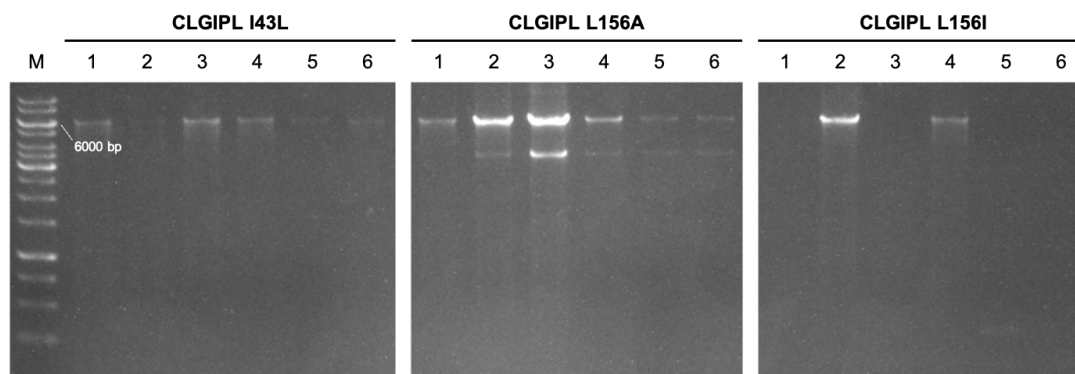


Figure S 4. Agarose gel analysis of QuikChange reaction for generation of AMDase CLGIPL I43L (left), L156A (center) and L156I (right) with temperature gradient ($T_1 = 63^\circ\text{C}$, $T_2 = 63.6^\circ\text{C}$, $T_3 = 65.3^\circ\text{C}$, $T_4 = 67.7^\circ\text{C}$, $T_5 = 69.3^\circ\text{C}$, $T_6 = 70^\circ\text{C}$). pET28a[AMDase_CLGIPL] (6005 bp) was used as template. M: GeneRuler 1 kb DNA Ladder (Thermo Scientific).

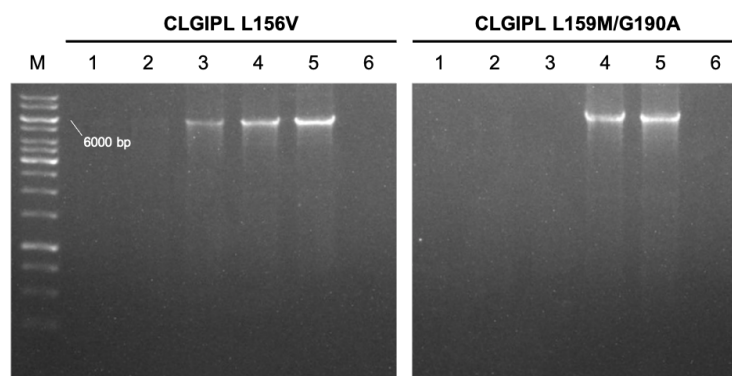


Figure S 5. Agarose gel analysis of QuikChange reaction for generation of AMDase CLGIPL L156V (left) and L159M/G190A (right) with temperature gradient ($T_1 = 63^\circ\text{C}$, $T_2 = 63.6^\circ\text{C}$, $T_3 = 65.3^\circ\text{C}$, $T_4 = 67.7^\circ\text{C}$, $T_5 = 69.3^\circ\text{C}$, $T_6 = 70^\circ\text{C}$). pET28a[AMDase_CLGIPL] (for L156V) and pET28a[AMDase_CLGIPL_L159M] (for L159M/G190A) (6005 bp) were used as template. M: GeneRuler 1 kb DNA Ladder (Thermo Scientific).

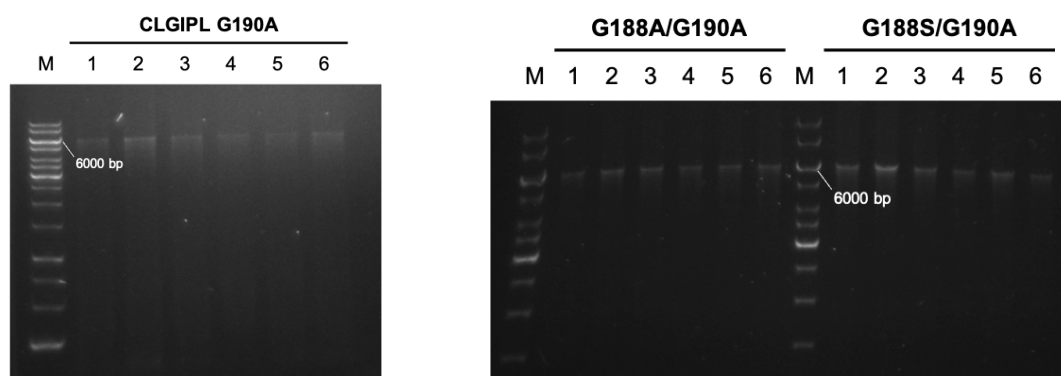


Figure S 6. Agarose gel analysis of QuikChange reaction for generation of AMDase CLGIPL G190A with temperature gradient ($T_1 = 63^\circ\text{C}$, $T_2 = 63.6^\circ\text{C}$, $T_3 = 65.3^\circ\text{C}$, $T_4 = 67.7^\circ\text{C}$, $T_5 = 69.3^\circ\text{C}$, $T_6 = 70^\circ\text{C}$). pET28a[AMDase_CLGIPL] (6005 bp) was used as template. M: GeneRuler 1 kb DNA Ladder (Thermo Scientific).

Figure S 7. Agarose gel analysis of QuikChange reaction for generation of AMDase CLGIPL G188A/G190A (left) and G188S/G190A (right) with temperature gradient ($T_1 = 63^\circ\text{C}$, $T_2 = 63.6^\circ\text{C}$, $T_3 = 65.3^\circ\text{C}$, $T_4 = 67.7^\circ\text{C}$, $T_5 = 69.3^\circ\text{C}$, $T_6 = 70^\circ\text{C}$). pET28a[AMDase_CLGIPL] (6005 bp) was used as template. M: GeneRuler 1 kb DNA Ladder (Thermo Scientific).

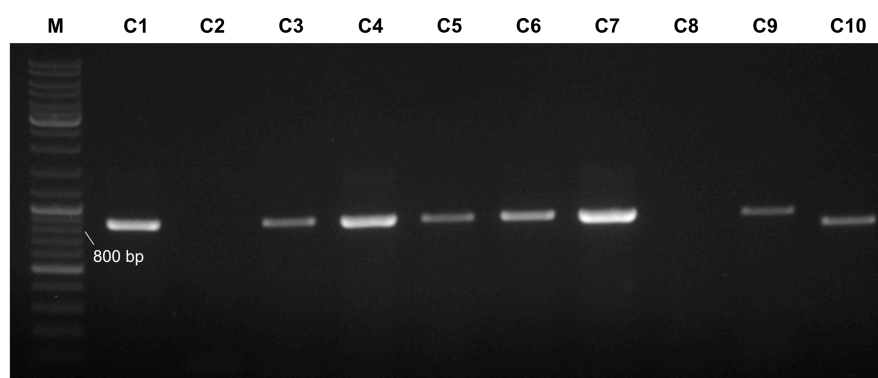


Figure S 8. Agarose gel analysis of results obtained from screening 10 transformants (C1-C10) via colony PCR. The correct construct pJET1.2_AvLDO_megaprimer would lead to a signal at 850 bp after amplification with pJET2.1 forward and reverse primer. M: GeneRuler DNA Ladder Mix (Thermo Scientific).

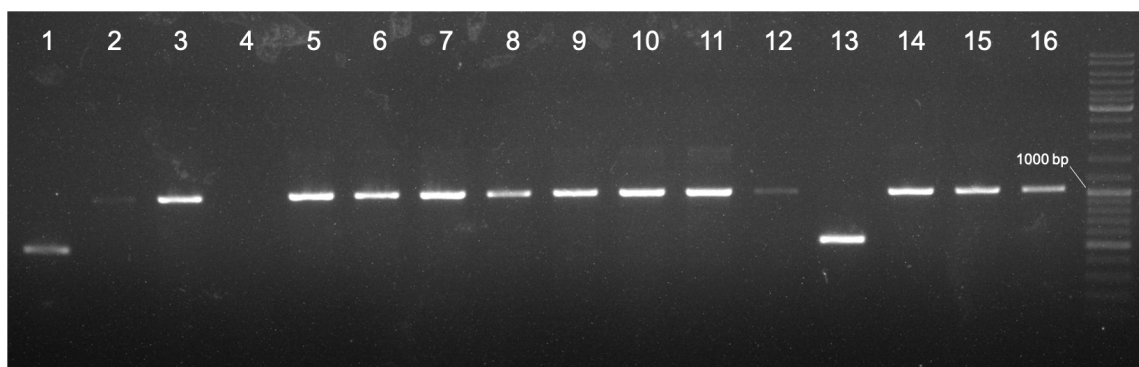


Figure S 9. Agarose gel analysis of results obtained from screening 16 clones via colony PCR after pET28a_AvLDO_Y67X/I70X/F78X/E158X/Y206X library transformation. Amplification with T7_pET_mod and T7_term should result in a fragment of 1049 bp. Slot 1 and 13 correspond to clone 7-8 in Table 26. M: GeneRuler DNA Ladder Mix (Thermo Scientific).

Appendix F SDS gels

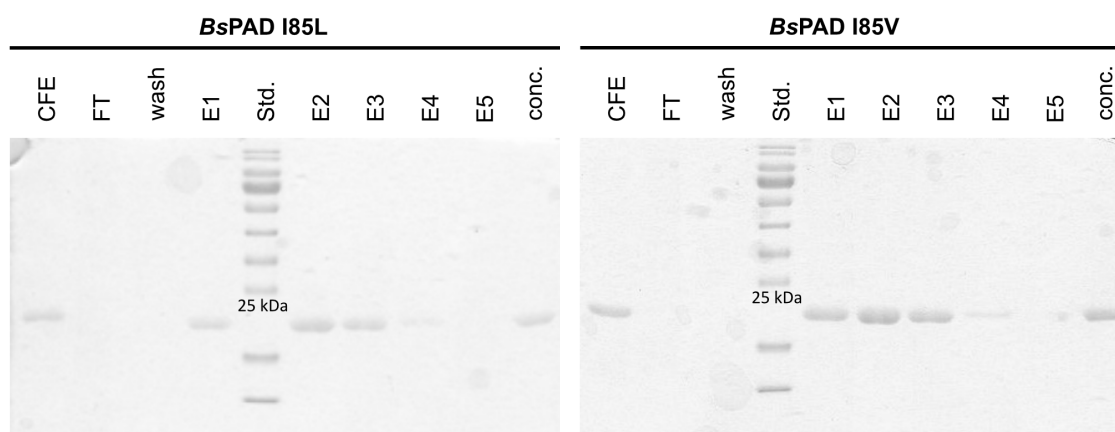


Figure S 10. SDS-PAGE analysis of different fractions obtained from Ni-affinity purification of BsPAD I85L (MW = 19.1 kDa). CFE: cell-free extract, FT: flow through, E: elution fraction, Conc.: desalted and concentrated combined elution fractions (E1-E7). PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

Figure S 11. SDS-PAGE analysis of different fractions obtained from Ni-affinity purification of BsPAD I85V (MW = 19.1 kDa). CFE: cell-free extract, FT: flow through, E: elution fraction, Conc.: desalted and concentrated combined elution fractions (E1-E7). PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

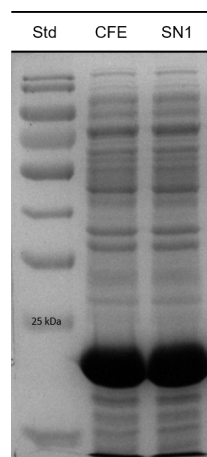


Figure S 12. SDS-PAGE analysis of *BsPAD*-containing cell-free extract (CFE) and supernatant fraction (SN1) obtained after binding to EziG™ 2 carrier material. Results correspond to Table 6, entry 1. *BsPAD* (19.1 kDa) appears slightly below the 25 kDa band of the standard (PageRuler Prestained Protein Ladder, Thermo Scientific).

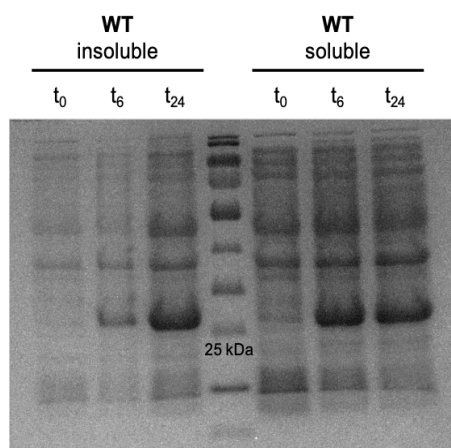


Figure S 13. SDS-PAGE analysis of AMDase WT (MW = 25.8 kDa) expression study in LB medium. Samples were taken 0 h (t₀), 6 h (t₆) and 24 h (t₂₄) after induction. Soluble and insoluble fraction were separated after sonication. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

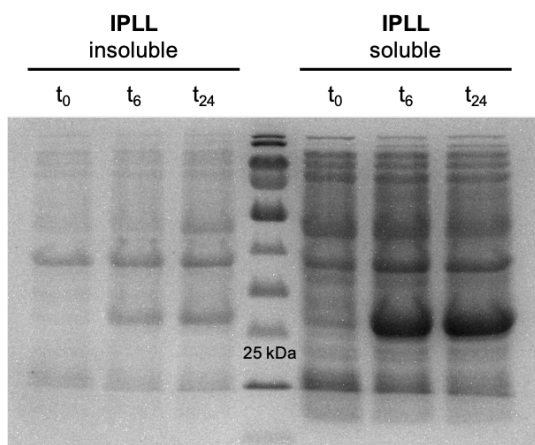


Figure S 14. SDS-PAGE analysis of AMDase IPLL (MW = 25.8 kDa) expression study in LB medium. Samples were taken 0 h (t₀), 6 h (t₆) and 24 h (t₂₄) after induction. Soluble and insoluble fraction were separated after sonication. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

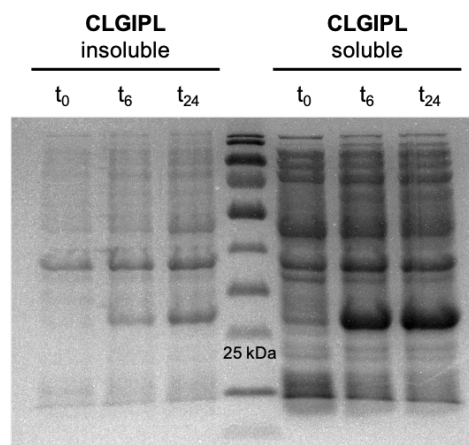


Figure S 15. SDS-PAGE analysis of AMDase CLGIPL (MW = 25.8 kDa) expression study in LB medium. Samples were taken 0 h (t₀), 6 h (t₆) and 24 h (t₂₄) after induction. Soluble and insoluble fraction were separated after sonication. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

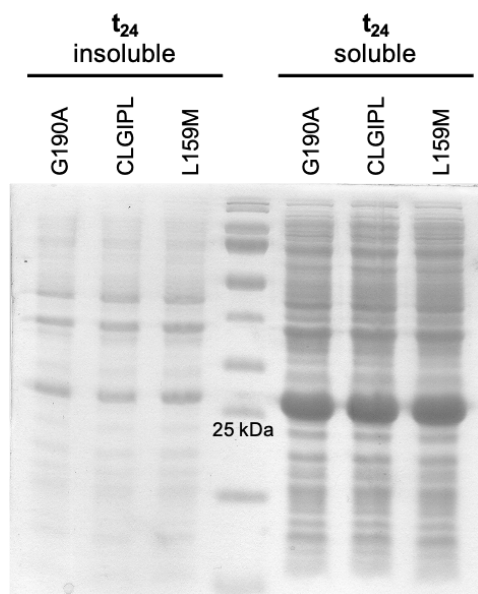


Figure S 16. SDS-PAGE analysis of AMDase CLGIPL G190A, CLGIPL and CLGIPL L159M (MW = 25.8 kDa) expression study in LB medium. Samples were taken 24 h (t_{24}) after induction. Soluble and insoluble fraction were separated after sonication. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

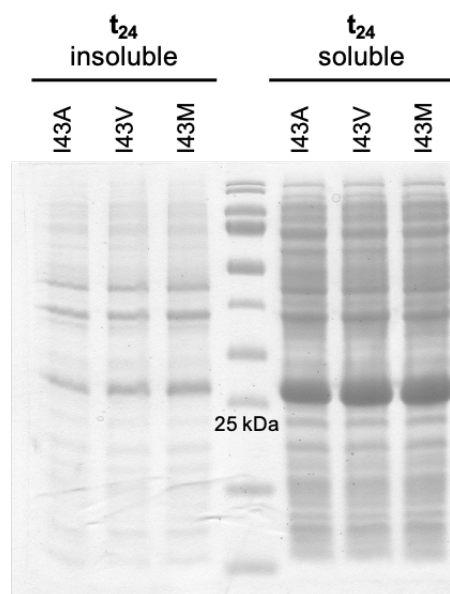


Figure S 17. SDS-PAGE analysis of AMDase CLGIPL I43A, CLGIPL I43V and CLGIPL I43M (MW = 25.8 kDa) expression study in LB medium. Samples were taken 24 h (t_{24}) after induction. Soluble and insoluble fraction were separated after sonication. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

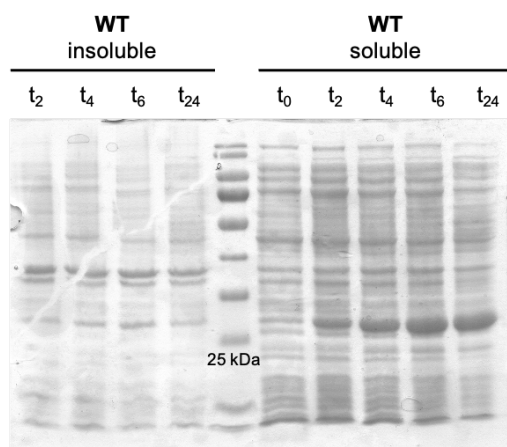


Figure S 18. SDS-PAGE analysis of AMDase WT (MW = 25.8 kDa) expression study in TB medium. Samples were taken 0 h (t_0), 2 h (t_2), 4 h (t_4), 6 h (t_6) and 24 h (t_{24}) after induction. Soluble and insoluble fraction were separated after sonication. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

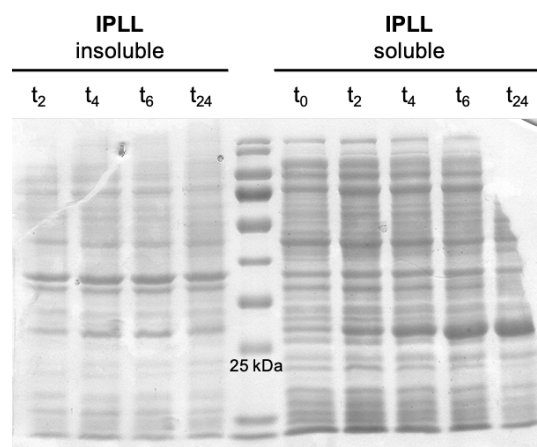


Figure S 19. SDS-PAGE analysis of AMDase IPLL (MW = 25.8 kDa) expression study in TB medium. Samples were taken 0 h (t_0), 2 h (t_2), 4 h (t_4), 6 h (t_6) and 24 h (t_{24}) after induction. Soluble and insoluble fraction were separated after sonication. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

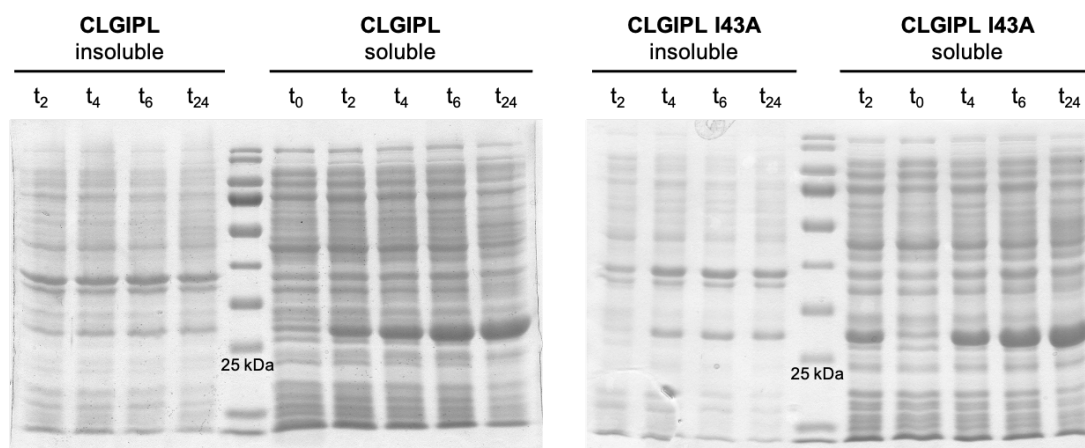


Figure S 20. SDS-PAGE analysis of AMDase CLGIPL (MW = 25.8 kDa) expression study in TB medium. Samples were taken 0 h (t_0), 2 h (t_2), 4 h (t_4), 6 h (t_6) and 24 h (t_{24}) after induction. Soluble and insoluble fraction were separated after sonication. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

Figure S 21. SDS-PAGE analysis of AMDase CLGIPL I43A (MW = 25.8 kDa) expression study in TB medium. Samples were taken 0 h (t_0), 2 h (t_2), 4 h (t_4), 6 h (t_6) and 24 h (t_{24}) after induction. Soluble and insoluble fraction were separated after sonication. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

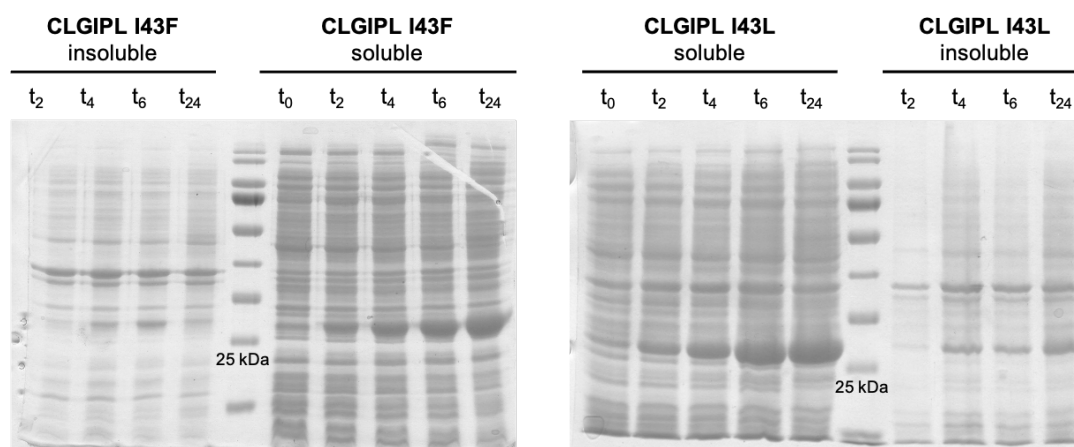


Figure S 22. SDS-PAGE analysis of AMDase CLGIPL I43F (MW = 25.8 kDa) expression study in TB medium. Samples were taken 0 h (t_0), 2 h (t_2), 4 h (t_4), 6 h (t_6) and 24 h (t_{24}) after induction. Soluble and insoluble fraction were separated after sonication. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

Figure S 23. SDS-PAGE analysis of AMDase CLGIPL I43L (MW = 25.8 kDa) expression study in TB medium. Samples were taken 0 h (t_0), 2 h (t_2), 4 h (t_4), 6 h (t_6) and 24 h (t_{24}) after induction. Soluble and insoluble fraction were separated after sonication. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

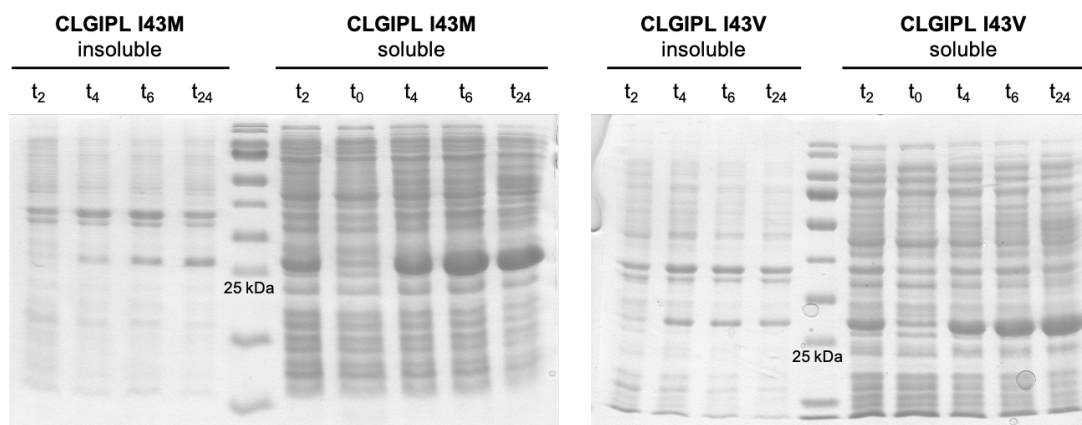


Figure S 24. SDS-PAGE analysis of AMDase CLGIPL I43M (MW = 25.8 kDa) expression study in TB medium. Samples were taken 0 h (t₀), 2 h (t₂), 4 h (t₄), 6 h (t₆) and 24 h (t₂₄) after induction. Soluble and insoluble fraction were separated after sonication. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

Figure S 25. SDS-PAGE analysis of AMDase CLGIPL I43V (MW = 25.8 kDa) expression study in TB medium. Samples were taken 0 h (t₀), 2 h (t₂), 4 h (t₄), 6 h (t₆) and 24 h (t₂₄) after induction. Soluble and insoluble fraction were separated after sonication. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

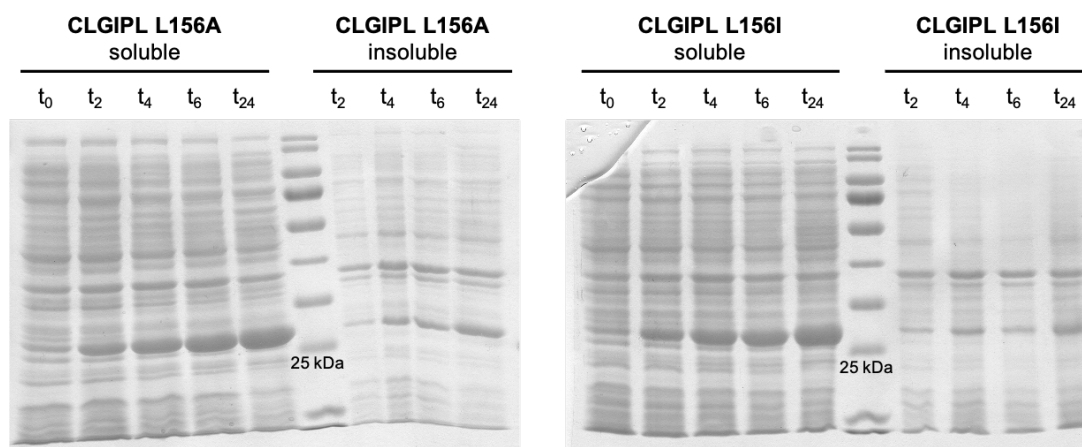


Figure S 26. SDS-PAGE analysis of AMDase CLGIPL L156A (MW = 25.8 kDa) expression study in TB medium. Samples were taken 0 h (t₀), 2 h (t₂), 4 h (t₄), 6 h (t₆) and 24 h (t₂₄) after induction. Soluble and insoluble fraction were separated after sonication. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

Figure S 27. SDS-PAGE analysis of AMDase CLGIPL L156I (MW = 25.8 kDa) expression study in TB medium. Samples were taken 0 h (t₀), 2 h (t₂), 4 h (t₄), 6 h (t₆) and 24 h (t₂₄) after induction. Soluble and insoluble fraction were separated after sonication. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

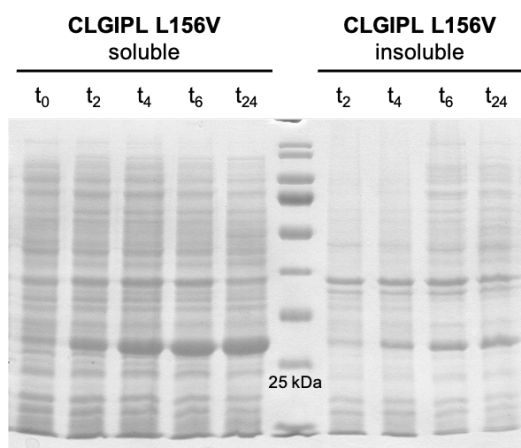


Figure S 28. SDS-PAGE analysis of AMDase CLGIPL L156V (MW = 25.8 kDa) expression study in TB medium. Samples were taken 0 h (t_0), 2 h (t_2), 4 h (t_4), 6 h (t_6) and 24 h (t_{24}) after induction. Soluble and insoluble fraction were separated after sonication. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

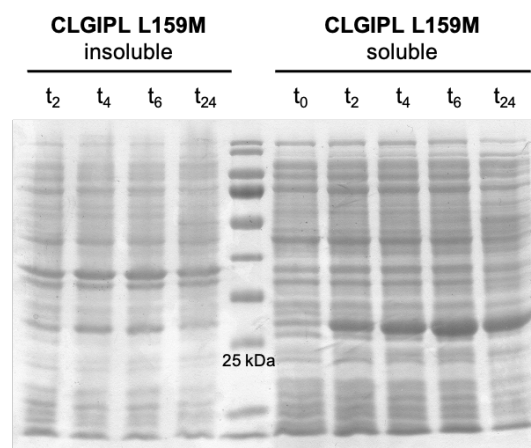


Figure S 29. SDS-PAGE analysis of AMDase CLGIPL L159M (MW = 25.8 kDa) expression study in TB medium. Samples were taken 0 h (t_0), 2 h (t_2), 4 h (t_4), 6 h (t_6) and 24 h (t_{24}) after induction. Soluble and insoluble fraction were separated after sonication. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

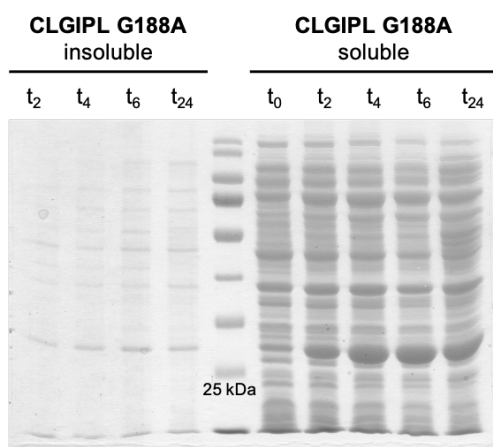


Figure S 30. SDS-PAGE analysis of AMDase CLGIPL G188A (MW = 25.8 kDa) expression study in TB medium. Samples were taken 0 h (t_0), 2 h (t_2), 4 h (t_4), 6 h (t_6) and 24 h (t_{24}) after induction. Soluble and insoluble fraction were separated after sonication. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

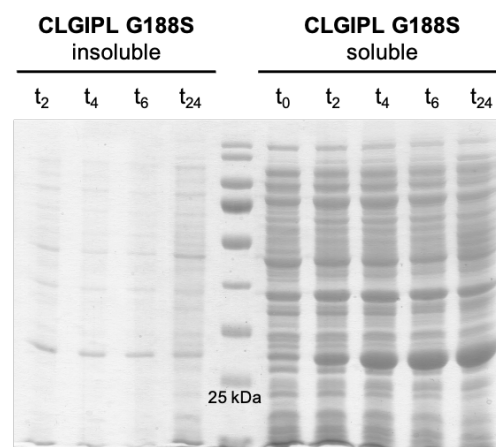


Figure S 31. SDS-PAGE analysis of AMDase CLGIPL G188S (MW = 25.8 kDa) expression study in TB medium. Samples were taken 0 h (t_0), 2 h (t_2), 4 h (t_4), 6 h (t_6) and 24 h (t_{24}) after induction. Soluble and insoluble fraction were separated after sonication. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

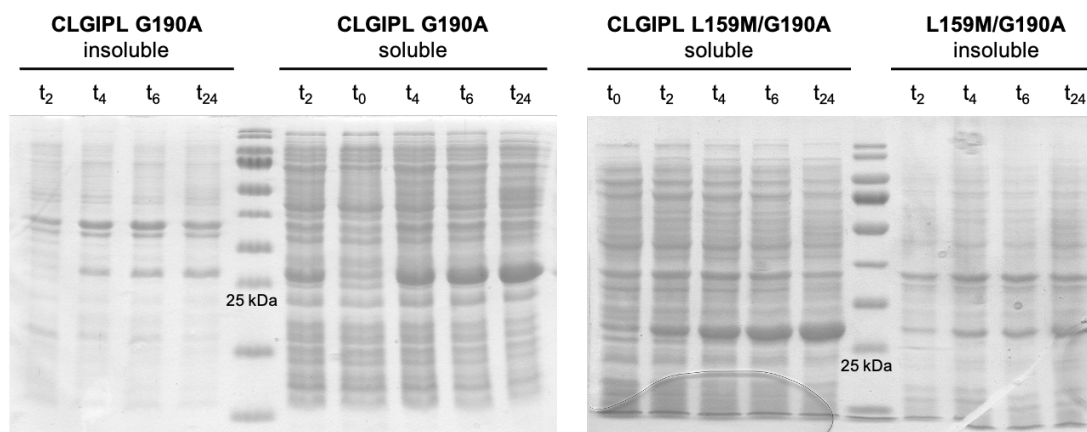


Figure S 32. SDS-PAGE analysis of AMDase CLGIPL G190A (MW = 25.8 kDa) expression study in TB medium. Samples were taken 0 h (t₀), 2 h (t₂), 4 h (t₄), 6 h (t₆) and 24 h (t₂₄) after induction. Soluble and insoluble fraction were separated after sonication. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

Figure S 33. SDS-PAGE analysis of AMDase CLGIPL L159M/G190A (MW = 25.8 kDa) expression study in TB medium. Samples were taken 0 h (t₀), 2 h (t₂), 4 h (t₄), 6 h (t₆) and 24 h (t₂₄) after induction. Soluble and insoluble fraction were separated after sonication. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

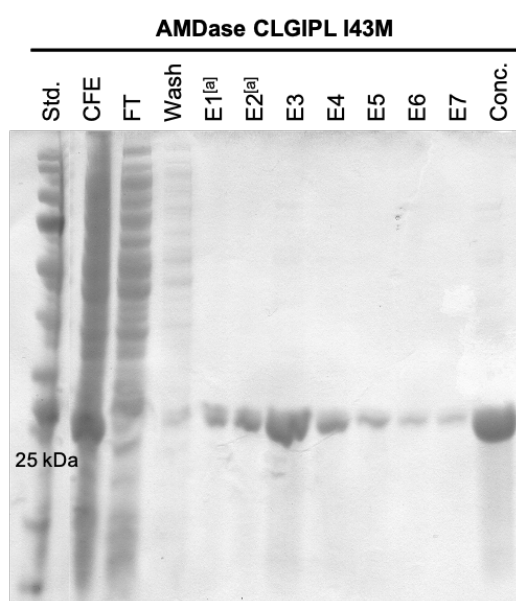


Figure S 34. SDS-PAGE analysis of different fractions obtained from Ni-affinity purification of AMDase CLGIPL I43M (MW = 25.8 kDa). CFE: cell-free extract, FT: flow through, E: elution fraction, Conc.: desalted and concentrated combined elution fractions (E1-E7). PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard. [a] diluted 1:20.

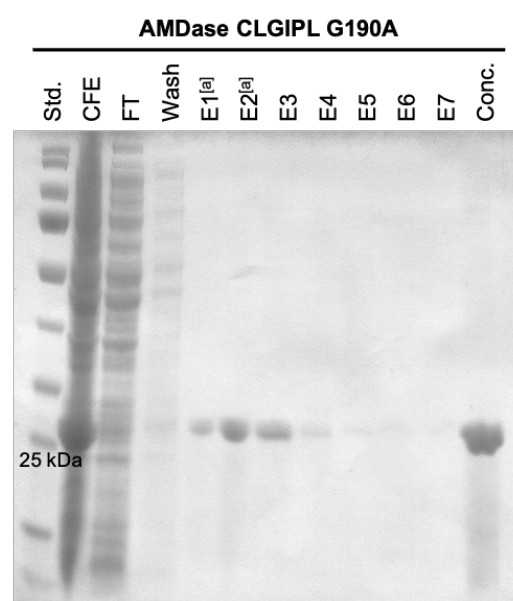


Figure S 35. SDS-PAGE analysis of different fractions obtained from Ni-affinity purification of AMDase CLGIPL G190A (MW = 25.8 kDa). CFE: cell-free extract, FT: flow through, E: elution fraction, Conc.: desalted and concentrated combined elution fractions (E1-E7). PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard. [a] diluted 1:20.

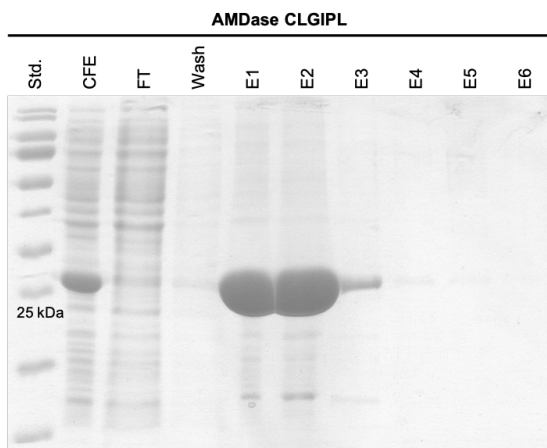


Figure S 36. SDS-PAGE analysis of different fractions obtained from Ni-affinity purification of AMDase CLGIPL (MW = 25.8 kDa). CFE: cell-free extract, FT: flow through, E: elution fraction. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

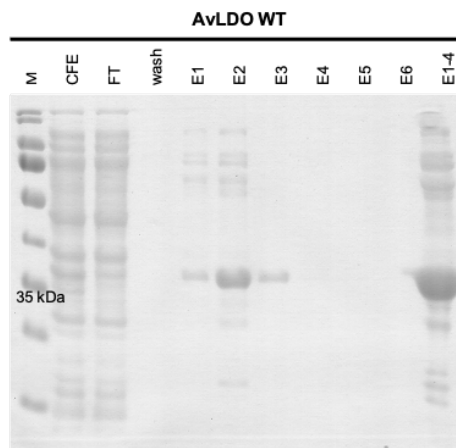


Figure S 37. SDS-PAGE analysis of different fractions obtained from Ni-affinity purification of AvLDO WT (MW = 32.5 kDa). CFE: cell-free extract, FT: flow through, E: elution fraction. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

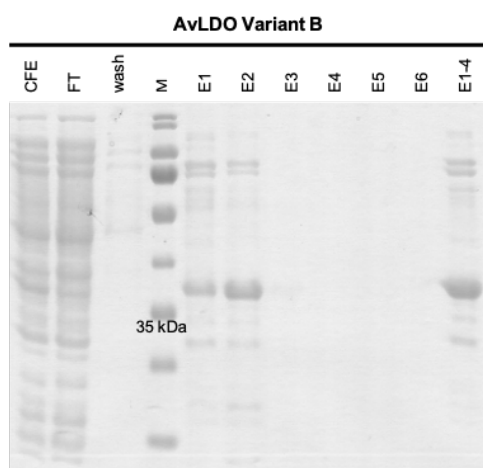


Figure S 38. SDS-PAGE analysis of different fractions obtained from Ni-affinity purification of AvLDO variant B (MW = 32.5 kDa). CFE: cell-free extract, FT: flow through, E: elution fraction. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

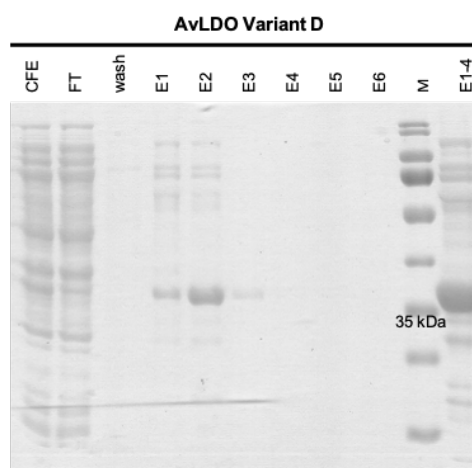


Figure S 39. SDS-PAGE analysis of different fractions obtained from Ni-affinity purification of AvLDO variant D (MW = 32.5 kDa). CFE: cell-free extract, FT: flow through, E: elution fraction. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

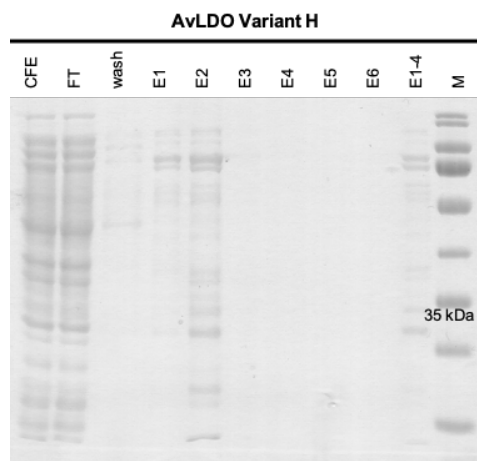
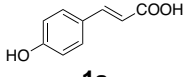
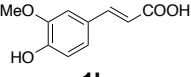
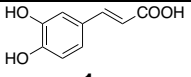
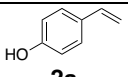
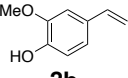
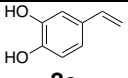
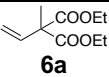
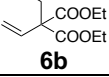
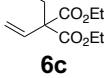
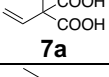
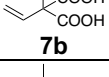
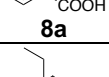
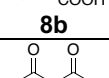
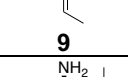
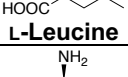
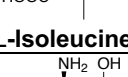
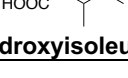


Figure S 40. SDS-PAGE analysis of different fractions obtained from Ni-affinity purification of AvLDO variant H (MW = 32.5 kDa). CFE: cell-free extract, FT: flow through, E: elution fraction. PageRuler Prestained Protein Ladder (Thermo Scientific) was used as standard.

Appendix G TLC

Table 60. Mobile phase, detection method and R_f values of compounds analyzed *via* TLC in this work.

Compound	Mobile phase	Detection	R_f
 1a	CH/EA = 1:1	UV, CAM	0.36
 1b	CH/EA/AcOH = 1:1:0.001	UV, CAM	0.46
 1c	CH/EA = 1:1	UV, CAM	0.14
 2a	CH/EA = 1:1	UV, CAM	0.78
 2b	CH/EA/AcOH = 1:1:0.001	UV, CAM	0.92
 2c	CH/EA = 1:1	UV, CAM	0.62
 6a	CH/EA = 5:1	KMnO ₄	0.56
 6b	CH/EA = 5:1	KMnO ₄	0.62
 6c	CH/EA = 5:1	KMnO ₄	0.50
 7a	CH/EA/AcOH = 1:1:0.001	KMnO ₄	0.10
 7b	CH/EA/AcOH = 1:1:0.001	KMnO ₄	0.18
 8a	CH/EA/AcOH = 1:1:0.001	KMnO ₄	0.58
 8b	CH/EA/AcOH = 1:1:0.001	KMnO ₄	0.68
 9	CH/EA = 5:1	KMnO ₄	0.44
 L-Leucine	1-butanol/AcOH/H ₂ O = 22:5:3	Ninhydrin	0.33
	acetone/1-butanol/AcOH/H ₂ O = 7:7:2:4	Ninhydrin	0.64
 L-Isoleucine	1-butanol/AcOH/H ₂ O = 22:5:3	Ninhydrin	0.36
	acetone/1-butanol/AcOH/H ₂ O = 7:7:2:4	Ninhydrin	0.60
 4-Hydroxyisoleucine	1-butanol/AcOH/H ₂ O = 22:5:3	Ninhydrin	0.18
	acetone/1-butanol/AcOH/H ₂ O = 7:7:2:4	Ninhydrin	0.50

Appendix H Raw data

Table 61. Determined enantiomeric excess of the product 2-methylbut-3-enoic acid (**8a**) obtained from AMDase CLGIPL I43M- or G190A-catalyzed decarboxylation of 2-methyl-2-vinylmalonic acid (**7a**) at different reaction temperatures. Data is presented in Figure 29 of the main text.

Reaction temperature	CLGIPL I43M	CLGIPL G190A
40°C	72 %ee (S)	41 %ee (S)
30°C	76 %ee (S)	47 %ee (S)
20°C	76 %ee (S)	53 %ee (S)
10°C	79 %ee (S)	56 %ee (S)

Table 62. Optical density (OD_{600}) determined during preliminary enrichment experiments of *E. coli* Δ sucA harboring either AvLDO WT or AvLDO Y67X/I70X/F78X/E158X/Y206X library. Conditions A-D and E-F are shown in Table 27 and Table 28. Data is presented in Figure 43- Figure 44.

	A	B	C		D		E		F	
	OD_{600} (Library)	OD_{600} (Library)	OD_{600} (WT)	OD_{600} (Library)	OD_{600} (WT)	OD_{600} (Library)	OD_{600} (WT)	OD_{600} (Library)	OD_{600} (WT)	OD_{600} (Library)
t_0	0.52	0.52	0.32	0.18	0.21	0.16	0.08	0.16	0.12	0.38
t_1	0.06	0.07	0.13	0.04	0.05	0.01	0.07	0.30	0.25	0.63
t_2	0.06	0.07	0.12	0.07	0.07	0.02			0.59	0.73
t_3		0.05	0.13	0.15	0.07	0.03			0.69	0.83
t_4		0.05	0.26	0.11	0.09	0.10			0.93	1.48
t_5					4.21	2.87			1.52	2.35
t_6					3.96	1.84			2.53	3.01
t_7					3.95	3.81				
t_8					3.20	4.13				
t_9					4.47	3.98				
t_{10}					3.97	3.94				
t_{11}					4.18	4.10				
t_{12}					3.94	4.02				
t_{13}					4.06	4.23				

Table 63. Determined optical density (OD_{600}) of *E. coli* Δ sucA harboring BflDO, AvLDO WT or variants B, D and H after two consecutive cultivation intervals in M9-Kan⁴⁰/IPTG/ α KG containing either L-Leu or L-Ile (5 mL, 37°C). The data is presented in Figure 46 of the main text.

Clone	L-Isoleucine containing M9		L-Leucine containing M9	
	OD_{600} (t_1)	OD_{600} (t_2)	OD_{600} (t_1)	OD_{600} (t_2)
BflDO	0.036	0.304	0.103	0.780
AvLDO WT	0.104	0.022	1.25	1.71
AvLDO B	0.521	0.559	0.501	0.882
AvLDO D	0.074	0.030	0.058	0.030
AvLDO H	1.40	1.34	0.941	1.275

Table 64. Determined optical density (OD_{600}) of *E. coli* $3\Delta sucA$ harboring AvLDO WT or variants 1-35 after two consecutive cultivation intervals in M9-Kan⁴⁰/IPTG/ α KG/L-Ile (5 mL, 37°C). The data is presented in Figure 48 of the main text.

Clone	OD_{600} (t_1)	OD_{600} (t_2)	Clone	OD_{600} (t_1)	OD_{600} (t_2)
WT1	0.015	0.07	18	0.834	0.858
WT2	0.04	0.003	19	0.836	0.886
1	0.895	0.905	20	0.814	0.871
2	0.621	0.723	21	0.817	0.861
3	0.851	0.892	22	0.782	1.015
4	0.941	0.946	23	0.879	0.937
5	0.803	0.918	24	0.849	0.945
6	0.911	0.925	25	0.825	0.908
7	0.866	0.9	26	0.835	0.931
8	0.908	0.897	27	0.827	0.913
9	0.416	0.422	28	0.844	0.893
10	0.905	0.895	29	0.819	0.933
11	0.909	0.938	30	0.825	0.895
12	0.897	0.947	31	0.856	0.939
13	0.49	0.692	32	0.876	0.993
14	0.866	0.898	33	0.82	0.89
15	0.842	0.9	34	0.865	0.941
16	0.888	0.889	35	0.75	0.826
17	0.883	0.887			

Appendix I GC supplementary data

Table 65. Detailed parameters of methods used for GC-MS analyses (see 7.5.2.1).

Parameter	GC-MS_A	GC-MS_B	GC-MS_C
Injection volume	1 μ L	1 μ L	1 μ L
Injection mode	Split	Split	Split
Injection temperature	250°C	250°C	250°C
Carrier gas	He	He	He
Flow control mode	Linear velocity	Linear velocity	Linear velocity
Total flow	11.5 mL min ⁻¹	15.0 mL min ⁻¹	13.1 mL min ⁻¹
Column flow	1.13 mL min ⁻¹	1.19 mL min ⁻¹	1.00 mL min ⁻¹
Linear velocity	39.5 cm sec ⁻¹	39.5 cm sec ⁻¹	36.7 cm sec ⁻¹
Purge flow	3.0 mL min ⁻¹	3.0 mL min ⁻¹	3.0 mL min ⁻¹
Split ratio	6.5	9.1	9.1
Oven program	100°C for 4 min, 20°C min ⁻¹ to 300°C, 300°C for 5 min (Total: 19.0 min)	50°C for 5 min, 40°C min ⁻¹ to 300°C, 300°C for 5 min (Total: 16.25 min)	70°C, 10°C min ⁻¹ to 170°C, 30°C min ⁻¹ to 280°C, 280°C for 3 min (Total: 16.6 min)
Column	Zebtron ZB-5MSi (30 m \times 0.25 mm \times 0.25 μ m)	Zebtron ZB-5MSi (30 m \times 0.25 mm \times 0.25 μ m)	Zebtron ZB-5MSi (30 m \times 0.25 mm \times 0.25 μ m)

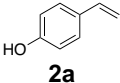
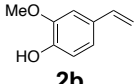
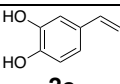
Continuation of Table 65.

Ion source temperature	250°C	250°C	250°C
Interface temperature	320°C	320°C	320°C
Solvent cut time	4.01 min	4.01 min	4.01 min
ACQ mode	Scan	Scan	Scan
Event time	0.13 sec	0.13 sec	0.13 sec
Scan speed	5000	5000	5000
Start <i>m/z</i>	50.00	50.00	50.00
End <i>m/z</i>	500.00	500.00	500.00

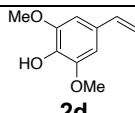
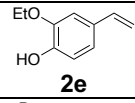
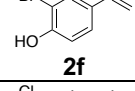
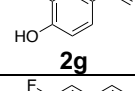
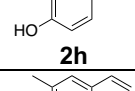
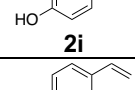
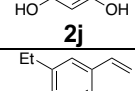
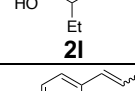
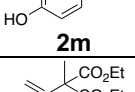
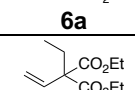
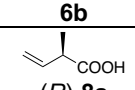
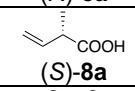
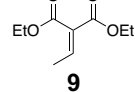
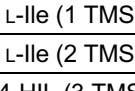
Table 66. Detailed parameters of methods used for chiral GC-FID analyses (see 7.5.2.2).

Parameter	GC-FID_D	GC-FID_E
Injection volume	1 μL	1 μL
Injection mode	Split	Split
Injection temperature	250°C	250°C
Carrier gas	N ₂	N ₂
Flow control mode	Velocity	Velocity
Total flow	35.4 mL min ⁻¹	43.6 mL min ⁻¹
Column flow	1.54 mL min ⁻¹	1.93 mL min ⁻¹
Linear velocity	44.8 cm sec ⁻¹	44.8 cm sec ⁻¹
Purge flow	3.0 mL min ⁻¹	3.0 mL min ⁻¹
Split ratio	20	20
Oven program	100°C (Total: 60 min)	80°C (Total: 60 min)
Column	HYDRODEX β -6TBDM (25 m \times 0.25 mm, proprietary df)	HYDRODEX β -6TBDM (25 m \times 0.25 mm, proprietary df)
Detection temperature	230°C	230°C
Makeup gas	N ₂	N ₂
Makeup flow	24.0 mL min ⁻¹	24.0 mL min ⁻¹
H ₂ flow	32.0 mL min ⁻¹	32.0 mL min ⁻¹
Air flow	200 mL min ⁻¹	200 mL min ⁻¹

Table 67. Analytical data for compounds analyzed by GC-MS and chiral GC-FID.

Compound	GC-MS			Chiral GC-FID	
	Method	Retention time (min)	Exact mass	Method	Retention time (min)
 2a	GC-MS_A	6.3	120.06	-	-
 2b	GC-MS_A	7.5	150.07	-	-
 2c	GC-MS_A	8.4	136.05	-	-

Continuation of Table 67.

 2d	GC-MS_A	9.3	180.08	-	-
 2e	GC-MS_A	7.9	164.08	-	-
 2f	GC-MS_A	7.3	197.97	-	-
 2g	GC-MS_A	6.4	154.02	-	-
 2h	GC-MS_A	5.0	138.05	-	-
 2i	GC-MS_A	7.1	134.07	-	-
 2j	GC-MS_A	8.5	136.05	-	-
 2l	GC-MS_A	8.8	176.12	-	-
 2m	GC-MS_A	7.3	134.07	-	-
 6a	GC-MS_B	8.5	200.10	-	-
 6b	GC-MS_B	8.8	214.12	-	-
 (R)-8a	-	-	-	GC-FID_D	11.8
				GC-FID_E	31.5
 (S)-8a	-	-	-	GC-FID_D	11.3
				GC-FID_E	29.8
 9	GC-MS_B	8.7	168.09	-	-
L-Ile (1 TMS)	GC-MS_C	6.4	203.13	-	-
L-Ile (2 TMS)	GC-MS_C	8.1	275.17	-	-
4-HIL (3 TMS)	GC-MS_C	10.7	363.21	-	-
α -Ketoglutarate (3TMS)	GC-MS_C	11.5	290.10	-	-
Succinate (2TMS)	GC-MS_C	8.3	262.11	-	-

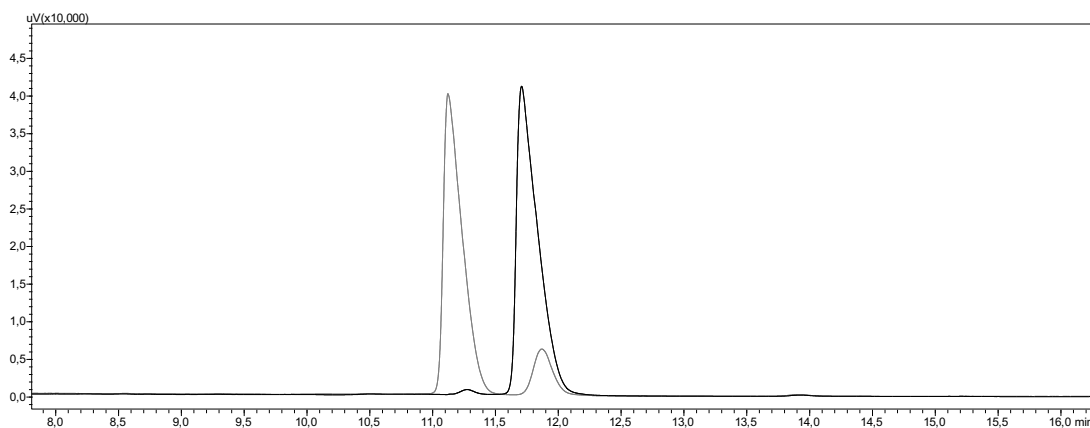


Figure S 41. Chromatograms obtained from chiral GC-FID analysis of (*S*)-**8a** ($t_R = 11.3$ min) and (*R*)-**8a** ($t_R = 11.8$ min). The shown traces correspond to the products obtained from using AMDase IPLL (98 %ee (*R*), black) and CLGIPL (72 %ee (*S*), gray). Method: GC-FID_D.

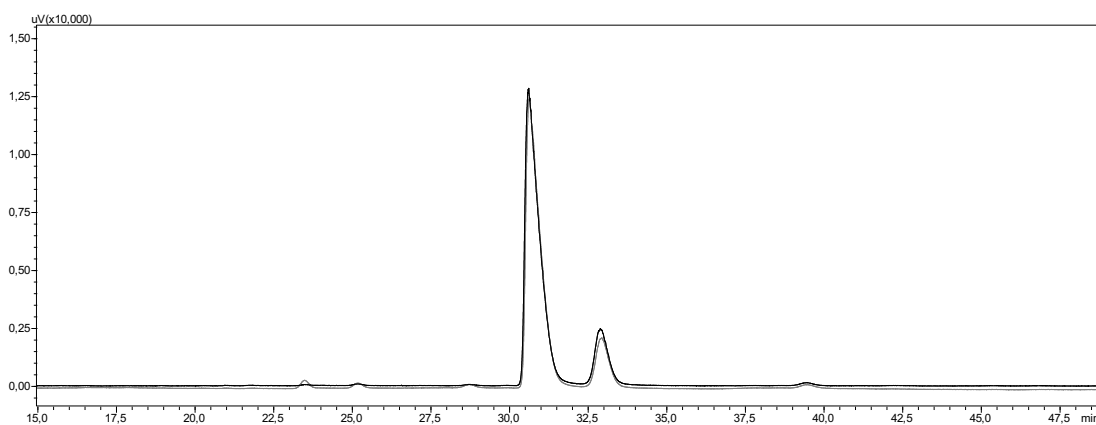
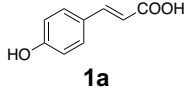
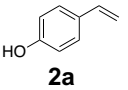
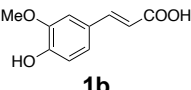
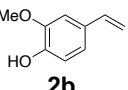
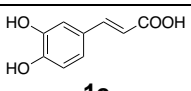
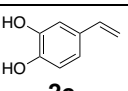
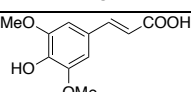
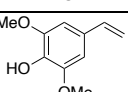
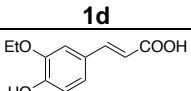
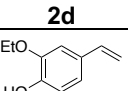
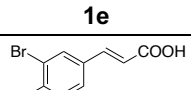
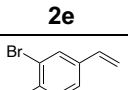
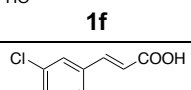
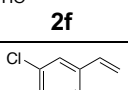
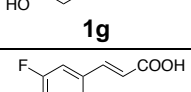
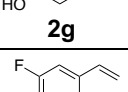
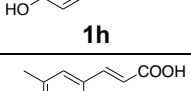
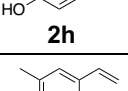
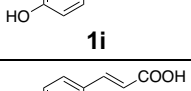
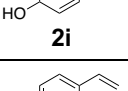
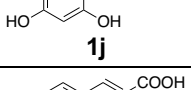
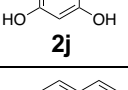
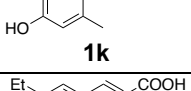
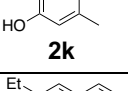
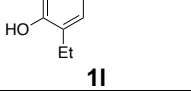
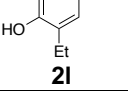
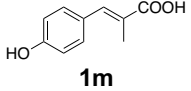
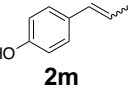


Figure S 42. Chromatograms obtained from chiral GC-FID analysis of (*S*)-**8a** ($t_R = 30.6$ min) and (*R*)-**8a** ($t_R = 32.9$ min). The shown traces correspond to the products obtained from reactions with purified AMDase CLGIPL (black) and AMDase CLGIPL cell-free extracts (gray). Method: GC-FID_E.

Appendix J HPLC supplementary data

Table 68. Overview on substrates and their respective retention times (t_R) and detection wavelengths used for quantification via HPLC and the corresponding products.

Substrate	t_R (Substrate) (min)	Detection wavelength (nm)	Product	t_R (Product) (min)
 1a	3.0	280	 2a	5.3
 1b	3.0	280	 2b	5.5
 1c	2.6	280	 2c	3.8
 1d	2.9	320	 2d	5.0
 1e	3.4	320	 2e	7.3
 1f	3.5	320	 2f	7.8
 1g	3.4	320	 2g	7.2
 1h	3.1	320	 2h	5.9
 1i	3.3	320	 2i	7.0
 1j	2.7	320	 2j	4.1
 1k	3.1	320	 2k	n.d.
 1l	5.5	320	 2l	n.d.
 1m	3.2	280	 2m	6.6
 1n	3.4	280	 2n	n.d.

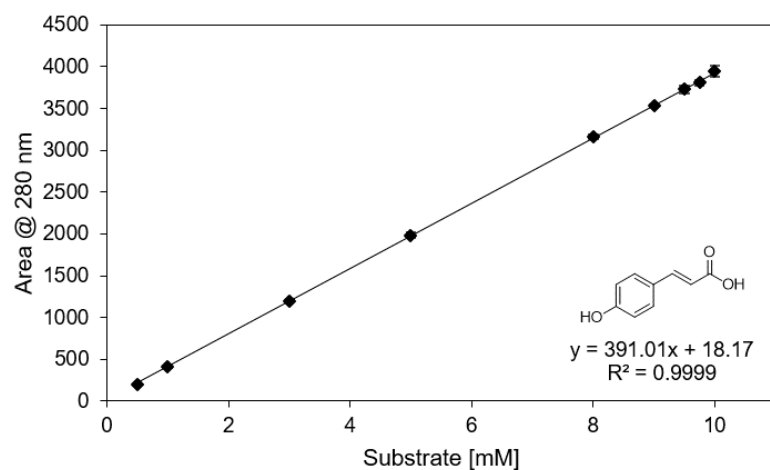


Figure S 43. Calibration curve for *p*-coumaric acid (**1a**) (2 μ L injection volume, $\lambda_{\text{detection}} = 280$ nm).

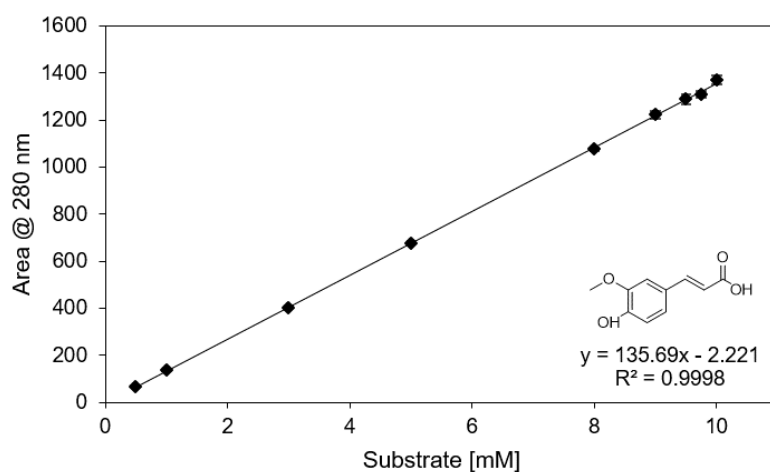


Figure S 44. Calibration curve for ferulic acid (**1b**) (2 μ L injection volume, $\lambda_{\text{detection}} = 280$ nm).

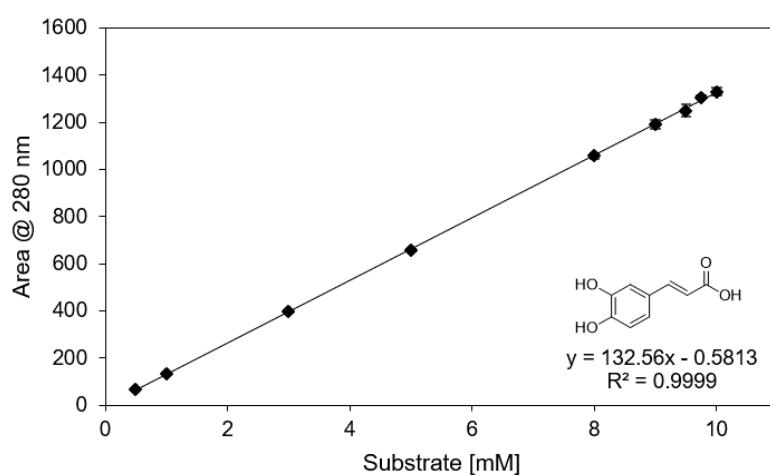


Figure S 45. Calibration curve for caffeic acid (**1c**) (2 μ L injection volume, $\lambda_{\text{detection}} = 280$ nm).

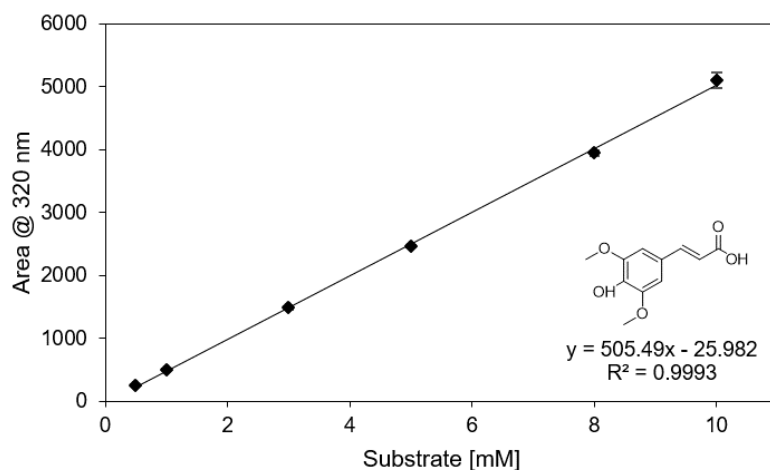


Figure S 46. Calibration curve for sinapic acid (**1d**) (2 μ L injection volume, $\lambda_{\text{detection}} = 320$ nm).

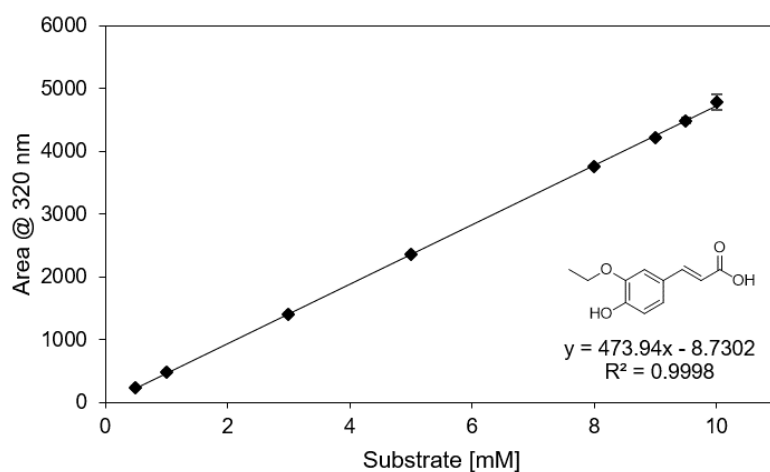


Figure S 47. Calibration curve for (*E*)-3-(3-ethoxy-4-hydroxyphenyl)acrylic acid (**1e**) (2 μ L injection volume, $\lambda_{\text{detection}} = 320$ nm).

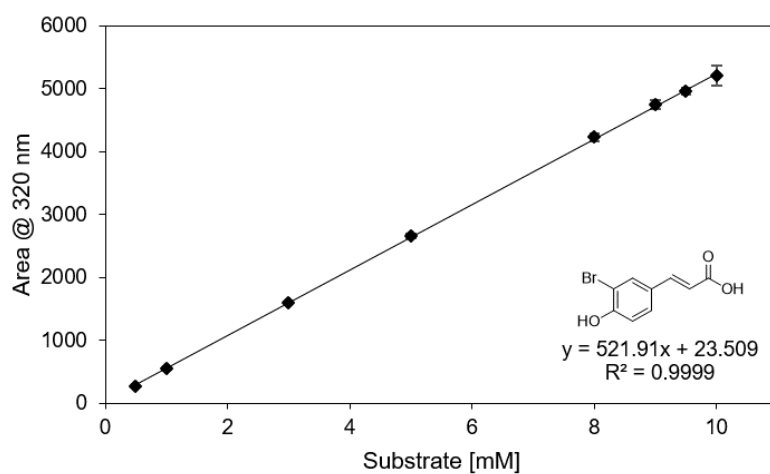


Figure S 48. Calibration curve for (*E*)-3-(3-bromo-4-hydroxyphenyl)acrylic acid (**1f**) (2 μ L injection volume, $\lambda_{\text{detection}} = 320$ nm).

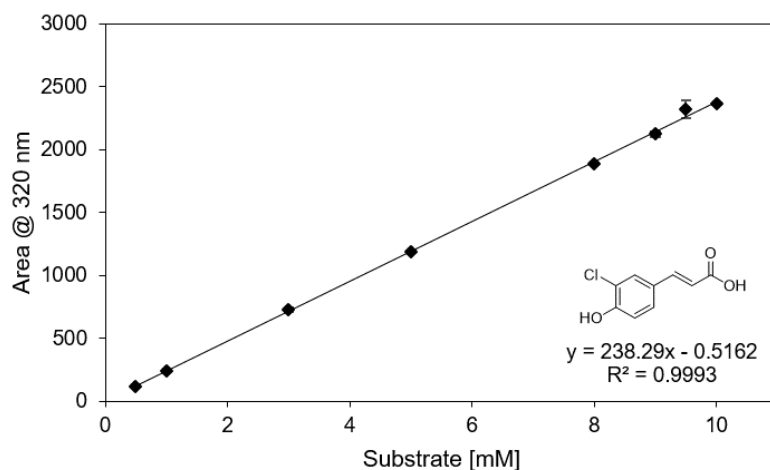


Figure S 49. Calibration curve for (*E*)-3-(3-chloro-4-hydroxyphenyl)acrylic acid (**1g**) (1 μ L injection volume, $\lambda_{\text{detection}} = 320$ nm).

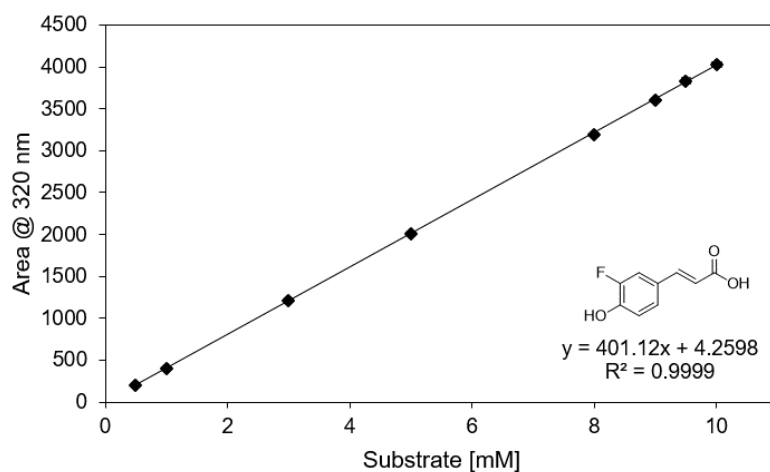


Figure S 50. Calibration curve for (*E*)-3-(3-fluoro-4-hydroxyphenyl)acrylic acid (**1h**) (2 μ L injection volume, $\lambda_{\text{detection}} = 320$ nm).

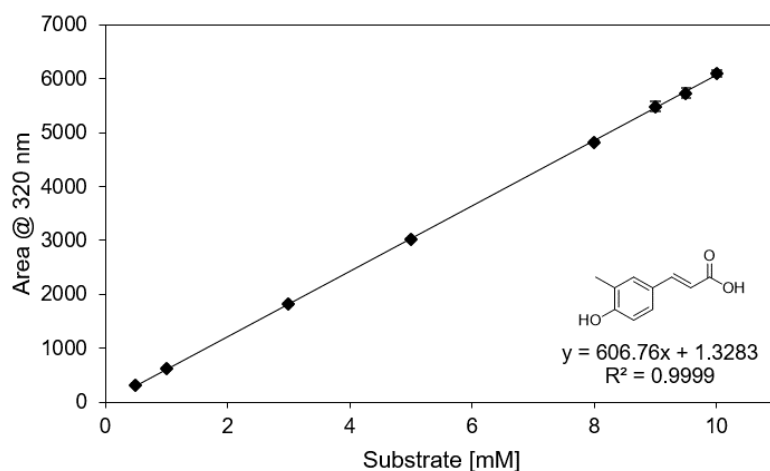


Figure S 51. Calibration curve for (*E*)-3-(4-hydroxy-3-methylphenyl)acrylic acid (**1i**) (2 μ L injection volume, $\lambda_{\text{detection}} = 320$ nm).

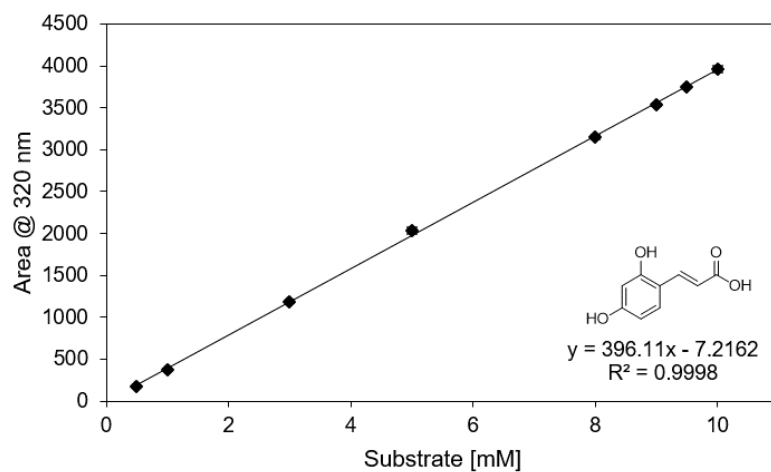


Figure S 52. Calibration curve for (*E*)-3-(2,4-dihydroxyphenyl)acrylic acid (**1j**) (2 μ L injection volume, $\lambda_{\text{detection}} = 320$ nm).

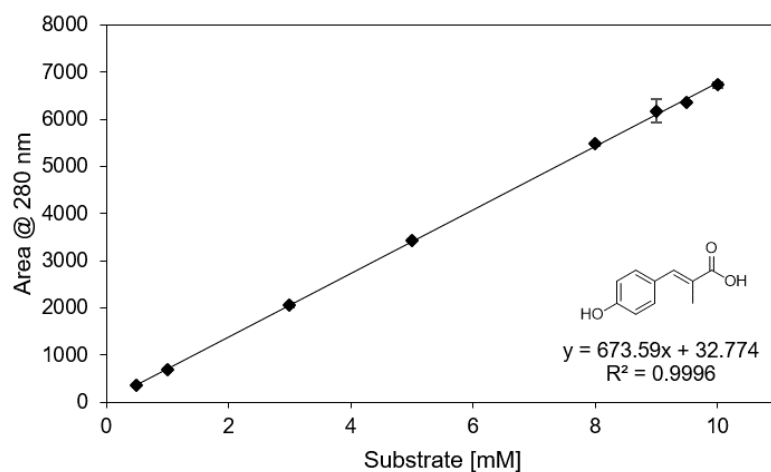
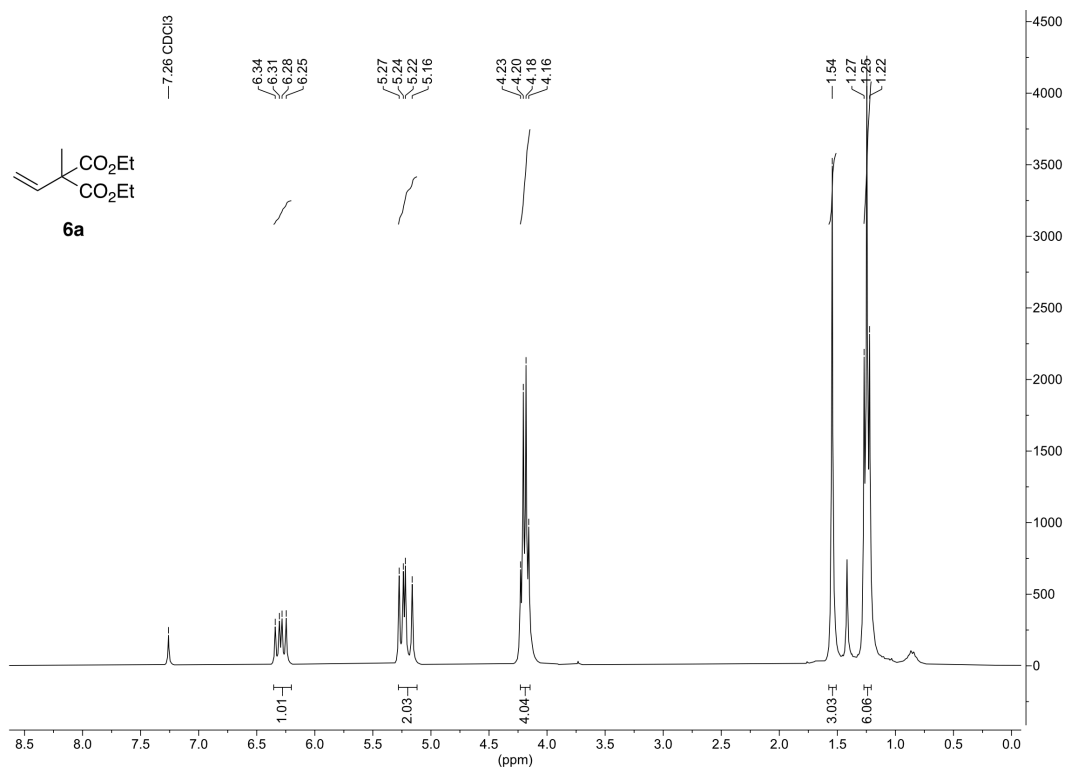
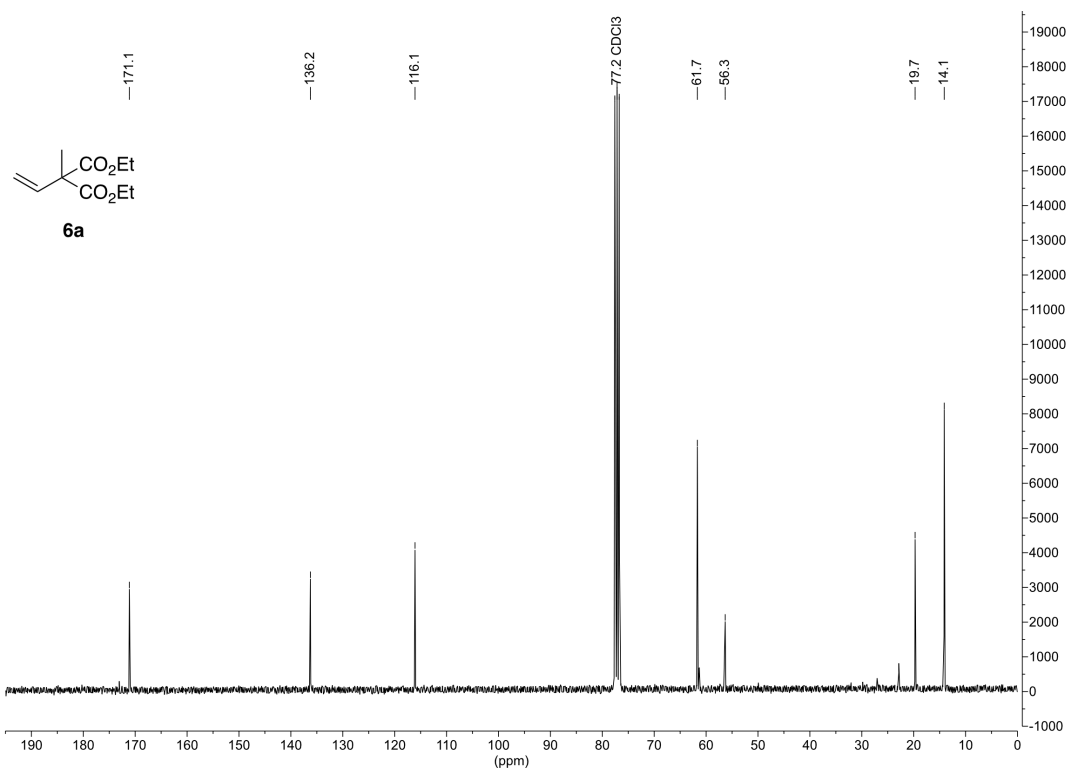


Figure S 53. Calibration curve for (*E*)-3-(4-hydroxyphenyl)-2-methylacrylic acid (**1m**) (2 μ L injection volume, $\lambda_{\text{detection}} = 280$ nm).

Appendix K NMR spectra

Figure S 54. ¹H-NMR of diethyl 2-methyl-2-vinylmalonate (**6a**) in CDCl₃.Figure S 55. ¹³C-NMR of diethyl 2-methyl-2-vinylmalonate (**6a**) in CDCl₃.

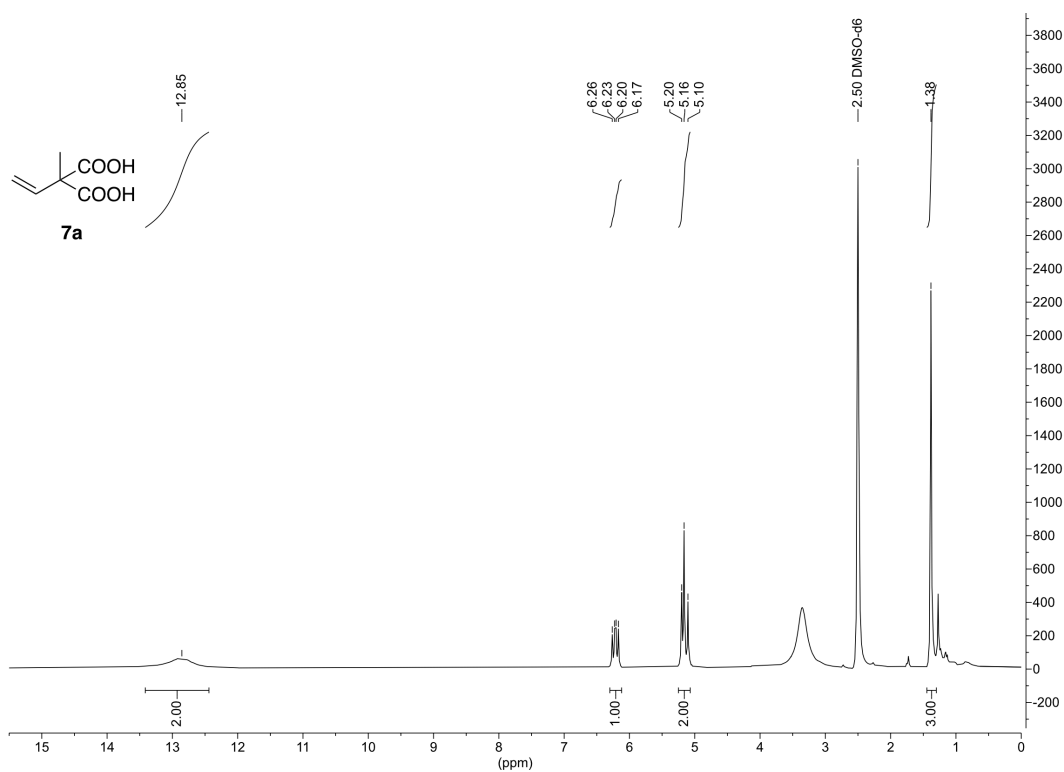


Figure S 56. $^1\text{H-NMR}$ of 2-methyl-2-vinylmalonic acid (**7a**) in DMSO-d^6 .

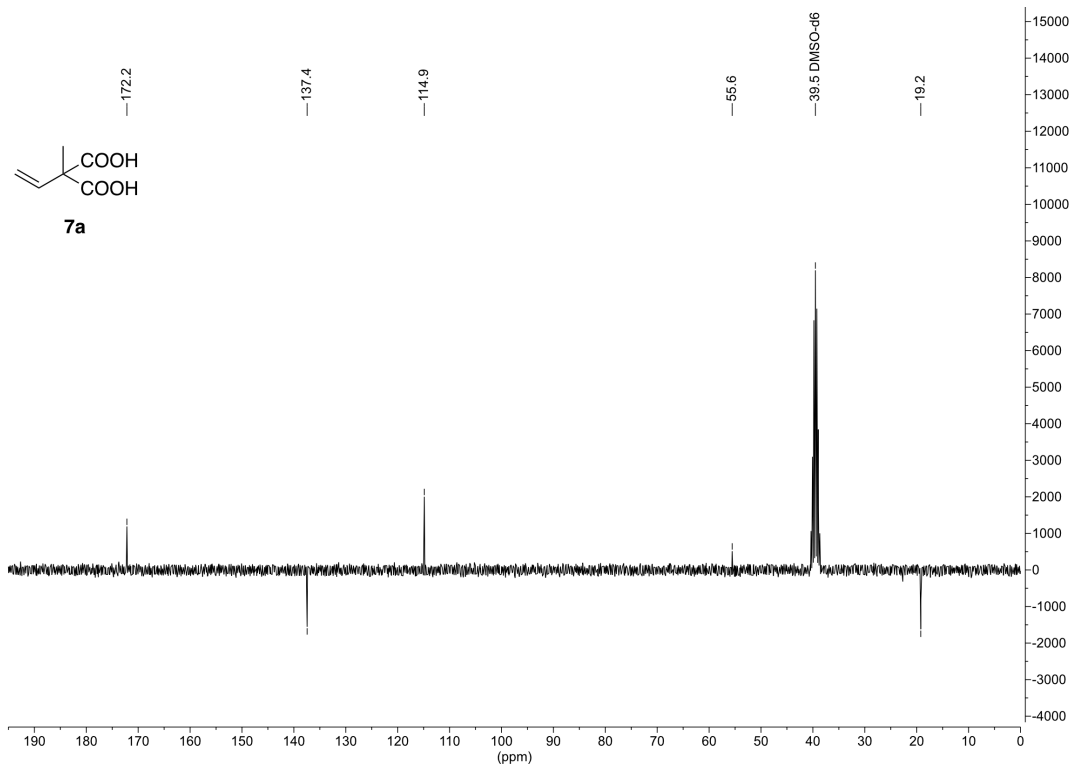


Figure S 57. $^{13}\text{C-NMR}$ (APT) of 2-methyl-2-vinylmalonic acid (**7a**) in DMSO-d^6 .

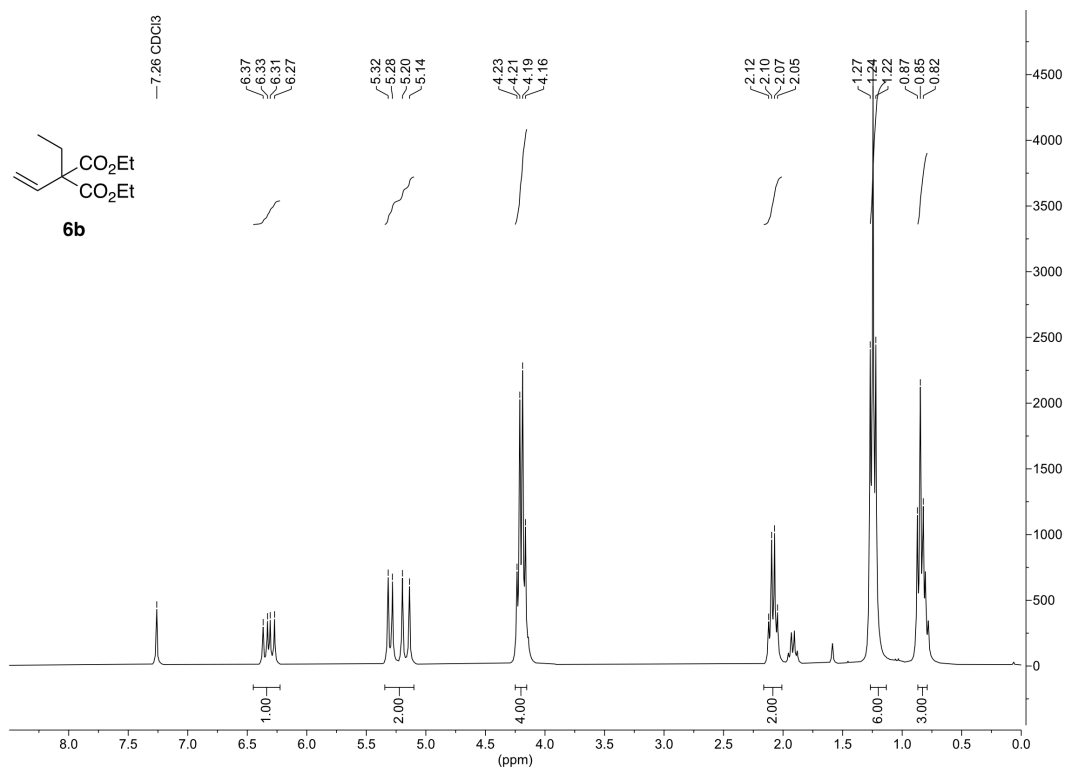


Figure S 58. ¹H-NMR of diethyl 2-ethyl-2-vinylmalonate (**6b**) in CDCl₃.

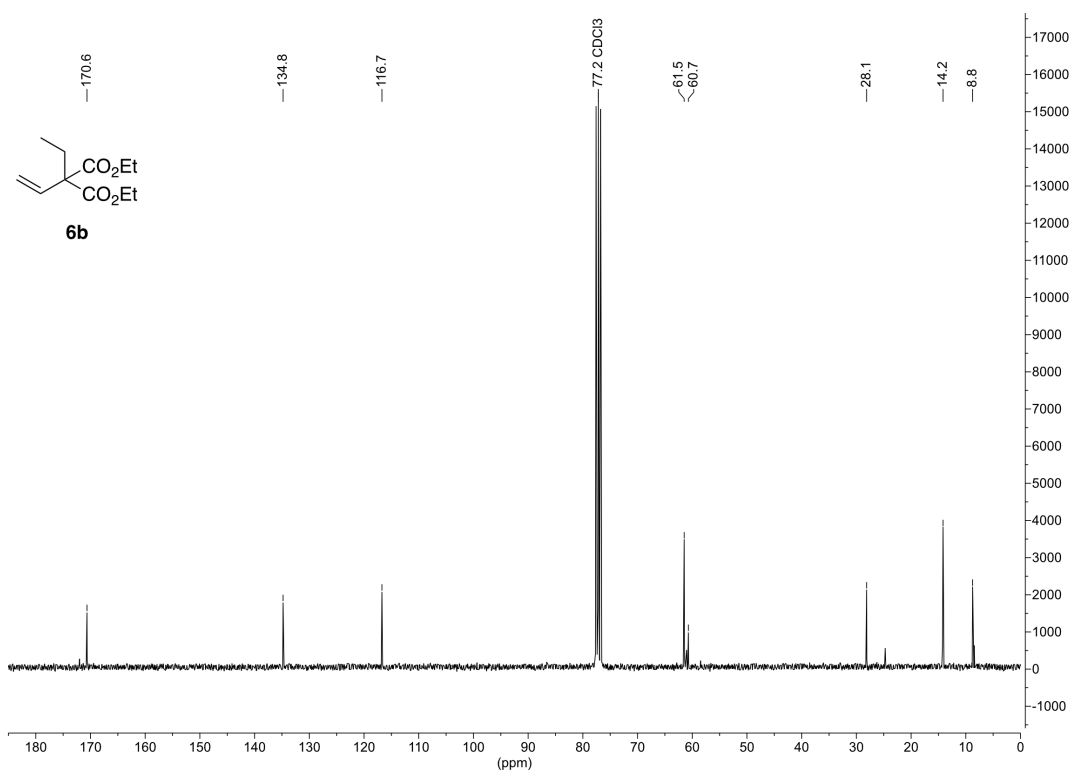


Figure S 59. ¹³C-NMR of diethyl 2-ethyl-2-vinylmalonate (**6b**) in CDCl₃.

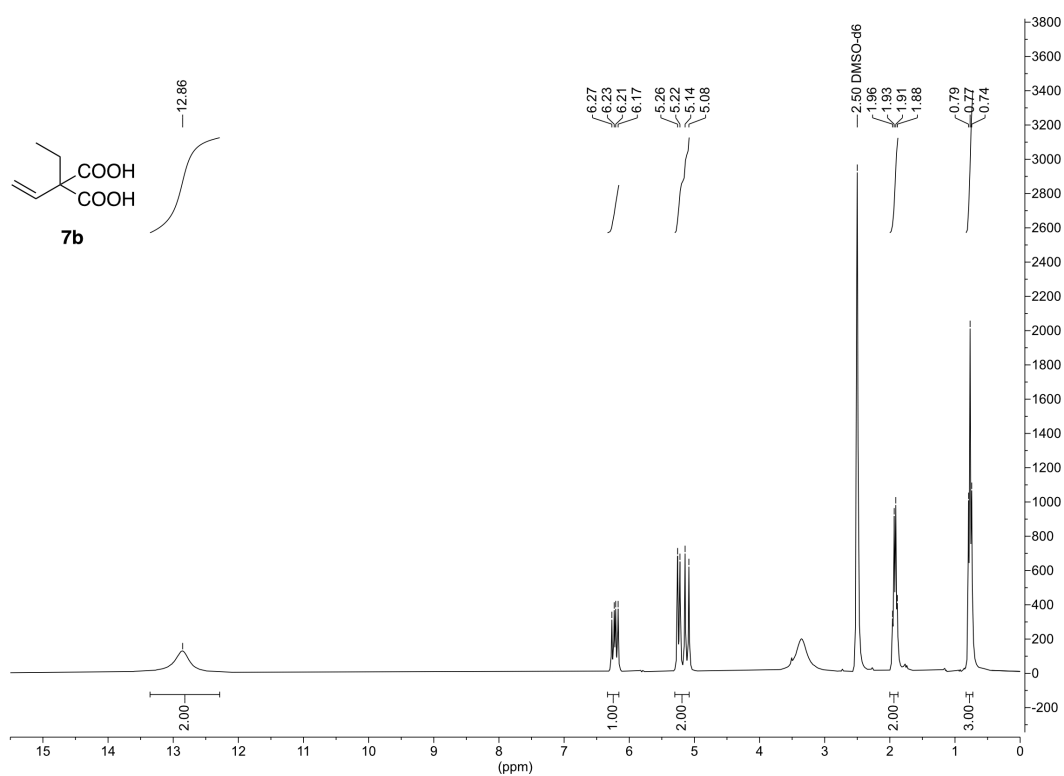


Figure S 60. $^1\text{H-NMR}$ of 2-ethyl-2-vinylmalonic acid (**7b**) in DMSO-d^6 .

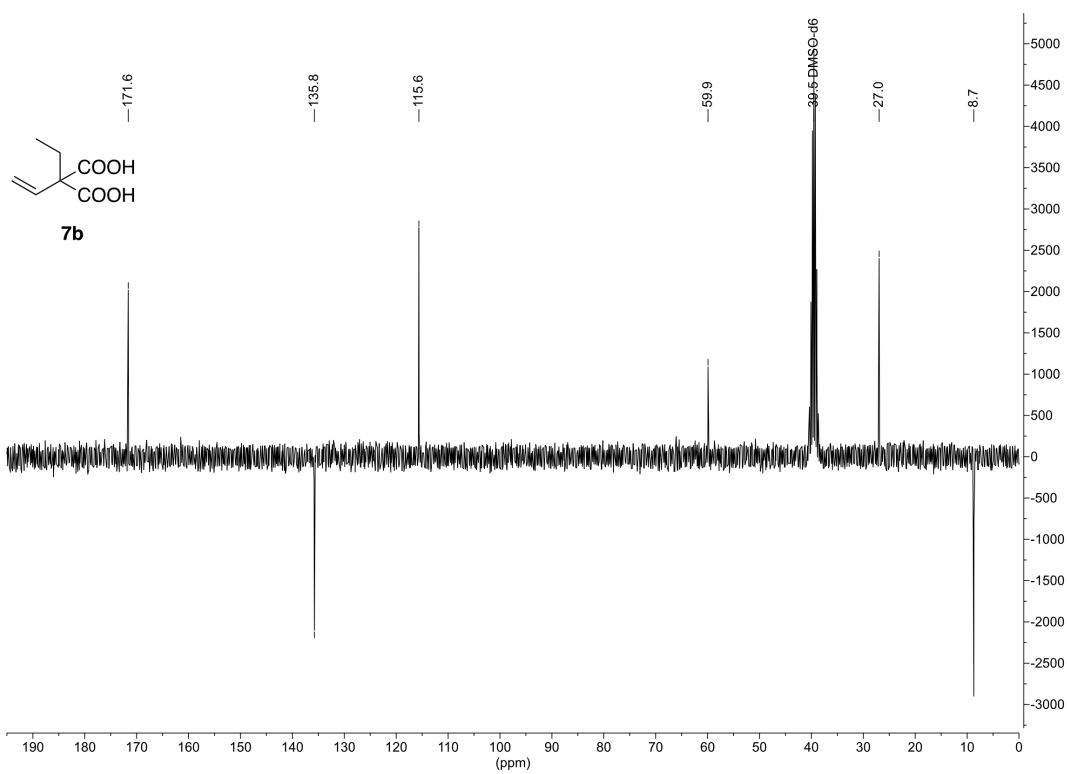


Figure S 61. $^{13}\text{C-NMR}$ (APT) of 2-ethyl-2-vinylmalonic acid (**7b**) in DMSO-d^6 .

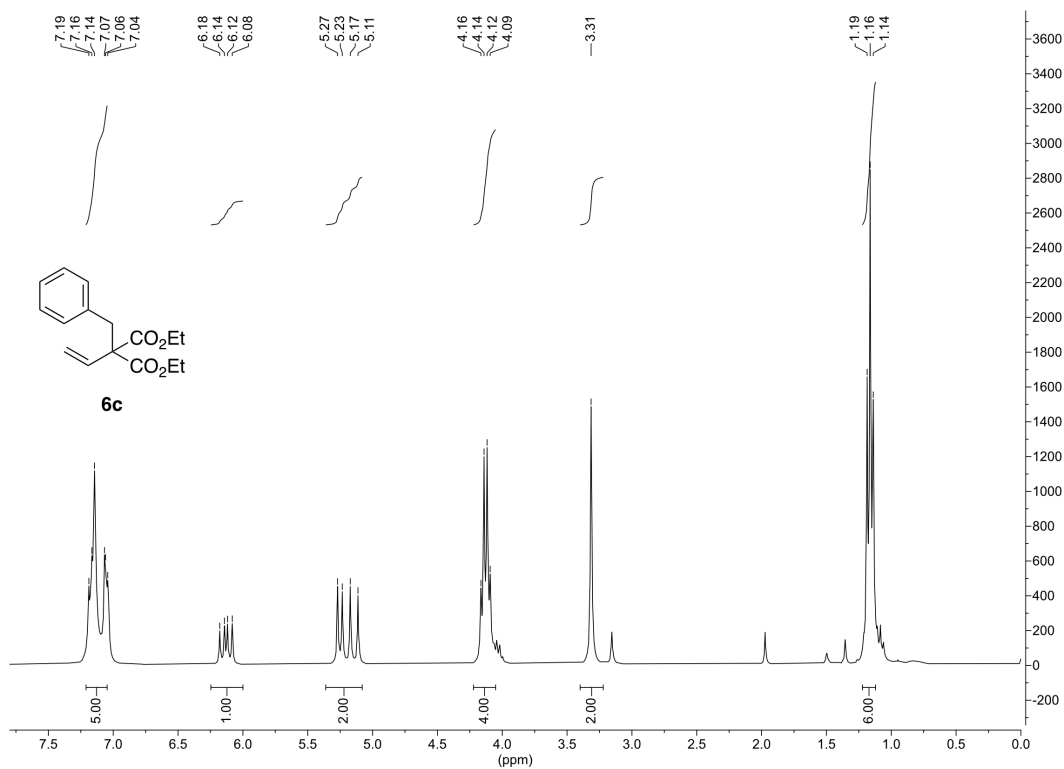


Figure S 62. ^1H -NMR of diethyl 2-benzyl-2-vinylmalonate (**6c**) in CDCl_3 .

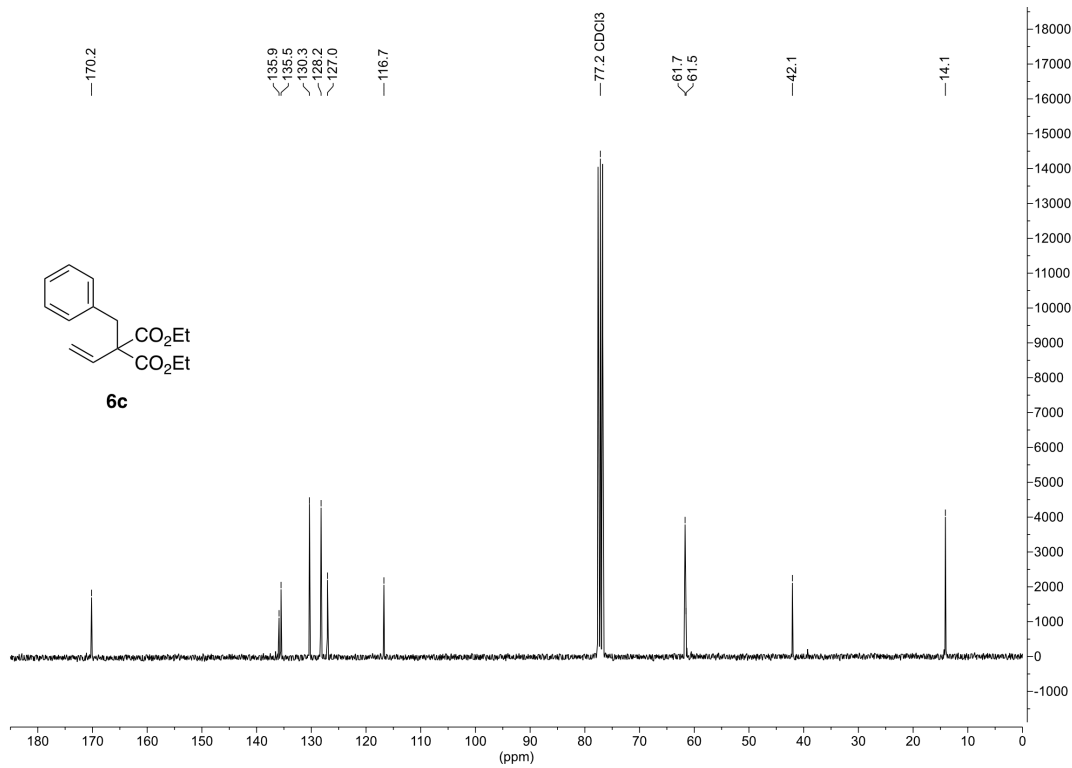


Figure S 63. ^{13}C -NMR of diethyl 2-benzyl-2-vinylmalonate (**6c**) in CDCl_3 .

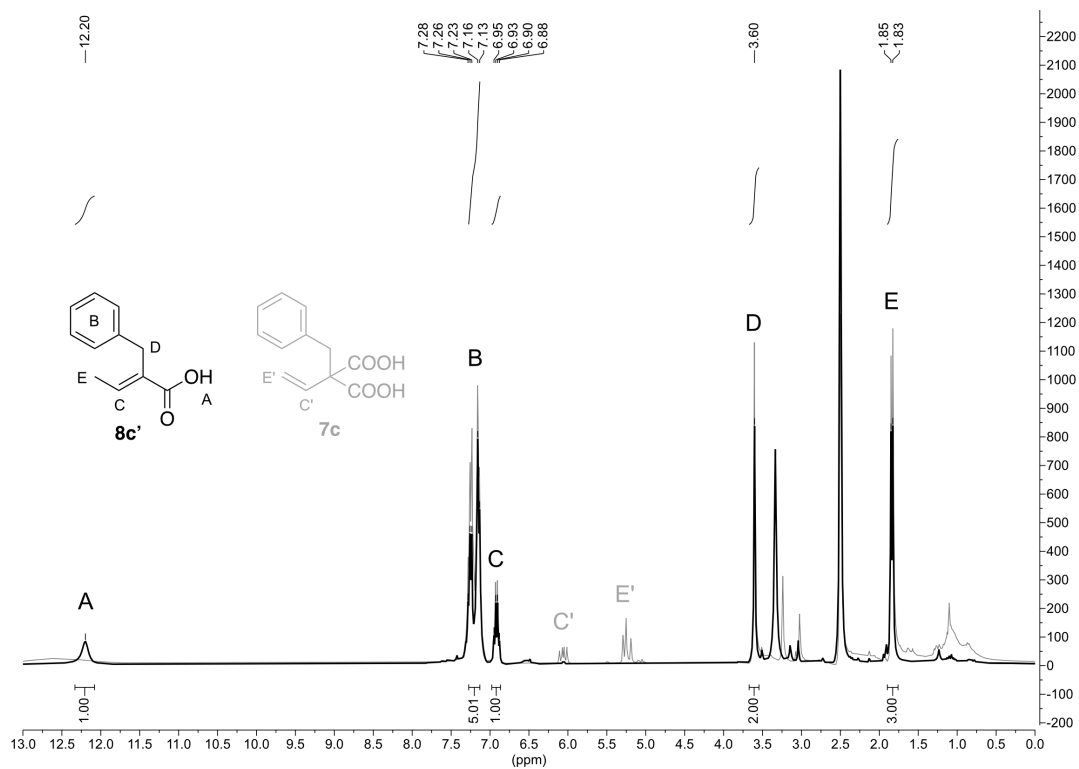


Figure S 64. Stacked $^1\text{H-NMR}$ of 2-benzylbut-2-enoic acid (**8c'**) (black) and a mixture of **8c'** and 2-benzyl-2-vinylmalonic acid (**7c**) (gray) in DMSO-d_6 .