

**Novel aminoacylases for the synthesis  
of acyl-amino acids as biosurfactants**

Inaugural dissertation

for the attainment of the title of doctor  
in the Faculty of Mathematics and Natural Sciences  
at the Heinrich Heine University Düsseldorf

presented by  
**Gerrit Haeger**  
from Dormagen

Dormagen, July 2023



The work of this thesis has been conducted at the Institute of Nano- and Biotechnologies, FH Aachen - University of Applied Sciences in cooperation with the Institute of Molecular Enzyme Technology (IMET), Institute of Bio- and Geosciences 1, Forschungszentrum Jülich, from October 2019 until July 2023 under the supervision of Prof. Dr. Petra Siegert, Prof. Dr. Johannes Bongaerts and Prof. Dr. Karl-Erich Jaeger.

Published by permission of the  
Faculty of Mathematics and Natural Sciences at  
Heinrich Heine University Düsseldorf

Examiner:

1. Prof. Dr. Karl-Erich Jaeger  
Institute of Molecular Enzyme Technology (IMET)  
Forschungszentrum Jülich GmbH  
Heinrich Heine University Düsseldorf

2. Prof. Dr. Petra Siegert  
Institute of Nano- and Biotechnologies  
FH Aachen – University of Applied Sciences

Date of the oral examination: 08.12.2023

## **Eidesstattliche Erklärung**

Ich, Herr Gerrit Haeger, versichere an Eides statt, dass die vorliegende Dissertation „Aminoacylases for the synthesis of acyl-amino acids as biosurfactants“ von mir selbstständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist. Diese Dissertation wurde weder in dieser noch in ähnlicher Form bei einer anderen Fakultät eingereicht. Bisher habe ich keine erfolglosen oder erfolgreichen Promotionsversuche unternommen.

---

Gerrit Haeger

## I. List of publications

**Haeger G**, Bongaerts J, Siegert P. A convenient ninhydrin assay in 96-well format for amino acid-releasing enzymes using an air-stable reagent. *Analytical Biochemistry* 2022; 654:114819. doi:10.1016/j.ab.2022.114819.

**Haeger G**, Wirges J, Tanzmann N, Oyen S, Jolmes T, Jaeger KE, Schörken U, Bongaerts J, Siegert P. Chaperone assisted recombinant expression of a mycobacterial aminoacylase in *Vibrio natriegens* and *Escherichia coli* capable of N-lauroyl-L-amino acid synthesis. *Microbial Cell Factories* 22:77 2023. doi:10.1186/s12934-023-02079-1.

**Haeger G**, Jolmes T, Oyen S, Jaeger KE, Bongaerts J, Schörken U, Siegert P. Novel recombinant aminoacylase from *Paraburkholderia monticola* capable of N-acyl-amino acid synthesis. *Applied Microbiology and Biotechnology* 108, (93) 2024. doi:10.1007/s00253-023-12868-8

**Haeger G**, Wirges J, Genesseeux L, Guiavarc'h Y, Humeau C, Paris C, Chevalot I, Jaeger KE, Bongaerts J, Siegert P. Biocatalytic potential of MsAA aminoacylase for synthesis of N-acyl-L-amino acids in aqueous media. Manuscript in preparation.

**Haeger G**, Probst J, Jaeger KE, Bongaerts J, Siegert P. Novel aminoacylases from *Streptomyces griseus* DSM 40236 and their recombinant production in *Streptomyces lividans*. *FEBS Open Bio* 13 (12) 2023. doi:10.1002/2211-5463.13723

Patent application filed in context of this thesis:

“Aminoacylase and methods of use thereof”; Applicant: BASF SE; European Patent Office;  
Application number 23176331.9; Priority date: 31.05.2023.

Publication not included in this work:

**Haeger G**, Grankin A, Wagner M. Construction of an *Aspergillus oryzae* triple amylase deletion mutant as a chassis to evaluate industrially relevant amylases using multiplex CRISPR/Cas9 editing technology. *Applied Research* 2023; e202200106. doi: 10.1002/appl.202200106.

## **II. Conference contributions**

**Haeger G**, Tanzmann N, Bongaerts J, Siegert P. Enhancement of soluble heterologous protein expression in *V. natriegens* by co-expression of chaperones. *Annual conference of the association for general and applied microbiology (VAAM)*, Düsseldorf, Germany, February 20-23, 2022.

**Haeger G**, Wirges J, Tanzmann N, Oyen S, Bongaerts J, Siegert P. Novel mycobacterial aminoacylase for synthesis of N-acyl-amino acids. *10th International Congress of Biocatalysis (Biocat2022)*, Hamburg, Germany, August 28-September 01, 2022.

### III. Table of contents

I. List of publications .....	III
II. Conference contributions.....	V
III. Table of contents .....	VI
IV. Zusammenfassung.....	VIII
V. Abstract .....	IX
1. Introduction .....	1
1.1. Acyl-amino acids as biosurfactants .....	1
1.2. Biocatalysis.....	5
1.2.1. Aminoacylases .....	8
1.2.2. Enzymes from the MH clan and M20 family of metallopeptidases .....	9
1.2.3. Enzymes from the MJ clan and M38 family of metallopeptidases.....	14
1.2.4. Biocatalytic synthesis of acyl-amino acids with aminoacylases.....	17
1.3. Heterologous expression of aminoacylases .....	26
1.3.1. Recombinant expression with <i>Escherichia coli</i> .....	26
1.3.2. Recombinant expression with <i>Vibrio natriegens</i> .....	29
1.3.3. Recombinant expression with <i>Streptomyces lividans</i> .....	30
1.4. Objective of the thesis .....	32
2. Results .....	33
2.1. Chapter I .....	34
2.2. Chapter II.....	40
2.3. Chapter III.....	55
2.4. Chapter IV .....	71
2.5. Chapter V.....	92
3. Discussion .....	108
3.1. Overview of results.....	108
3.2. Novel aminoacylases were identified .....	110
3.3. Conserved residues of metallopeptidases in the novel aminoacylases.....	112
3.4. Predicted structures of the novel aminoacylases .....	115
3.5. The aminoacylases could be overexpressed in various hosts .....	122
3.6. Development of an optimized ninhydrin-based activity assay .....	126
3.7. The purified aminoacylases were characterized and compared .....	129
3.7.1. The aminoacylases have different substrate specificities .....	129
3.7.2. Comparative biochemical characterization.....	132
3.8. Biocatalytic synthesis .....	135

3.9. Comparison to chemical synthesis .....	138
3.10. Outlook .....	139
3.10.1. Aminoacylase PmAcy .....	139
3.10.2. Aminoacylase MsAA .....	140
3.10.3. Streptomycetal aminoacylases SgAA and SgELA .....	141
3.11. Conclusion .....	142
4. References .....	144
5. Abbreviations .....	163
6. Acknowledgements .....	165
7. Supplementary material.....	166

#### IV. Zusammenfassung

Der Markt für Detergenzien und Tenside ist der Größte innerhalb der Chemieindustrie. Insbesondere als Konsumgüter steigt die Nachfrage an biobasierten Tensiden aus nachhaltiger Produktion. Die Zielmoleküle dieser Arbeit, namentlich Acyl-Aminosäuren, stellen eine wichtige Substanzklasse innerhalb der grünen Tenside dar. Diese finden als milde Tenside bereits Einsatz in hochwertigen Kosmetika. Die kommerzielle Synthese der Moleküle wird konventionell mit der Schotten-Baumann Methode durchgeführt, bei der Acylchloride eingesetzt werden. Obwohl es sich bei den Produkten um biobasierte Tenside handelt, ist die chemische Synthese nicht nachhaltig und umweltschädlich. Neue Aminoacylasen sollen eine nachhaltige Synthese ermöglichen. Durch die Fähigkeit, freie Fettsäuren ohne vorherige Aktivierung als Substrate zu nutzen, bieten diese Enzyme großes Potential als Biokatalysatoren. Der Einsatz von Aminoacylasen für die Acylierung von Aminosäuren ist jedoch industriell noch nicht genutzt und wissenschaftlich wenig erforscht.

Es wurden neue Aminoacylasen über Datenbanksuchen identifiziert und für die rekombinante Expression kloniert. Nach der Produktion und Aufreinigung in geeigneten Wirten wurden die Aminoacylasen charakterisiert und für den Einsatz in der Synthese evaluiert. So wurden Aminoacylasen aus *Paraburkholderia monticola* DSM 1000849 (PmAcy), *Mycolicibacterium smegmatis* MKD 8 (MsAA), und *Streptomyces griseus* DSM 40236<sup>T</sup> (SgAA und SgELA) untersucht. Die Enzyme konnten in *E. coli* BL21(DE3), *V. natriegens* Vmax<sup>TM</sup> und *S. lividans* TK23 exprimiert werden, wobei auch durch Co-Expression von molekularen Chaperonen die Produktion von löslichem Enzym ermöglicht oder verbessert wurde. Die Aminoacylase PmAcy stellte sich als herausragender Biokatalysator heraus. Diese weist nicht nur eine sehr gute Stabilität gegenüber hohen Temperaturen und breitem pH-Spektrum auf, sondern katalysiert auch die Acylierung von verschiedenen Aminosäuren zu hohen Umsätzen von teilweise über 70 %. Die  $\alpha$ -Aminoacylase MsAA konnte insbesondere für die Synthese von Lauroyl-Methionin verwendet werden, sodass ein Umsatz von 67 % erreicht werden konnte. Auch weitere Acyldonoren und -akzeptoren konnten für die Synthese verwendet werden. Die homologe  $\alpha$ -Aminoacylase SgAA konnte ebenfalls Lauroyl-Methionin synthetisieren. Für SgELA konnte die putative Funktion als  $\epsilon$ -Lysin-Acylyase nachgewiesen werden. Durch geringe Enzymausbeute in *S. lividans* konnte allerdings nur eine initiale Charakterisierung durchgeführt werden. Zusammenfassend konnten biokatalytische Synthesen etabliert werden. Insbesondere für die  $\alpha$ -Acylierung übertreffen die Prozesse der rekombinanten Enzymproduktion und der Acylierungen bisher beschriebene Produktivität und Ausbeuten für Aminoacylasen.

## V. Abstract

The market for detergents and surfactants is the largest within the chemical industry. Especially as consumer products, the demand for bio-based surfactants from sustainable production is increasing. The target molecules of this work, namely acyl-amino acids, represent an important substance class within green surfactants. They are already used as mild surfactants in high-quality cosmetics. The commercial synthesis of the molecules is conventionally carried out by the Schotten-Baumann method using acyl chlorides. Although the products are biobased surfactants, the chemical synthesis is unsustainable and environmentally harmful due to the chlorine chemistry. New aminoacylases are anticipated to enable sustainable synthesis. With the ability to use free fatty acids as substrates without prior activation, these enzymes offer great potential as biocatalysts. However, the use of aminoacylases for the acylation of amino acids has not yet been exploited industrially and has been scientifically underexplored.

New aminoacylases were identified via database searches and cloned for recombinant expression. After production and purification in suitable hosts, the aminoacylases were characterized and evaluated for use in synthesis. Thus, aminoacylases from *Paraburkholderia monticola* DSM 1000849 (PmAcy), *Mycolicibacterium smegmatis* MKD 8 (MsAA), and *Streptomyces griseus* DSM 40236<sup>T</sup> (SgAA and SgELA) were studied. The enzymes could be expressed with *E. coli* BL21(DE3), *V. natriegens* Vmax<sup>TM</sup>, and *S. lividans* TK23, and soluble enzyme production was furthermore enabled or enhanced by co-expression of molecular chaperones. The aminoacylase PmAcy turned out to be an outstanding biocatalyst. The enzyme not only exhibits very good stability towards high temperatures and a broad pH spectrum, but also catalyzes the acylation of various amino acids to high conversions, sometimes exceeding 70 %. In particular, the  $\alpha$ -aminoacylase MsAA could be used for the synthesis of lauroyl methionine, resulting in a conversion of 67 %. Other acyl donors and acceptors could also be used the synthesis. The homologous  $\alpha$ -aminoacylase SgAA was also able to synthesize lauroyl-methionine. For SgELA, putative function as an  $\epsilon$ -lysine acylase was demonstrated. However, due to low enzyme yields from *S. lividans*, only an initial characterization could be performed.

In summary, biocatalytic syntheses of acyl amino acids could be established with the help of newly identified aminoacylase. Especially for  $\alpha$ -acylation, the processes of recombinant enzyme production and acylations exceed previously described productivity and yields for aminoacylases.



# 1. Introduction

## 1.1. Acyl-amino acids as biosurfactants

Surfactants and detergents are one of the most used chemicals in households. They are found in soap, laundry detergent, or household cleaning agents, but also in cosmetic products like toothpaste, skin care products, and shampoos and many more. Surfactants or emulsifiers also play an important role as food additives. Due to the often antimicrobial properties, surfactants can be added in disinfectant formulations. Furthermore, surfactants find technical uses in many other industries, like paints or coating, or paper processing. Surfactants act by lowering the surface tension of water and are amphiphilic molecules, consisting of a hydrophobic and a hydrophilic part. The hydrophobic moiety or tail-group is most often aliphatic hydrocarbons. The hydrophilic moiety, also called head-group, is more diverse render the surfactants either non-ionic, anionic, cationic, or amphoteric [1]. Several qualitative measures exist for the properties of surfactants. The efficiency of a surfactant to reduce the surface tension, mostly water, determines its quality. Soaps, salt of fatty acids, are already very effective in that regard. The surface tension of water at room temperature is 72 mN/m, and sodium laurate can lower the tension up to 25 mN/m. However, these substances are highly skin-irritating and deplete the skin of moisture. With sodium lauroyl ether sulfate, also called sodium laureth sulfate (SLES), lowest surface tensions of 30.6 mN/m were measured. Sodium lauryl sulfate (SDS) can lead to a minimal surface tension of 39 mN/m [2]. The limit of surface tension for each surfactant is influenced by their critical micelle concentration (CMC). Since the molecular structure consists of a hydrophilic and hydrophobic part, the surfactant molecule orients itself at the water-air interface, with the hydrophilic head group pointed towards the aqueous phase. At the CMC, the surfactant molecules can no longer occupy the interface and become completely submerged in the water to form micelles. Hence, they can no longer contribute to lowering the surface tension.

The use of soap originates back to ancient civilizations and were made from animal fat or plant oils. With industrialization, detergents from petrochemical origin became more important and surpassed the consumption of soaps in the US by 1953 [3]. However, due to environmental concerns, detergents from natural resources, called biosurfactants or green surfactants, become increasingly sought after. Completely petrochemistry-based synthetic surfactants raise concerns of low biodegradability, high toxicity, and harmful impact on water systems, potentially causing eutrophication. Not least, diminishing petrochemical stocks will lead to increasing shortage of

## 1. Introduction - 1.1. Acyl-amino acids as biosurfactants

raw materials [4]. The use of biosurfactants addresses these concerns, as they are inherently biobased and sustainable, have high biodegradability and are comparable environmentally benign. Regarding their quality, they are not only skin-friendly and low inflammatory, but also have a low CMC, meaning that they are more effective at reducing the surface tension [4].

Due to their low environmental impact, biosurfactants are considered the next generation of industrial surfactants [3]. Biosurfactants can be of various compositions. Phospholipids are an abundant biosurfactant, as they are contained in the cell membrane. A prominent example used in food industry are lecithines, which can be composed of a glycerol molecule linked to phosphorylcholine and two fatty acids. Glycolipids consists of one or more fatty acids linked to a sugar molecule or an oligosaccharide. A third group of biosurfactants are lipopeptides. These are composed of one or more fatty acids linked to oligopeptides. Often, these peptides are circular, like in the case of surfactin from *B. subtilis*, and have antimicrobial properties. In fact, surfactin has been described as one of the strongest known biosurfactants [3]. Furthermore, glycolipopeptides like herbicolin A exist [5]. The oligopeptide moiety of lipopeptides is commonly synthesized non-ribosomically by multienzyme complexes. Hence, like glycolipids, lipopeptides are usually produced *via* microbial fermentation, because *in vitro* synthesis is too complex to be commercially feasible.

Amino acid surfactants (AAS) are simpler compounds related to lipopeptides. In acyl-amino acids, the polar head group represents an amino acid and is linked to a fatty acid. Most commonly, an amide bond between the fatty acid and the  $\alpha$ -amino group of the amino acid is established. For amino acids with alcohol residues, O-acylation can be performed. In general, acyl-amino acids have favorable properties compared to common surfactants like SDS or SLES. Mostly, they are mild, carry little inflammatory potential and low toxicity, exhibit desirable foaming properties, and are biodegradable. Through variation of both the amino acid and the fatty acid moieties, diverse compounds can be synthesized [6]. Depending on the length of the fatty acid, mild surfactants are gained from e.g. lauric acid (C12) or myristic acid (C14), while longer acyl chains like stearic acid (C18) or oleic acid (C18:1) yield emulsifiers. The high chemical diversity of amino acids leads to different characteristics of their acylation products. Most N-acyl-amino acids are anionic, because of the free  $\alpha$ -carboxylic group. N-acyl-glutamates or -aspartates have two negative charges, and are thus relatively more harsh compared to other AAS [2]. The acylation products of lysine and arginine, which have a positive charge in their side chain, yield amphoteric surfactants. O-ethyl-N-lauroyl arginine (LAE) is a cationic surfactant by esterification of the  $\alpha$ -carboxylic group. Reduction of surface tension to 25.4 mN/m has been reported for this compound [7]. Commercial examples from

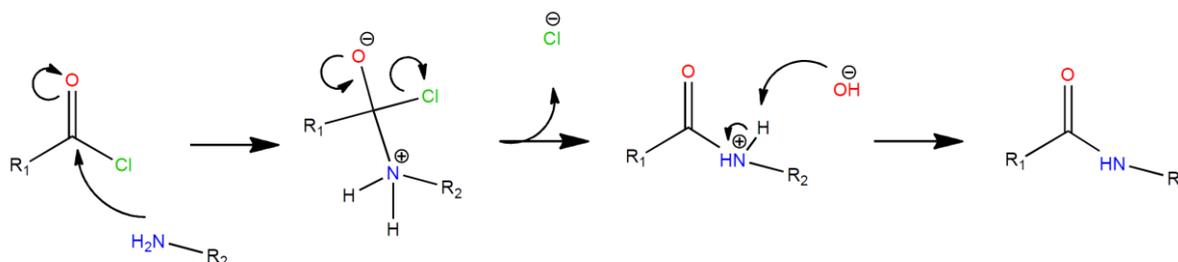
acylation of glutamic acid can be found as N-lauroyl-L-glutamic acid or N-oleoyl-L-glutamic acid. N<sub>α</sub>-lauroyl-L-arginine is commercialized as a mild surfactant and hair conditioner. Being a valuable compound in cosmetics, AAS are remarkably skin-protective. The skin's natural moisturizing factor comprises of amino acids by 40 %. Furthermore, intercellular lipids in the stratum corneum play an important role in prevention of dehydration [8]. Bacterial inhabitants of human skin have been described to carry enzymes that are able to hydrolyze acyl-amino acids [9, 10]. Thus, when decomposed after application, acyl-amino acids can act as skin-protecting agents.

Acyl-amino acids are not only valuable biosurfactants, but also have various physiological functions. They exhibit structural resemblance the endocannabinoid N-arachidonoyl-ethanolamine [11]. Several acyl-amino acids have been found in mammalian brains [12], and they can act as signal molecules, interacting with G-protein coupled receptors or other proteins. Furthermore, they were shown to stimulate mitochondrial oxidative metabolism through uncoupled respiration [13]. Some short-chain acyl-amino acids or -dipeptides are neurotransmitters. N-acetyl-L-aspartate and N-acetyl-aspartyl-glutamate are the most abundant acyl-amino acid and dipeptide found in the mammalian brain, respectively [14, 15]. The neurophysiological and metabolic effects of the compounds open new market prospects as therapeutics. The *in vivo* biosynthesis from fatty acids and amino acids has not been completely clarified, but involves activation of the fatty acid [16]. This can be realized in the form of acyl-phosphates, acyl-adenylates, acyl-CoA thioesters, or with the help of acyl carrier proteins (ACP). On the other hand, the level of endocannabinoid-like acyl-amino acids is regulated by hydrolysis from fatty acid amide hydrolase (FAAH) and the extracellular N-acyl-amino acid hydrolase PM20D1 [17].

To chemically synthesize acyl-amino acids from fatty acids and amino acids, a prior activation of the fatty acid through halogenation is necessary. A system of amino acids and fatty acids is not reactive without a specialized catalyst. To form an amide bond, the free electron pair of the  $\alpha$ -amino group attacks the carbonyl carbon of the fatty acid. Therefore, a sufficiently high pH is necessary to deprotonate the amino group. At basic pH values, however, the fatty acid is deprotonated, thus negatively charged. This renders the fatty acid resistant to a nucleophilic attack by the amino group, as this would lead to two oxyanions in the tetrahedral intermediate. In chemical synthesis, this reaction is performed by using fatty acyl chlorides, called the Schotten-Baumann reaction. The halogenation leads to an uncharged molecule, and the chloride ion is a good leaving group, so that an amide bond is readily established. The mechanism of the Schotten-Baumann reaction for N-acylation is shown in figure 1. However, the halogenation

## 1. Introduction - 1.1. Acyl-amino acids as biosurfactants

and the use of acyl chlorides is what renders the reaction environmentally harmful. As a first step, the fatty acids must be chlorinated with thionyl chloride, phosgene, phosphorous trichloride, or -pentachloride. These chemicals are all highly toxic and pose hazards in their production. The fatty acyl chloride products are more corrosive and irritant than their fatty acid counterparts. The acylation itself is conducted at alkaline pH by the addition of NaOH, which leads to the stoichiometric formation of NaCl as a waste product. The reaction medium is often a system consisting of water and an organic solvent to solubilize the amino acid and the fatty acyl chloride, respectively. The organic solvent can be water-miscible like acetone, tetrahydrofuran, dioxane, or tert-butyl alcohol [18], or a biphasic system with diethyl ether can be used [19]. Not least, for amino acids like lysine with nucleophilic groups in their side chain, protective groups must be introduced for selective synthesis. The need for hazardous chemicals and chlorinated fatty acids, use of organic solvents, and the occurrence of waste materials demonstrate the desire for sustainable, biocatalytic synthesis of N-acyl-L-amino acids. Aminoacylases, the enzymes in focus of this thesis, are promising catalysts for alternative green synthesis of amino acid-based surfactants.



**Figure 1:** Mechanism of the Schotten-Baumann reaction with acyl chloride and amines. The reaction is performed under alkaline conditions through the addition of NaOH. After nucleophilic attack of the amino group at the carbonyl carbon and release of the chloride ion, NaCl and H<sub>2</sub>O are formed as a side product.

## 1.2. Biocatalysis

Enzymes are proteins with catalytic properties. They can lower the activation energy of chemical reactions and thus enable reactions that would not occur, or only slowly, without the catalyst. Enzymes are involved in every cellular process, from metabolism to DNA replication. The use of enzymes for the fermentation of alcoholic beverages reaches back to at least 7000 B.C. and has been found documented in ancient China [20]. Evidence suggests that in ancient Egypt, enzymes were used for brewing and cheese making [21, 22]. Back then, enzymes were used unknowingly, and detailed research unveiling the nature and mechanisms of these biocatalysts only started in the 19<sup>th</sup> century but were already well understood at the end of the 20<sup>th</sup> century. The term enzyme was coined in 1877 by Wilhelm Kühne, and in 1897, Eduard Buchner could show fermentation activity from dead yeast extract similar to vital yeast cells [23]. The substrate specificity of enzymes was explained with the lock-and-key model by Emil Fischer in 1894. The notion of the formation of an enzyme-substrate complex prior to conversion of substrates into products has described by Leonor Michaelis and Maud Menten, to which the mathematical model is nowadays known as the Michaelis-Menten equation [23]. The theory was further elaborated in 1948 by Linus Pauling by proposing that enzymes would stabilize a transition state during the reaction [24]. The idea of the “induced-fit”, that the binding of the substrate to the active site of the enzyme induces a conformational change, was termed by Daniel Koshland in 1958 [25].

The great diversity of catalyzed reactions and accepted substrates hints to the vast number of enzymes found in nature. Hence, means to classify enzymes are necessary. The Enzyme Commission (EC) numbers represent such a system. Based on the catalyzed reaction, they are assorted to various enzyme classes, namely oxidoreductases (EC 1), transferases (EC 2), hydrolases (EC 3), lyases (EC 4), isomerases (EC 5), ligase (EC 6), and translocases (EC 7). Further classifiers narrow down the specific reactions. Aminoacylases, the focal enzymes of this thesis, are assigned the EC number 3.5.1.14, defining hydrolases (3) that act on carbon-nitrogen bonds, other than peptide bonds (3.5) in linear amides (3.5.1), which are N-acyl-aliphatic-L-amino acid amidohydrolases (3.5.1.14).

Even while the scientific community was still working on understanding and unraveling the nature of enzymes, early industrial applications of the biocatalyst can be found. In 1894, the first enzyme that was industrially produced was the fungal amylase takadiastase from *Aspergillus oryzae* [26]. The use of enzymes in laundry detergent was patented as early as 1915 by Otto Röhm [27]. At these early times, the industrial application of enzyme was hampered to

## 1. Introduction - 1.2. Biocatalysis

insufficient accessibility of the biocatalysts leading to great endeavors to obtain the enzymes, often from animal resources. With the arrival of recombinant gene technology, heterologous expression of genes in suitable, often bacterial host organisms, removed previous boundaries. Since then, many biotechnological applications were established, from bulk chemical productions over fine chemicals and pharmaceuticals to enzymes as molecular biology toolboxes. The largest market share comes from technical enzymes, followed by food enzymes and then by enzymes used in animal feed industry. Hydrolases are most frequently used, regarding their market size and their occurrence in various products. Besides proteases, which are dominating the market through their intensive use in laundry detergent [28] and dairy industries, amylases, lipases, cellulases, or further esterases and glycosidases and many more enzymes find industrial use. The previously mentioned use of aminoacylases for obtaining L-amino acids from racemic mixtures of acetyl-amino acids is another example for the use of hydrolytic enzymes.

In view of the use of enzymes for synthesis of chemicals, not only do enzymes enable reactions that are otherwise difficult or impossible to realize with organic chemistry. Often, they allow processes to be realized at relatively mild conditions. While organic synthesis often needs high temperature, high pressure, extreme pH values, organic solvents, toxic chemicals or catalysts, or previously activated substrates, enzymes might realize the reaction under mild conditions. Referring to the synthesis of N-acyl-L-amino acids, the target molecule of this work, this becomes apparent. The use of aminoacylases present an alternative to the Schotten-Baumann reaction, enabling synthesis from free fatty acids and amino acids. The biocatalytic synthesis satisfies many principles of green chemistry [29]:

- (i) *Prevent waste* in form of NaCl.
- (ii) *Better atom economy* since some synthesis steps are abolished.
- (iii) *Less hazardous synthesis*, since chlorination by toxic phosgene or thionyl chloride is circumvented, hence dangerous acyl chlorides are not used, and acylation itself does not need extremely basic pH.
- (iv) *Design of benign chemicals*, as acyl-amino acids are very mild surfactants and less irritating than many petrochemical-derived detergents.
- (v) *Use of benign solvents and auxiliaries*, since synthesis can occur in aqueous buffers and no harmful chemicals need to be eliminated, like the organic solvents tetrahydrofuran (THF) or acetone.
- (vi) *Design for energy efficiency*, as no extreme temperatures or high pressure needs to be applied, and the biocatalytic acylation could also occur at room temperature.
- (vii) *Use of renewable feedstocks*, since both fatty acids and amino acids can be obtained from biological resources and even agricultural wastes, for example sunflower press cakes.
- (viii) *Reduce derivatives*, because chlorinated acyl-derivates are prevented. Furthermore, since acyl-chlorides can also unspecifically react with alcohol groups and may need blocking groups. Their use is often needless in biocatalysis due to the specificity of enzymes.
- (ix) *Catalysis* is, by definition, the principle of biocatalysis, in contrast to stoichiometric use of reagents, in this case for chlorination.
- (x) *Design for degradation* is a given in the case of N-acyl-L-amino acids, since they can be broken down by enzymes, thus being are not harmful to soil and water.
- (xi) *Real-time analysis for pollution prevention* needs to be established in context of green analytical chemistry.
- (xii) *Inherently safer chemistry for accident prevention* also applies when replacing the Schotten-Baumann reaction by aminoacylase biocatalysis.

### 1.2.1. Aminoacylases

L-Aminoacylases are promising enzymes for the biocatalytic synthesis of N-acyl-L-amino acids and are thus the focus of this thesis. The acylation of amino acids with aminoacylases has been shown for streptomycetal aminoacylases, namely the  $\epsilon$ -lysine aminoacylase from *S. mobaraensis* (SmELA) [30] and the  $\alpha$ -aminoacylase (SamAA) and  $\epsilon$ -lysine aminoacylase (SamELA) from *S. ambofaciens* [31], the porcine aminoacylase-1 (pAcy1) [32] and an aminoacylase from *Burkholderia* sp. [33]. Despite being only recently explored for amino acid acylation, aminoacylases have a long history in industrial biotechnology. The industrial use of aminoacylases focused on the hydrolytic activity against acetyl-amino acids. With the enzymes, enantiomerically pure L-amino acids can be obtained from racemic mixtures of chemically synthesized acetyl-amino acids. In fact, the fungal aminoacylase from *Aspergillus oryzae* was one of the first enzymes that found industrial application and was the first enzyme to be used in an immobilized form in 1969 by Tanabe Seiyaku Co. Ltd. [34]. The *Aspergillus* aminoacylase has also been investigated for acylation, but yields were unsatisfactory [35].

Functionally and structurally, aminoacylases share similarities with some homologous peptidases. Peptidases are assigned the EC number 3.4., which defines hydrolases that act on peptide bonds. A peptide bond is formed by two  $\alpha$ -amino acids between the  $\alpha$ -carbonyl group of one amino acid and the  $\alpha$ -amino group of the second amino acid. In N-acyl-amino acids, the carbonyl-nitrogen amide bond is formed between the amino group of the amino acid and an organic acid. Aminoacylases are therefore assigned the EC number 3.5., defined by hydrolysis of C-N bonds that are not peptides. However, from an evolutionary view, regarding sequence conservation, protein structure, and catalytic mechanism, a strong link between aminoacylases and peptidases can be found. Hence, the MEROPS classification for peptidases enables a better overview and description of aminoacylases. The MEROPS database contains information about proteolytic enzymes, their inhibitors and substrates [36]. A holotype, along with a MEROPS identifier, is assigned for each protein species as the first biochemically characterized specimen, and one holotype is selected as the type example for each (sub-)family. Uncharacterized homologs of the holotype are assigned the same MEROPS identifier [37]. It provides a system to identify and narrow down peptidases to families, clans, and individual sequences. The families are distinguished by the catalytic type of their members. All aminoacylases known so far belong to the family of metallopeptidases.

Metallopeptidases cleave peptide bonds by a water molecule activated with a divalent metal cation. This is most often a zinc ion, but cobalt, manganese, nickel, or copper ions can be found,

too. Depending on the number of metal ions in the active site, metallopeptidases can be classified into two groups. The first group only contains a catalytic metal ion, while the second requires two cocatalytic metal ions. Cobalt- or manganese-metallopeptidases are always cocatalytic and all cocatalytic metallopeptidases are exopeptidases [38]. Furthermore, additional structural metal binding sites can be found. The most common metal ligands are histidine, glutamic acid, aspartic acid, and carboxylated lysine (Kcx), but cysteine can also be found. Furthermore, glutamic acid is often described as a non-metal binding residue necessary for catalysis that acts as a general base [38]. There are multiple families of metallopeptidases, which have been grouped to 14 clans. Metallopeptidases from clans MA, MC, MD, ME, MM, MO, MP, and MT contain one metal ion. Members of clans MF, MG, MH, MJ, MN, and MQ have cocatalytic metal sites [38]. By using the MEROPS BLAST tool, metallopeptidases and their non-peptidase homologs, like aminoacylases, can be associated with their respective family [39]. The sequences of most L-aminoacylases can be assigned to the M20 and M38 metallopeptidase families. Some exceptions are the human aspartoacylase (ASPA), also known as aminoacylase-2 (Acy2), which cleaves acetyl-aspartate [40], or the mammalian aminoacylase-3 (Acy3), which deacetylates mercapturic acids, but also acts on N-acetyl-aromatic amino acids [41]. These two enzymes belong to the M14 family of the MC clan. In the following, the catalytic and structural characteristics of the M20 and M38 families are reviewed.

### 1.2.2. Enzymes from the MH clan and M20 family of metallopeptidases

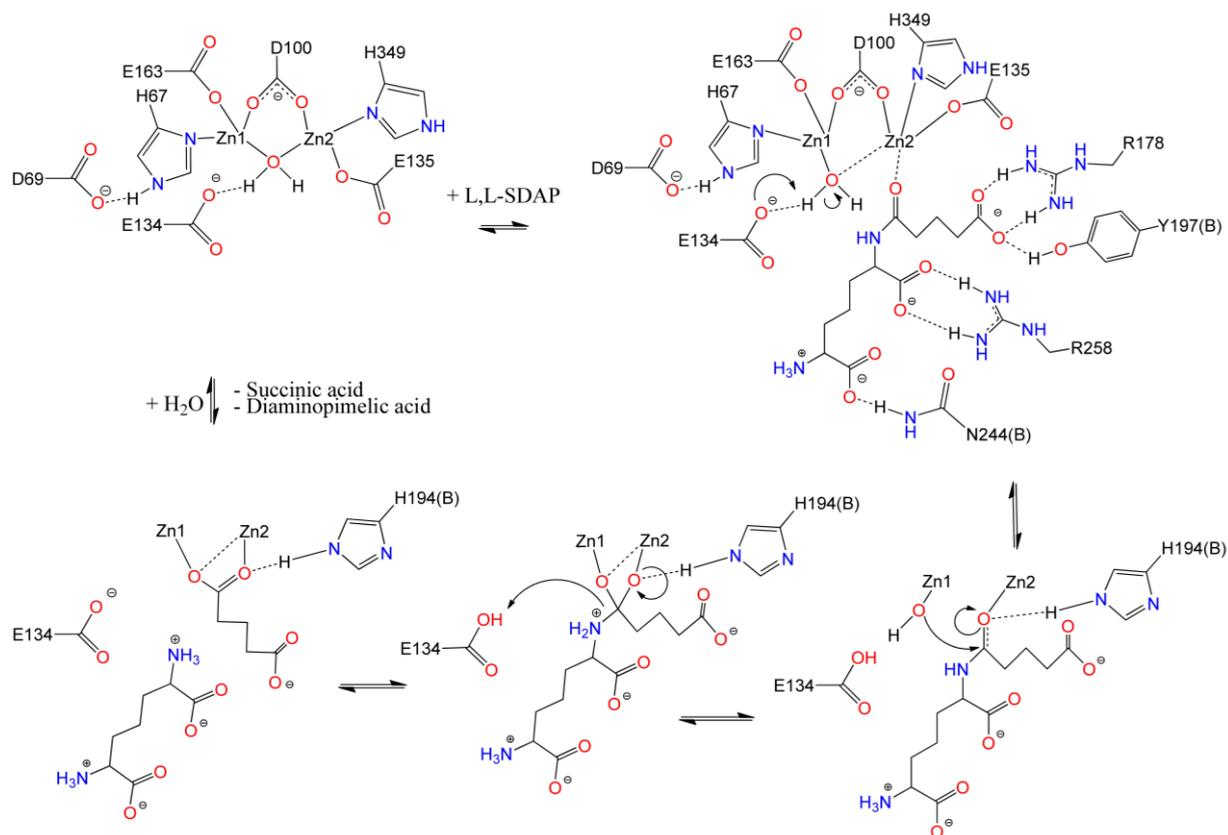
Given that several aminoacylases can be assigned to the MH clan and M20 family, the characteristics of these groups will be explained in some detail. Enzyme members of the MH clan are characterized by possessing two cocatalytic zinc ions, for which the binding sites are conserved among its families. The most common metal ligands in metallopeptidases are histidine, glutamic acid, aspartic acid, and lysine residues, but there are cases where cysteine is a ligand as well. The family M20 is further divided into four subfamilies, namely M20A, M20B, M20C, M20D, and M20F. M20A contains dipeptidases, like peptidase V from *Lactobacillus delbrueckii* (PepV), carboxypeptidases and non-peptidase homologs, like acetylornithine deacetylase from *E. coli* (ArgE), N-succinyl-L,L-diaminopimelic acid desuccinylase from *Haemophilus influenzae* (HiDapE), or aminoacylases like pAcy1. The subfamily M20B contains the aminopeptidase peptidase T from *E. coli*, which acts only on tripeptides. M20C contains a Xaa-His peptidase from *E. coli*, which is specific for dipeptide carnosine  $\beta$ -Ala-His.

## 1. Introduction - 1.2. Biocatalysis

M20D includes the peptidase HmrA from *Staphylococcus aureus*, which binds one zinc ion with H95 and H156, and the second by E129 and H341, while both are linked by C93 [42]. In the recently published crystal structure of aminoacylase CsAga from *Corynebacterium striatum* Ax20, the two cocatalytic zinc ions are also bridged by a cysteine residue [43], suggesting that CsAga belongs to M20D as well. Lastly, the M20F includes the murine carnosine dipeptidase II.

In metallopeptidases, the divalent metal ions, usually zinc, play a central role in the catalytic mechanism. Zinc ions are often coordinated by four or five ligands, but six ligands can also occur, making the ion a versatile coordination partner. It is not involved in redox reactions but acts as a Lewis acid in catalysis if the coordinated ion retains a positive charge [44]. The zinc site acts as a base or nucleophile by deprotonation, converting the bound water to a hydroxide ion [45]. In M20A peptidases, aspartic acid is the bridging residue, and the zinc ions are further coordinated by two histidines and two glutamic acids. The non-ligating nitrogen of histidine can form hydrogen bonds with carboxylates of neighboring glutamic acid and aspartic acid residues, which can assist orientation of the histidine residue by restricting its movement, leading to stronger coordination and decreased Lewis acidity of the zinc ions [46].

The mechanism of hydrolysis of M20A peptidases [38, 45] has been studied, in particular for succinyl-diaminopimelate desuccinylase HiDapE from *Haemophilus influenzae* [47]. The reaction is initiated by binding of the substrate, which interacts with one zinc ion with its peptide- or amide carbonyl oxygen, disrupting the bridging water molecule. The catalytic glutamate acts as a general base by deprotonating the water molecule. The formed hydroxide ion can attack the carbonyl carbon of the substrate, forming a tetrahedral intermediate. The catalytic glutamic acid can then donate a proton to the amide nitrogen, so that the intermediate decomposes, and the products are released. The proposed mechanism of HiDapE is shown in figure 2. Zinc is usually described as the metal for cocatalytic ions for members of the MH clan, but  $\text{Co}^{2+}$  led to higher hydrolytic activity of HiDapE than  $\text{Zn}^{2+}$ , and HiDapE was also active with  $\text{Cd}^{2+}$  and  $\text{Mn}^{2+}$  incorporated to the active site. Mixed ion forms were active as well [48].

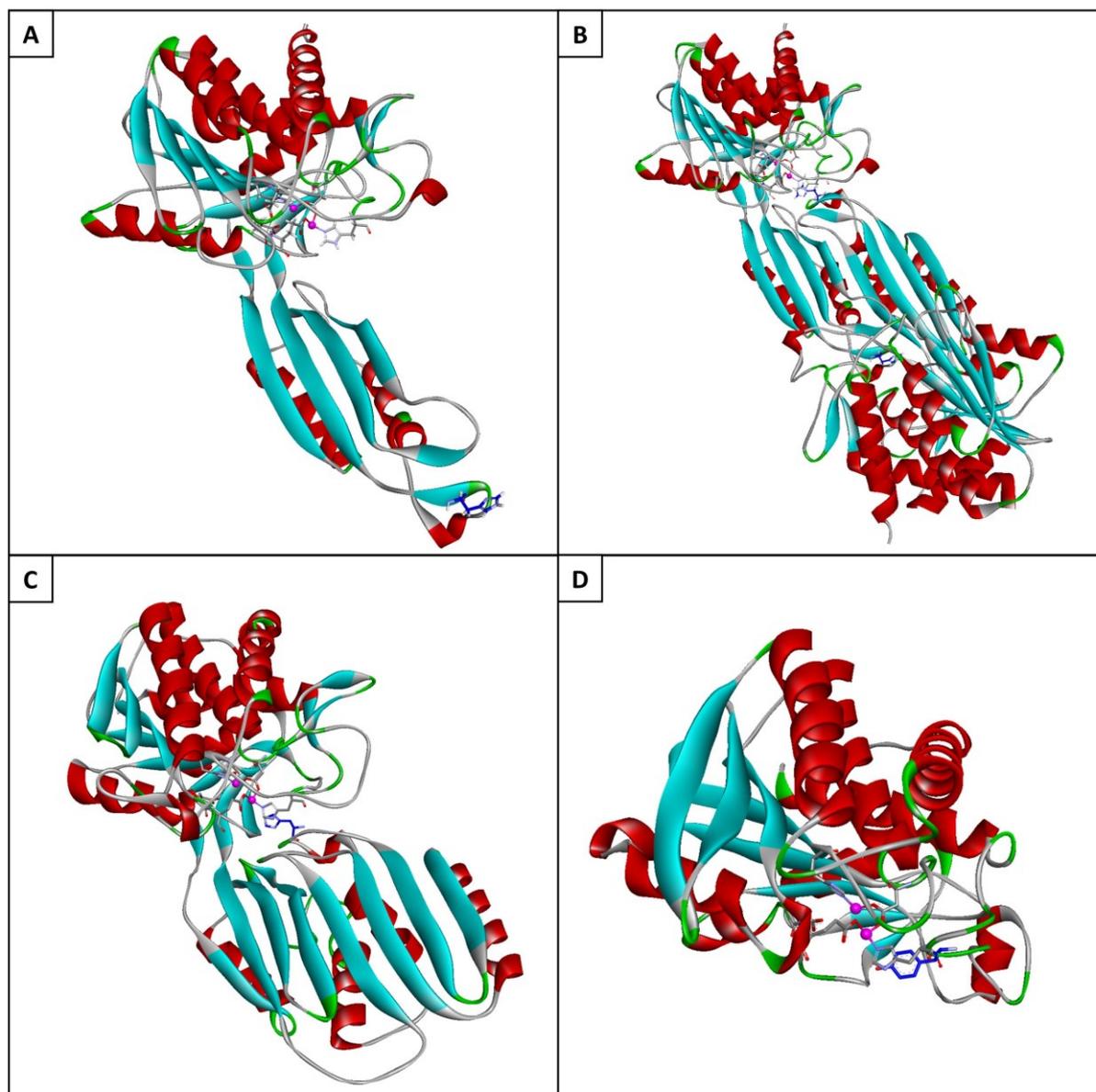


**Figure 2:** Proposed mechanism of the HiDapE-catalyzed hydrolysis of L,L-succinyl diaminopimelate. Adapted and modified from [47]. The electron pushing arrows were also taken from the mechanism of  $\beta$ -alanine synthase from *Saccharomyces kluyveri* [49]. The hydrolytic reaction is initiated by binding of the substrate and deprotonation of the zinc-bound water molecule by E134. The formed hydroxide ion attacks the substrate's carbonyl carbon under the formation of a tetrahedral intermediate. H194(B) contributes to the formation of an oxyanion hole, thus stabilizing the intermediate state, which decomposes and gets protonated by E134. Upon release of the products, a new water molecule can be bound in the active site.

Studies on HiDapE have furthermore shown a conformational change during catalysis and have elucidated the importance of further amino acids residues apart from the cocatalytic zinc site [47]. Members of the M20 family of metallopeptidases have a distinct three-dimensional structure. It is comprised of a catalytic domain and a dimerization or lid domain. The catalytic domain of HiDapE consists of eight adjacent  $\beta$ -sheets, that are twisted around a rotational axis and surrounded by seven  $\alpha$ -helices (PDB 5VO3; Figure 3A). It contains the binuclear zinc active site. The active site of HiDapE consists of the metal-binding residues H67, D100, E135, E163, and H349, and the catalytic residues D69 and E134. These are conserved among members of the M20A family. A similar architecture can be found in the zinc-binding domain of human aminoacylase-1 (PDB 1Q7L), in the structure of carboxypeptidase G2 (PDB 1CG2) from *Pseudomonas* sp. RS-16, and in aminoacylase CsAga (PDB 6SLF) from *C. striatum* Ax20. Since M20 peptidases can be either monomeric or homodimeric, the second domain is called lid- or dimerization domain, respectively. It has been shown for HiDapE, and it is assumed for other members like aminoacylase-1, that dimerization is required for catalysis [50]. The dimeric

## 1. Introduction - 1.2. Biocatalysis

enzyme exists in two conformations, open and closed, with the region connecting its two domains acting as a hinge. Without any bound substrate, the enzyme is in an open conformation. Upon addition of the substrate to the active site, the enzyme undergoes a conformational change, and the hinge closes. This leads not only to the two domains approaching each other, but also to the dimerization domain of the second monomer protruding into the active site of the first monomer. The substrate binding pocket of the active site is bipartite; one side is responsible for binding the amino acid moiety, the other binds the carboxylic acid moiety. In the open conformation, the crescent-shaped cavity is exposed to the solvent. Upon closing of the structure due to substrate binding, the residues of the dimerization domain assist in substrate binding, especially of the amino acid moiety. The apical end of the second dimer also reaches into the active site (Figure 3B). While some residues may also assist in substrate binding, like Y197 of HiDapE, the crucial catalytic function in the context of dimerization lies in the residue H194 (Fig. 2). This histidine residue is described to contribute to the formation of an oxyanion hole for the tetrahedral intermediate with its nitrogen-bound hydrogen and is conserved in human or porcine aminoacylase-1. However, it has been shown for pAcy1 that a H205L mutant exhibited 35 % decreased activity, indicating that the residue is not strictly essential for catalysis [51]. The human aminoacylase-1 (hAcy1) was found to be a dimeric M20A metallopeptidase as well [52]. The crystal structure of the catalytic, zinc-binding domain of a hAcy1 variant has been published (PDB 1Q7L). The porcine pAcy1 and the human hAcy1 are highly homologous and show 87.7 % sequence identity and 91.9 % sequence similarity by the Needleman-Wunsch algorithm [53]. Again, both enzymes have conserved metal-binding (H80, D113, E148, E175, and H373; hAcy1 numbering) and catalytic (D82 and E147; hAcy1 numbering) residues. The human N-fatty-acyl-amino acid hydrolase PM20D1 also belongs to the M20A family and shows sequence identity of 21.5 % to hAcy1.



**Figure 3:** (A) Monomeric HiDapE (5VO3). (B) Dimeric HiDapE (5VO3). (C) PepV (1LFW). (D) SgAP (1QQ9). The  $\alpha$ -helices are shown in red,  $\beta$ -sheets in light blue, the oxyanion hole-forming residue is highlighted in dark blue, and the zinc ions are shown as magenta balls.

Examples for monomeric members of the M20A family are the aminopeptidase PepV [54] and the metallopeptidase from *S. aureus* (Sapep) [55]. In PepV (PDB 1LFW), the histidine that forms the oxyanion hole in dimeric M20 metallopeptidases is not conserved. However, the function of the oxyanion hole is realized in a different manner. In the dimeric HiDapE enzyme, the dimerization domain mainly consists of four parallel  $\beta$ -sheets on top of two  $\alpha$ -helices. At the tip, a short  $\alpha$ -helix and two further short  $\beta$ -sheets are present, with one  $\beta$ -sheet containing the oxyanion-hole forming histidine. In the monomeric PepV enzyme, one half of the lid domain has the same structure of four parallel  $\beta$ -sheets with two  $\alpha$ -helices. However, the lid domain is structurally duplicated, so that the second half contains the same architecture, connected by loops, leading to eight parallel  $\beta$ -sheets on top of four  $\alpha$ -helices (Figure 3C). This

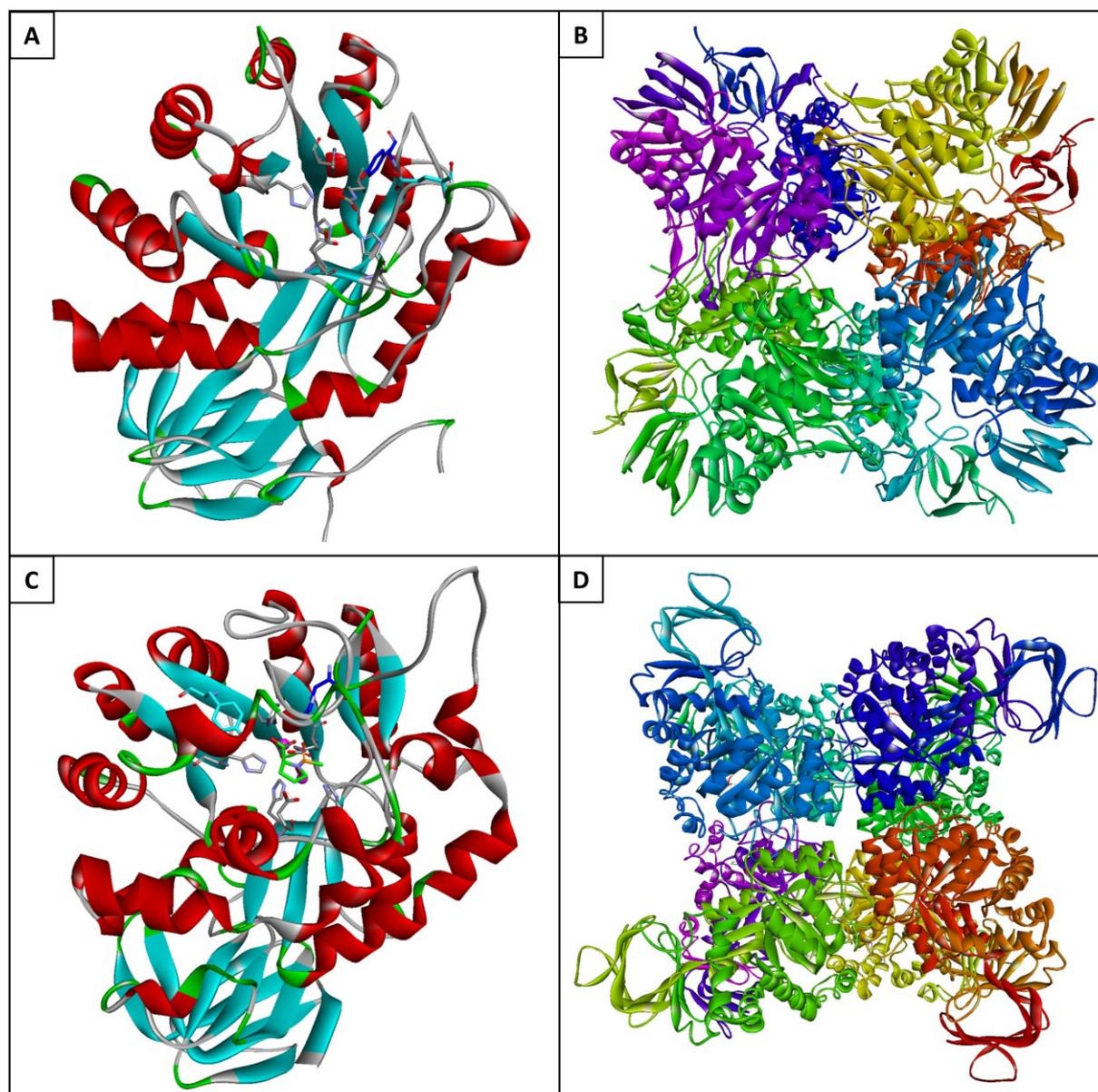
results in the apical tip region having an inverted direction in monomeric PepV, with a histidine residue protruding to the active site. This expanded lid domain has a structure resembling a dimerized form of the dimeric M20A enzymes. A similar structure was observed for Sapep (PDB 3KI9). An exception to the architecture consisting of a catalytic and dimerization or lid domain is the aminopeptidase from *Streptomyces griseus* (SgAP; PDB 1QQ9 or 1CP7) from the M28A peptidase family of the MH clan [56, 57]. This extracellular enzyme only has a catalytic domain and no dimerization or lid domain (Figure 3D). The function of the oxyanion hole-forming residue is filled by a tyrosine residue, instead of a histidine, present in a loop near the active site.

### 1.2.3. Enzymes from the MJ clan and M38 family of metallopeptidases

The amidohydrolase superfamily contains members of the M38 family of the MJ clan. The structure and mechanism of isoaspartyl-dipeptidase from *E. coli* (PDB 1ONW), which catalyzes the hydrolysis of  $\beta$ -aspartyl dipeptides, has been described [58]. The enzyme has a homo-octameric structure and contains a binuclear metal active site for each subunit. The monomeric structure of isoaspartyl-dipeptidase can be divided into an N-terminal domain composed of eight mixed  $\beta$ -strands and a C-terminal domain, that shows a  $(\beta/\alpha)_8$ -barrel fold and contains the active site. In the active site, one metal ion is coordinated by H68, H70, Kcx162, and D285. The second metal ion is coordinated by Kcx162, H201, and H230. The lysine residue is carboxylated at the  $\epsilon$ -amino group to form a carbamate, and bridges the two ions [58]. The tyrosine residue Y137 contributes to the formation of an oxyanion hole and the stabilization of the transition state [59]. Regarding substrate-binding residues, E77 binds the  $\alpha$ -amino group of dipeptide substrates. The  $\alpha$ -carboxylic group of the substrate is bound by the backbone amide groups of G75, T106, and S289. Ion pairs with the  $\alpha$ -carboxylate group of the C-terminal amino acid substrate are formed by R169 and R233 [58]. The structure of 1ONW is shown in figure 4A and 4B.

More recently, further members of the amidohydrolase superfamily and M38 family that show similarities with isoaspartyl-dipeptidase have been analyzed [60]. The structure and active site of the prolidases Sgx9260b from *Paraburkholderia phytofirmans* (PDB 3MKV) and Sgx9260c from *Burkholderia lata* (PDB 3FEQ) have been published. Furthermore, the structure of Sgx9260c was solved in complex with the N-methylphosphonate derivative of L-proline (PDB 3N2C; figure 4C and 4D). Likewise, the structure a homologous carboxypeptidase Cc2672 from *Caulobacter crescentus* was solved in complex with the N-methylphosphonate derivative of L-

arginine (PDB 3MTW). The enzymes have been designated as prolidases and catalyze the hydrolysis of various Xaa-Pro dipeptides, but also N-acyl-prolines like N-acetyl-L-proline or N-propionyl-L-proline. All these enzymes are homo-octamers, with a similar three-dimensional structure to isoaspartyl-dipeptidase. They fold with a N-terminal domain containing eight  $\beta$ -sheets and a C-terminal catalytic domain composed of a  $(\beta/\alpha)_8$ -barrel. The metal-binding residues are also similar in identity, being composed of four histidines, an aspartate and a carboxylated lysine residue (Sgx9260b: H65, H67, Kcx193, H234, H254, D326; Sgx9260c: H65, H67, Kcx190, H231, H251, D323). However, only the first two histidines of the residues are conserved in the primary sequence, the remaining four residues are not. Still, the function of the residues as metal ligands is conserved. The tyrosine residue Y231 (Sgx9260c numbering) interacts with the  $\alpha$ -carboxylic acid of the inhibitor (PDB 3N2C). Furthermore, the oxyanion hole-forming function of these amidohydrolases has been attributed to H142 (Sgx9260c numbering) and is not conserved in sequence position to Y137 of isoaspartyl-dipeptidase.



**Figure 4:** (A) Monomeric and (B) octameric isoaspartyl-dipeptidase (1ONW). (C) Monomeric and (D) octameric prolidase Sgx9260c (3N2C). The  $\alpha$ -helices are shown in red,  $\beta$ -sheets in light blue, the oxanion hole-forming residue is highlighted in dark blue, the  $\alpha$ -carboxylic group-binding residue is highlighted in light blue, the inhibitor is shown with green carbon atoms, and the zinc ions are shown as magenta balls. The octameric structures are colored by enzyme monomers.

An aminoacylase from *Burkholderia* sp. strain LP5\_18B, capable of high-yield acyl-amino acid synthesis, has been identified and characterized [33]. It shows sequence identities of 28-36 % to Sgx9260b, Sgx9260c, and Cc2672. All the metal-binding residues are conserved, as well as the oxanion-hole forming histidine. Despite this similarity to the prolidases, it did not show synthesis of N-lauroyl-L-proline, while multiple other amino acids were efficiently acylated. This indicates differences in substrate binding. Nevertheless, regarding MEROPS classification, the aminoacylase can be assigned to the M38 family as well. The  $\epsilon$ -lysine acylase (EC 3.5.1.17) SmELA from *S. mobaraensis* also shares conserved residues with other members

of the M38 peptidase family, especially with Sgx9260b, Sgx9260c, and the *Burkholderia* sp. aminoacylase. The “HXH”-motif is conserved (H68 and H70), as well as H323, H358, and D424. However, the absence of the conserved (carboxylated) lysine distinguishes the  $\epsilon$ -lysine acylase from the other above mentioned M38 enzymes.

#### 1.2.4. Biocatalytic synthesis of acyl-amino acids with aminoacylases

Most of the investigated aminoacylases belong to the M20A metallopeptidase family. Despite overall low sequence similarity, their members have conserved active site residues and highly conserved structure. The human hAcy1 and the porcine pAcy1 are highly homologous and have been studied in most detail among the aminoacylases. The publications focus more on the function and mechanism of the enzymes, but synthesis of acyl-amino acids has been investigated as well. The aminoacylase pAcy1 has been shown to be a dimeric enzyme [61]. The zinc-binding (H80, D113, E148, E175, H373)- and catalytic residues (D82, E147) are conserved, including the oxyanion hole-forming H206 residue, highlighting the importance of a dimeric structure [62]. Both human and porcine aminoacylase-1 have a broad substrate specificity and can hydrolyze  $N_\alpha$ -acetyl-L-amino acids with diverse side chains. For pAcy1, highest activity was measured with norleucine ((2S)-2-aminohexanoic acid), glutamate, leucine, glutamine and methionine. The human hAcy1 showed a similar substrate scope, with highest activity determined with norleucine, glutamate, leucine, methionine, and glycine [63]. Synthesis of various lauroyl-amino acids has been reported with high conversions in a glycerol-water system [64]. Highest conversion of 81.8 % was observed for the synthesis of lauroyl-arginine, which was explained by low product solubility. Lauroyl-glutamic acid was also produced well with 44.4 % conversion. Other amino acids were acylated with lauric acid with yields between 0.9 % - 35.1 %. Only tyrosine and proline were not accepted among tested substrates. In the study, amino acid concentrations were up to 0.5 M, or 1.0 M for glutamic acid, and lauric acid concentrations were 6.2 mM or 8.3 mM. The water content of the biocatalytic reaction was 12.5 – 33 %, with further reaction parameters being the use of 100 mM phosphate buffer, pH 7.5, a temperature of 37 °C, 24 h reaction time, and 0.33 % w/v pAcy1 used. Because of the small amount fatty acid used for the acylations, final product concentrations remained low. The optimal pH for synthesis was pH 7.0 – 7.5. With the homologous hog kidney aminoacylase-1, conversion yield for the synthesis of N-acetyl-methionine was only 18 % in an aqueous system consisting of 200 mM methionine, 20 mM sodium acetate, 0.12 U aminoacylase, 10  $\mu$ M ZnCl<sub>2</sub> in 200 mM sodium phosphate buffer pH 6.0

[65]. Through protein engineering, a D346A variant of pAcy1 has been generated, which improved synthesis to hydrolysis ratio [32]. However, not least due to aggregation-prone recombinant expression in *E. coli*, predominantly yielding inclusion bodies, and overall low product concentration in synthesis, commercial and industrial biocatalytic applications have not been established with aminoacylase-1 [66]. The human enzyme PM20D1 was also shown to have synthetic activity, especially for oleoyl-phenylalanine. However, possibly due to sub-millimolar substrate concentrations, the conversion was only 1.2 %, so that hydrolytic activity prevailed [67].

The aminoacylase SamAA isolated from *S. ambifaciens* ATCC 23877 also belongs to the M20A family with its conserved active site residues. The enzyme catalyzes the acylation of amino acids with a broad substrate scope and has been characterized with the focus on 10-undecenoyl-phenylalanine. The reaction of 10-undecenoic acid resulted in 5-23 % conversion yield with the non-polar amino acids glycine, alanine, valine, leucine, methionine, phenylalanine, as well as the positively charged lysine and arginine, and polar amino acids serine, cysteine, threonine, glutamine and asparagine [31]. Regarding acyl chain length, the enzyme prefers middle-chain fatty acids like undecenoic acid and lauric acid. The aminoacylase has been characterized with the focus on 10-undecenoyl-phenylalanine. The addition of 0.1 mM CoCl<sub>2</sub> led to the highest activation of the enzyme, whereas 0.1 mM ZnSO<sub>4</sub> had no effect. The optimal pH and temperature for the synthesis were pH 8.0, and 45 °C, respectively. As substrates, 200 mM amino acid and 100 mM fatty acid were used in 25 mM Tris-HCl and 50 mM NaCl, and synthesis was performed for 3 d. The enzyme was isolated from wild type and has not yet been recombinantly produced. Therefore, a partly purified crude extract was used for synthesis at 1 g/l concentration [31, 68].

A homologous aminoacylase isolated from *S. mobaraensis* IFO 13819 (SmAA) has been cloned for expression in *S. lividans* [69]. The enzyme has been characterized regarding its biochemistry and based on its hydrolytic activities. No synthesis has been shown yet with this enzyme. The substrate specificity of the aminoacylase against acetyl-amino acids was relatively broad. The highest hydrolytic activity was measured with acetyl-methionine, followed by acetyl-cysteine and acetyl-alanine. Acetyl-arginine, acetyl-histidine and acetyl-asparagine were also hydrolyzed. The acetyl-derivates of glutamic and aspartic acid were not accepted in hydrolysis. In most cases, activity against acetyl-amino acids was higher than their lauroyl-counterpart. In a hydrolytic assay with acyl-methionines of varying chain length from acetyl- to palmitoyl-residues, octanoyl-methionine was hydrolyzed best. The optimal pH and temperature for hydrolysis were pH 7.0 and 50 °C, respectively. The enzyme was stable in a

pH range of 8.0 to 9.0 at 37 °C for 1 h and at temperatures below 40 - 45 °C at pH 7.5 for 1 h. No loss of activity was measured after addition of 1 mM ethylenediaminetetraacetic acid (EDTA) to the purified enzyme. On the other hand, incubation with 1 mM of the chelating agent 1,10-phenanthroline led to a loss of 90 % activity (15 min, 37 °C). The effect on various divalent metal ions was investigated on restoration of hydrolytic activity (5  $\mu$ M, 1 h, 37 °C). The addition of ZnCl<sub>2</sub> led to highest restored activity, followed by NiCl<sub>2</sub>. However, in contrast to SamAA, CoCl<sub>2</sub> reduced the activity even further. A reduction of activity was also observed with CaCl<sub>2</sub>, CuSO<sub>4</sub>, MgSO<sub>4</sub>, FeSO<sub>4</sub>, and MnSO<sub>4</sub>. The enzyme was found to be monomeric by native gel electrophoresis. However, because of the conserved histidine residue important for the function as a dimeric enzyme [47], this might still be in question.

The aminoacylase CsAga from *C. striatum* Ax20 has been identified and cloned for recombinant expression. No biocatalytic synthesis has been published with this enzyme, but it is worth mentioning because its crystal structure has been published [43]. It belongs to the M20D family of metallopeptidases and has a conserved binuclear zinc-binding site with a cysteine residue bridging the two ions (H107, C105, E141, H168, H370). Crystallography revealed a homotetrameric structure of CsAga, which was shown to be caused by inhibitor binding, and the enzyme naturally occurs as a homodimer [70]. The aminoacylase is specific for N $\alpha$ -acyl-glutamines, but promiscuous towards the acyl moiety, hydrolyzing lauroyl- or decanoyl-glutamine, benzoyl-glutamine, and branched chain 3-methyl-2-hexenoic-glutamine and 3-hydroxy-3-methylhexanoic-glutamine [10]. Other examples for M20D peptidases with aminoacylase activity are carboxypeptidase from *Pyrococcus horikoshii* [71], aminoacylase from *Geobacillus stearothermophilus* [72], or HmrA from *S. aureus* [73] which also exhibit dipeptidase activity.

Another aminoacylase produced by *S. mobaraensis* IFO13819 was found to be an  $\epsilon$ -lysine aminoacylase (SmELA), specific towards lysine acylated at the N $\epsilon$ -position [74]. The conversion rates for lysine acylation even reached 100 % with decanoic acid, lauric acid, and myristic acid, albeit with a high 50-fold excess of lysine and a low fatty acid concentration of 10 mM. The acylations were performed in 100 mM Tris-HCl pH 7.0, at 45 °C, and for 3 d [74]. In a later publication, the aminoacylase gene was cloned and heterologously expressed in *S. lividans*. Yields of 90 – 100 % were observed with 500 mM lysine and 50 mM, 100 mM and 250 mM lauric acid after 6 h, 9 h, and 24 h reaction, respectively, at 37 °C [30]. The aminoacylase is a monomeric protein with an apparent molecular mass of 60 kDa. The enzyme has a binuclear metal active site. Its hydrolytic activity was reduced by 89 % by the chelating agent 1,10-phenanthroline, and divalent zinc, cobalt, magnesium, and manganese ions restored

the activity the most. Zinc restored the activity only at 1-10  $\mu\text{M}$  concentration. The optimal temperature and pH for the hydrolysis of  $\text{N}_\epsilon$ -acetyl-lysine were 55  $^\circ\text{C}$  and pH 8.0-9.0, respectively. The enzyme is stable at a pH range of 6.5 to 10.5 for 1 h at 37  $^\circ\text{C}$ , and for 1 h at 55  $^\circ\text{C}$  at pH 8.0. Since lauroyl-lysines are compounds that are used in high-value cosmetics, the  $\epsilon$ -lysine aminoacylase from *S. mobaraensis* is a promising enzyme for the biocatalytic production due to the high conversions.  $\epsilon$ -Lysine aminoacylases have also been described from other organisms, like the enzyme from *Achromobacter pestifer*, but no synthesis has been shown [75, 76].

The aminoacylase isolated from *Burkholderia* sp. LP5\_18B is a homooctameric enzyme and shows similarity to several amidohydrolases that can be assigned to the M38 metallopeptidase family. The enzyme is capable to synthesize N-lauroyl-L-amino acids in high yields [77]. For example, N-lauroyl-arginine was synthesized with conversions of 89 %. The authors explain the high conversion with precipitation of lauroyl-arginine during the biocatalytic synthesis, which shifts the reaction equilibrium towards the product side. Other lauroyl-amino acids were produced efficiently as well, with highest conversion rates for lauroyl-phenylalanine (51 %), lauroyl-lysine (28 %), and lauroyl-valine (23 %). The reaction conditions for these syntheses were 200 mM amino acid and 100 mM sodium laurate in 100 mM Na-borate pH 9.0, at 25  $^\circ\text{C}$  with 60 h reaction time, using 2  $\mu\text{g}/\text{ml}$  enzyme. The determination of the hydrolytic substrate specificity revealed a preference for hydrophobic amino acids, with lauroyl-alanine, -phenylalanine, and -valine being preferred substrates.  $\text{N}_\alpha$ -lauroyl-lysine was also hydrolyzed well, but  $\text{N}_\epsilon$ -lauroyl-lysine was not accepted. Some lauroyl-amino acids with a polar side chain, like lauroyl-glutamine and lauroyl-serine were hydrolyzed, but to a lesser extent. Lauroyl-amino acids with negatively charged side chain, namely lauroyl-glutamic acid and -aspartic acid, were neither hydrolyzed nor synthesized. The optimal pH-value for synthesis for the enzyme is 9.0. Regarding hydrolytic activity against lauroyl-amino acids, the pH optimum is strongly shifted to basic pH-values, and maximum hydrolytic activity was observed at pH 12.0. At a pH range of 5.0 to 12.0, no loss of activity was detected after 60 min incubation. The highest hydrolytic activity was furthermore measured at 70  $^\circ\text{C}$ , and the enzyme was stable at 70  $^\circ\text{C}$  for 60 min without substantial loss of activity. The effect of several inhibitors and chelating agents on the aminoacylase were tested, but none severely affected hydrolytic activity. Strongest reduction of the enzyme activity was measured with dithiothreitol (DTT) at 78 %, with 2-mercaptoethanol at 81 %, and with EDTA at 87 %. The chelating agents 1,10-phenanthroline and 8-quinolinol had no effect [33]. This indicates that the zinc ions, are tightly bound and hard to remove from the enzyme. In summary, the enzyme is very resistant

to thermal denaturation, and stable in a broad pH range and against chelating agents. Together with the high conversion for acyl-amino acid synthesis, this enzyme is very interesting for industrial applications. An overview of L-aminoacylases described in literature is shown in table 1.

Penicillin acylases (EC 3.5.1.11) are industrially used in synthesis of semi-synthetic antibiotics [23], but were also shown to catalyze the acylation of amino acids. A penicillin V acylase from *S. mobaraensis* NBRC13422 could acylate several amino acids, either with free lauric acid [78], or by acyl transfer from methyl laurate [79]. Despite partially high conversions, final product concentration remained low. A homologous enzyme was identified in *S. ambofaciens*, but was found not to be responsible for the main acylation activity of the cell-free extract [31]. Lipases (EC 3.1.1.3) are also widely used catalysts. Due to their sometimes high substrate promiscuity, they can catalyze further reactions than hydrolysis of lipid-ester bonds, and also act on amides [80]. Lipases remain active in organic solvents, especially when immobilized, which makes them attractive for acylations [81]. N-Acyl-amino alcohols can be synthesized in organic solvent systems to high yields [82]. With lipase B from *Candida antarctica* (CalB), N<sub>ε</sub>-lauroyl-lysine could be synthesized [83]. However, amino acids are insoluble in most organic solvents, and best soluble in aqueous systems. Recently, acylation of glycine in aqueous media was described with an engineered lipase from *Rhizomucor miehei*, that surpassed yields obtained with aminoacylases for this compound [84].

**Table 1:** L-Aminoacylases described in the literature.

The enzyme names were given from the respective publications. In case of  $\epsilon$ -lysine acylases, "(ELA)" was listed as a name to highlight the specificity. If a sequence was given, they were subjected to MEROPS BLAST for classification. If a protein structure has been solved and published, the PDB-ID is annotated. Furthermore, if recombinant synthesis was performed, the respective expression host is listed. Unsuccessful expression of soluble enzyme due to the formation of inclusion bodies is indicated by "IB". Finally, "+" indicates if synthesis of acyl-amino acids was shown in principle with the enzyme, and if the enzyme showed remarkable conversions or product concentrations, it was designated with "+ +". Only enzymes were listed that were at least initially characterized, excluding solely hypothetical proteins. The list is extensive but not necessarily exhaustive, not least due to promiscuous activity of some metalloproteases.

Organism	Name	MEROPS	PDB	Recombinant	Synthesis	Reference
<i>Achromobacter pestifer</i> EA	(ELA)	-	-	-	-	[75, 76]
<i>Alcaligenes denitrificans</i> DA181	-	-	-	-	-	[85]
<i>Arabidopsis thaliana</i>	ILR1, IAR3 ILL1-6	M20D	1XMB	<i>E. coli</i>	-	[86]
<i>Aspergillus melleus</i>	-	-	-	-	-	[87]
<i>Aspergillus oryzae</i>	-	-	-	-	-	[88]
Avian/hog/rat	(ELA)	-	-	-	-	[89]
<i>Bacillus subtilis</i> JH642	AmhX	M20D	-	<i>E. coli</i>	-	[90]
<i>Bacillus thermoglucosidius</i> DSM 2542	-	-	-	-	-	[91]
<i>Brassica campestris</i> L.	-	-	-	-	-	[92]
<i>Burkholderia cepacia</i>	BcepM20D	M20D	-	<i>E. coli</i>	-	[73]

Table 1, continued.

Organism	Name	MEROPS	PDB	Recombinant	Synthesis	Reference
<i>Burkholderia lata</i>	Sgx9260c	M38	3N2C	<i>E. coli</i> , IB	-	[60]
<i>Burkholderia</i> sp. LP5_18B	-	M38	-	<i>E. coli</i> , unsuccessful	++	[33]
<i>Campylobacter jejuni</i> TGH9011	HipO	M20D	-	<i>E. coli</i>	-	[93]
<i>Caulobacter crescentus</i> CB15	Cc2672	M38	3MTW	<i>E. coli</i>	-	[60, 94]
<i>Corynebacterium striatum</i> Ax20	CsAga	M20D	6SLF	<i>E. coli</i>	-	[10, 43]
<i>Deinococcus radiodurans</i> BCRC12827	LAA	M20D	-	<i>E. coli</i>	-	[95]
<i>Escherichia coli</i>	ArgE	M20A	7RSF	<i>E. coli</i>	-	[96]
<i>Geoacillus stearothersophilus</i> NCIB 8224	Ama	M20D	-	<i>E. coli</i>	-	[72, 97]
<i>Heliothis virescens</i>	L-ACY-1	M20A	-	<i>E. coli</i>	-	[98]
<i>Homo sapiens</i>	hAcy1	M20A	1Q7L (partial)	<i>S. frugiperda</i>	-	[62]
<i>Homo sapiens</i>	PM20D1	M20A	-	293A cells	+	[41]
<i>Homo sapiens</i>	ASPA, Acy2	M14	2I3C	<i>E. coli</i>	-	[99, 100]

Table 1, continued.

Organism	Name	MEROPS	PDB	Recombinant	Synthesis	Reference
<i>Lactococcus lactis</i> ssp. <i>cremoris</i> MG1363	Amd1	M20D	-	<i>L. lactis</i>	-	[101]
<i>Micrococcus agilis</i> CCM 2131	-	-	-	-	-	[102]
<i>Mus musculus</i>	AcyIII	M14	3NH4	HEK 293T	-	[41, 103]
<i>Mycobacterium avium</i> Takeo	-	-	-	-	-	[104]
<i>Mycobacterium smegmatis</i> ATCC 607	-	-	-	-	-	[9, 105]
<i>Mycobacterium smegmatis</i> ATCC 607	-	-	-	-	-	[105]
<i>Paraburkholderia phytofirmans</i>	Sgx9260b	M38	3MKV	<i>E. coli</i> , IB	-	[60]
<i>Parkinsonia aculeata</i> L.	-	-	-	-	-	[106]
<i>Pseudomonas diminuta</i> IFO12699	-	-	-	-	-	[107]
<i>Pseudomonas</i> sp. AK2	LpipACY	M38	-	<i>E. coli</i>	-	[108]
<i>Pyrococcus furiosus</i> DSM 3638	-	M20D	-	<i>E. coli</i> , IB	-	[109]
<i>Pyrococcus horikoshii</i> OT3	PhoACY	M20D	-	<i>E. coli</i>	-	[71, 110]
<i>Rhodotorula glutinis</i> AKU 4828	-	-	-	-	-	[111]

Table 1, continued.

Organism	Name	MEROPS	PDB	Recombinant	Synthesis	Reference
<i>Staphylococcus aureus</i>	HmrA	M20D	3RAM	<i>E. coli</i>	-	[73]
<i>Streptomyces ambofaciens</i> ATCC23877	SamAA	M20A	-	<i>E. coli</i> , IB	++	[31, 112]
<i>Streptomyces ambofaciens</i> ATCC23877	SamELA	M38	-	<i>E. coli</i> , IB	+	[112, 113]
<i>Streptomyces coelicolor</i>	ScELA	-	-	<i>S. lividans</i> , <i>E. coli</i> , <i>C. glutamicum</i>	++	[77]
<i>Streptomyces mobaraensis</i> IFO13819	SmAA	M20A	-	<i>S. lividans</i>	-	[69]
<i>Streptomyces mobaraensis</i> IFO13819	SmELA	M38	-	<i>S. lividans</i> , <i>E. coli</i> , <i>C. glutamicum</i>	++	[30, 77]
<i>Streptomyces mobaraensis</i> IFO13819	-	-	-	-	+	[114]
<i>Streptoverticillium olivoreticuli</i> 62-17	-	-	-	-	-	[115]
<i>Sulfolobus solfataricus</i>	CPS <sub>so</sub>	M20F	4MMO	<i>E. coli</i>	+	[116, 117]
<i>Sus scrofa</i>	pAcy1	M20A	-	<i>E. coli</i> , <i>S. frugiperda</i>	+	[51, 66, 118]
<i>Thermococcus litoralis</i>	TliACY	M20D	-	<i>E. coli</i>	-	[119, 120]

### 1.3. Heterologous expression of aminoacylases

#### 1.3.1. Recombinant expression with *Escherichia coli*

In early developments of biocatalytic processes, obtaining the enzyme material was often tedious and a limiting factor, involving costly preparation from the producing organisms or tissues. The arrival of recombinant expression technologies accelerated the development of industrial enzyme applications. However, heterologous expression comes with challenges and production the recombinant protein of interest is not guaranteed. As a host for recombinant protein expression, *E. coli* is often the first and most popular choice. This is due to its genetic accessibility, ease of cultivation and high potential yields of recombinant protein. Around 60 % of recombinant protein production described in research articles use *E. coli* as the expression host [121]. Heterologous expression with *E. coli* also comes with pitfalls. Most prominent is inclusion body (IB) formation, which occurs when misfolded and insoluble proteins aggregate in the cytoplasm. In general, especially when investigating novel proteins, it is desired to prevent IB formation to obtain soluble and active proteins from the cell extract. Different factors in expression of recombinant proteins in *E. coli* influence solubility and the chance of inclusion body formation, including induction mechanisms, cultivation media and parameters or properties of various strains used for expression. The most popular induction mechanism for recombinant expression in *E. coli* is the T7 promoter system [122]. Strains that are compatible with this expression system have a T7 polymerase gene genomically integrated under control of the *lacUV5* promoter. The gene of interest (GOI) is cloned under control of the T7 promoter, which is specific for the T7-polymerase. Transcription is further controlled by the *lacO* operator, preventing leaky expression in the absence of inducer by the LacI repressor. In the absence of inducer, which can be isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) or lactose, both the expression of the T7 polymerase and the GOI are repressed [122]. Upon induction, the T7 system is capable of high-yield expression, with up to 50 % of total cell protein comprised of recombinant protein [123]. However, the induction of the T7-system with IPTG has been described as an “all-or-nothing” approach, as IPTG also induces, besides expression of the gene of interest, the lactose operon, which eventually leads to even higher uptake of IPTG by the lactose permease LacY. The *E. coli* Tuner™ strain (Novagen), a *lacY* deletion mutant that allows to adjust the cellular IPTG concentration, has been developed as a possible solution [124]. Furthermore, autoinduction of the T7-system by cultivation in media that contain glucose and lactose has been proposed, where expression is induced after cells have consumed the glucose, in a milder manner compared to IPTG induction [125].

The lactose autoinduction or deletion of *lacY* can be classified as attempts to tune down the production rate. Further attempts to achieve expression of soluble protein are lowering cultivation temperature, reducing recombinant gene dosage or adjusting promoter strength or choice of promoters [126]. Another solution for expression of aggregation-prone proteins is the co-expression of molecular chaperones. Most commonly, in *E. coli*, the main chaperones GroEL/S and DnaK/J/GrpE are used for co-expression [127]. Some chaperones are constitutively expressed and therefore always present in the bacterial cell, whereas others are stress induced, most prominently by high temperatures, thus coining the term “heat shock proteins” for some chaperones. They were also found to be upregulated in heterologous overexpression, as a form of cellular stress response. However, an overload of the bacterial quality control system due to insufficient chaperones and IB formation might even be supported under stress conditions. This way, misfolded proteins can be temporarily stored for either refolding or proteolysis [128]. In literature, improvement of soluble expression is often achieved by co-expression of GroEL/S, DnaK/J/GrpE, Trigger factor (TF) or combinations thereof [126, 129]. Chaperones typically bind unfolded polypeptides at exposed hydrophobic patches. They hinder nascent chains from forming incorrect intra- and intermolecular interactions, preventing misfolding and aggregation. They can act as holdases by binding to the nascent chains, which can be energy-independent, or as foldases, actively mediating correct folding. In the latter case, energy from adenosine triphosphate (ATP) hydrolysis is used to assist conformational changes.

The GroEL/S chaperonin is a heteromultimeric protein that forms a barrel-like structure. GroEL proteins assemble as two heptameric ring-structures, that can be closed by a lid of a heptameric GroES-complex on both sides. For refolding, a misfolded protein substrate enters the cavity formed by the GroEL rings and gets entrapped by the GroES lid. Multiple rounds of ATP hydrolysis, inducing conformational changes of the chaperonin, partially unfold the protein substrate and mediate assisted refolding. The protein is finally released from the complex. The size of the cavity determines the maximal molecular size of the protein substrate, with limits of approximately 60 kDa being reported [130]. Another chaperone complex that functions together is DnaK/J/GrpE. The DnaK chaperone exercises the main folding activity. It binds substrate protein stretches of four to five hydrophobic residues flanked by basic residues. With ATP bound to DnaK, the substrate binding pocket is open. Upon ATP hydrolysis, a conformation change is induced. This results in increased substrate binding and partial unfolding of the protein substrate, for an attempt to fold properly. The co-chaperone DnaJ

## 1. Introduction - 1.3. Heterologous expression of aminoacylases

increases the rate of ATP hydrolysis and targets the proteins to DnaK. On the other hand, the exchange of ADP with ATP is accelerated by GrpE [127]. Trigger Factor is the first chaperone to interact with nascent polypeptides, as it is associated to ribosomes. Its structure can be separated into three domains: An N-terminal domain for ribosome binding, a peptidylprolyl isomerase domain with auxiliary chaperone activity, and a C-terminal domain responsible for main chaperone activity [131]. Involved in the initial folding steps, Trigger Factor can act co-translationally, preventing premature folding or degradation and restricting access of downstream chaperones. Because a molar excess of TF to ribosomes was observed, TF may also act post-translationally and can be found in a dimeric form at high cytosolic concentrations. While TF associates to ribosomes as a monomer, dimeric TF can bind to proteins to hold them in a folding-competent state and cooperates with the DnaK/J/GrpE system for refolding [132].

In the case of aminoacylases, recombinant expression is often described to be unsuccessful. As an alternative to isolation from tissue, the human and porcine aminoacylase-1 have been produced with the Baculovirus expression system in *Spodoptera frugiperda* insect cells. Initial attempts to express pAcy1 in *E. coli* BL21(DE3) cannot be considered successful. The enzyme could be purified, but barely showed any activity [118]. The use of *E. coli* BL21 Rosetta™ (DE3), which expresses additional tRNAs for rare eukaryotic codons, led to specific pAcy1 activities compared to the enzyme isolated from porcine kidney [51]. Expression of pAcy1 in *E. coli* BL21(DE3) could be enhanced using a codon optimized *pAcy1* gene and by co-expression of molecular chaperones, especially of GroEL/S. Specific activity of the purified enzyme was however only half of the activity from pAcy1 isolated from porcine kidney [66]. The human aminoacylase PM20D1 was expressed with recombinant human 293A cell culture [67]. The aminoacylases SmAA and SmELA from *S. mobaraensis* have only been heterologously produced with *S. lividans* TK24 in their respective scientific publications [30, 69]. However, in a patent filed by Ajinomoto and Okayama University, the expression of SmELA and a homolog from *S. coelicolor* (ScELA) were expressed with *E. coli* JM109, albeit with much lower yields than with *S. lividans* [77]. The homologous aminoacylases SamAA and SamELA have been isolated from the natural producer *S. ambofaciens* ATCC 23877 [31, 113]. Production attempts in *E. coli* did not yield soluble protein [112]. Expression of the aminoacylase from *Burkholderia* sp. LP5\_18B in *E. coli* BL21 (DE3) yielded only traces of aminoacylase activity in the cell extract, and no visible overexpression could be detected in SDS-PAGE [33]. Other aminoacylases that were successfully produced in recombinant *E. coli*

were CsAga from *C. striatum* [10], aminoacylase from *G. stearothermophilus* [72, 97], and aminoacylase from *P. horikoshii* [110].

### 1.3.2. Recombinant expression with *Vibrio natriegens*

*Vibrio natriegens* is a gram-negative bacterium that has drawn attention in biotechnology in recent years and has been proposed as an alternative production host to *E. coli*. The bacterium has already been isolated in 1958 [133], first designated as *Pseudomonas natriegens*, and its strikingly short generation time of less than 10 minutes has been observed (in BHI medium with 1.5 % salt at 37 °C) [134]. It grows on multiple sugars and carbon sources such as D-glucose, D-mannitol, fructose, glycerol, L-rhamnose, sucrose, L-arabinose, D-mannitol, N-acetyl-glucosamine, maltose, gluconate, malic acid, citric acid, but cannot utilize lactose [135, 136]. The fast growth of *V. natriegens* may be attributed to the extraordinary high number of up to 115,000 ribosomes per cell, compared to *E. coli* with 70,000 ribosomes per cell [135, 137].

In 2013, the draft genome sequence of *V. natriegens* DSM 759 (ATCC 14048), the original strain isolated in 1958, has been published [138] and it was revealed that its genome is organized in two chromosomes [139]. Not only the fast growth and availability of the genome, but also the close genetic proximity to *E. coli* renders *V. natriegens* interesting for biotechnology, with many genetic elements and established molecular biology tools also functioning in the latter organism. For example, the *tet* promoter [139], *trc* promoter, *araBAD* promoter, and several common origins of replication were shown to be functional in *V. natriegens*. To establish a host for recombinant protein expression alternative to *E. coli*, a functional T7 expression system has been constructed to generate the strain *V. natriegens* Vmax™ [140]. The T7 polymerase gene has been put under the control of the IPTG-inducible *lacUV5* promoter, regulated by *lacI*, so that the T7 promoter can be used for gene expression. The T7 polymerase expression cassette has been integrated into the *dns* locus, and the resulting knockout of the exonuclease improved plasmid stability [140].

Another research group has generated the strain *V. natriegens* VnDX in a similar manner as *V. natriegens* Vmax [141]. This strain was tested and compared to *E. coli* BL21(DE3) in pET-based expression of 196 genes of interest (GOIs). From these, 65 % and 75 % experienced soluble expression in *V. natriegens* VnDX and *E. coli* BL21(DE3), respectively. Higher expression levels in *V. natriegens* were observed for 20 GOIs, which was about 10 % of the investigated constructs. In contrast, 102 GOIs were better expressed in *E. coli*. In both hosts,

## 1. Introduction - 1.3. Heterologous expression of aminoacylases

47 GOIs yielded no overexpression and 27 GOIs were expressed equally. The GOI library was intended to represent a diverse set of genes, including most enzyme classes from various biological origins, and with a broad span of sequence lengths. The *V. natriegens* VnDX system was further improved for T7 expression by two approaches. First, the *lacUV5* promoter was changed to the tetracycline promoter *Ptet*, leading to 109 % increased recombinant glucose dehydrogenase activity. Second, the ribosome binding site was varied, which led to an increase of recombinant enzyme activity by 12.6 % [142]. Furthermore, expression of multisubunit membrane protein complexes has been shown in *V. natriegens* [143] and even codon-suppression based incorporation of non-canonical amino acids in heterologous proteins could be achieved [144]. These results show the high potential of *V. natriegens* in biotechnology and the convenience of many technologies developed for *E. coli* functioning in the organism.

### 1.3.3. Recombinant expression with *Streptomyces lividans*

Streptomycetes are aerobic, Gram-positive bacteria, which exhibit filamentous growth. They carry a linear genome with a high G+C content over 70 % [145]. The genus is commonly known as natural producers of antibiotics and fungicides. As saprobionts, they have a high level of secretion of extracellular enzymes. Biotechnologically, *S. lividans* is widely used for recombinant expression. Secretion of recombinant proteins is most often realized through the Sec pathway, but Tat secretion is also efficient in this species [146, 147]. Some streptomycetal aminoacylases were expressed in *S. lividans* [30, 69], which is the most often used expression host among *Streptomyces*. The close genetic relation with the donors of the respective gene of interest can facilitate successful expression. Furthermore, the naturally high G+C content can abolish the need for codon optimization of streptomycetal genes.

*Streptomyces* does not tend to form inclusion bodies, which one of the prevalent issues of recombinant expression with *E. coli* [148]. Furthermore, particularly *S. lividans* are described as having low endogenous proteolytic activity [149]. Hence, *S. lividans* can be considered as a valuable alternative expression host to *E. coli*, especially for enzymes of streptomycetal origin. Commonly used strains are *S. lividans* TK23 (*spc-1*, SLP2<sup>-</sup>, SLP3<sup>-</sup>) and TK24 (*str-6*, SLP2<sup>-</sup>, SLP3<sup>-</sup>), which are sensitive to spectinomycin and streptomycin, respectively, and lack the small linear plasmids (SLP) [145]. Replicative, high-copy vectors are often used for expression, but low-copy or integrative plasmids have also been used [149]. Expressed proteins are often hydrolases, and the GOIs are sourced most often from Gram-positive bacteria. Constitutive

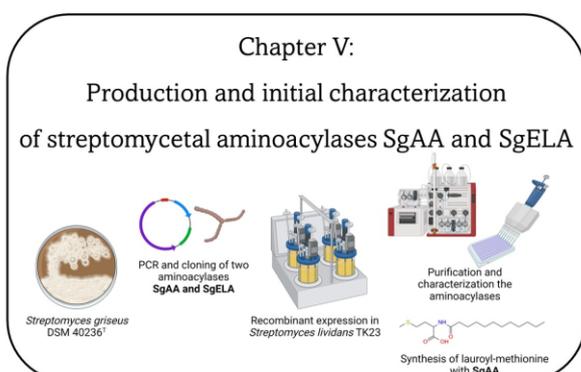
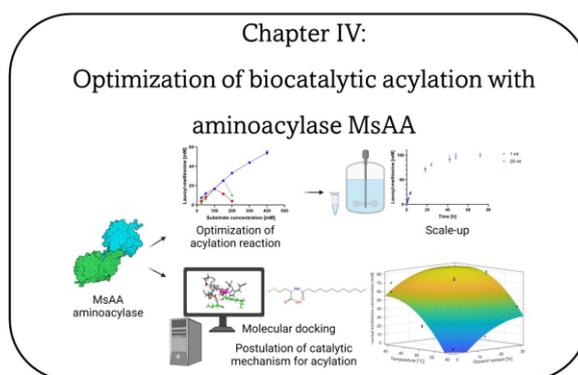
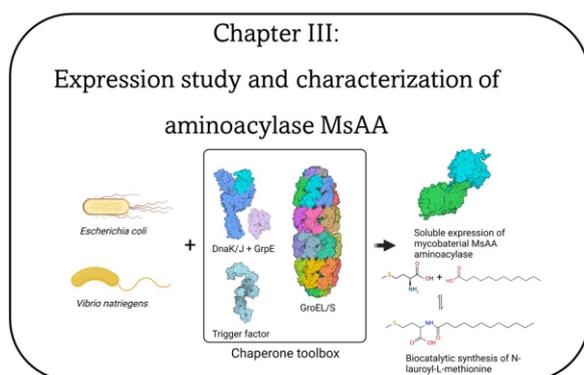
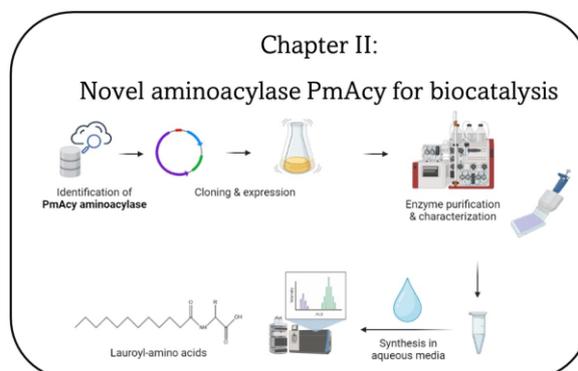
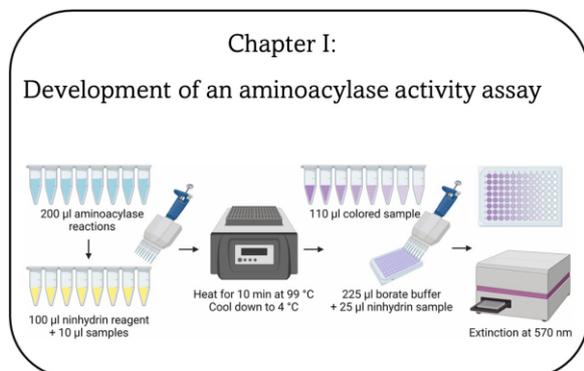
promoter systems have the highest usage, followed by inducible promoters, and by the respective native promoter [149]. Examples for promoters frequently found in literature are *P<sub>vsI</sub>* from *S. venezuelae*, *PermE\** from *Saccharopolyspora erythraea*, or the inducible *P<sub>lysA</sub>* from *S. halstedii* [149]. In summary, *S. lividans* is an interesting expression host with established recombinant technologies and should be considered for expression of streptomycetal aminoacylases.

## 1.4. Objective of the thesis

The work of this thesis aims at establishing the biocatalytic synthesis of N-acyl-L-amino acids in aqueous media. These amino acid-based surfactants find use in cosmetic products and have very desirable tensioactive and low-inflammatory properties. By using enzymes, it is intended to present an alternative to replace the conventional chemical Schotten-Baumann reaction that uses fatty acyl chlorides. Biocatalysis should allow the use of free fatty acids and amino acids. Hence, the synthesis of completely biobased surfactants should be enabled. The enzymes in focus are L-aminoacylases, which are primarily known for racemic cleavage of N-acetyl-DL-amino acids to produce enantiomerically pure L-amino acids. The reverse reaction catalyzes the synthesis of N-acyl-amino acids. However, the use of aminoacylase for surfactant production is still underexplored in literature and unexploited in industry. Novel aminoacylases are intended to be identified from homology search. The criteria for suitable enzymes comprise mainly their synthetic potential, with further important characteristics being stability of the catalysts, a broad substrate scope, and preferably a bacterial origin.

The strategy to pursue this objective was to first recombinantly express the selected aminoacylases, initially in *E. coli*, and to perform purification *via* an affinity tag. Because heterologous expression of aminoacylase has often been described as difficult, the establishment of suitable expression platforms was part of this thesis. The aminoacylases were planned to be biochemically characterized upon purification, regarding pH- and temperature optima and stabilities, substrate specificity, dependency on metal ions, or occurrence of multimers. For the measurement of hydrolytic aminoacylase activity, a suitable activity assay needed to be developed. Following the characterization by hydrolysis, the aminoacylases should be evaluated regarding their biocatalytic potential. First, a screening of synthetic activity with all proteinogenic amino acids using an acyl donor should elucidate the synthetic substrate scope. Afterwards, the reaction conditions were intended to be optimized, focusing on maximizing conversion for one product compound. Typically, optimal pH-value, temperature, and substrate concentrations need to be determined. Another approach was the bioinformatical assessment of protein sequences and prediction of protein structures, both to gain insight into mode of action, mechanism, or substrate binding, and to even assess possibilities for rational protein engineering. The novel aminoacylases selected in this thesis were PmAcy from *Parakburkholderia monticola* DSM 100849, MsAA and MsELA from *Mycolicibacterium smegmatis* MKD 8, and SgAA and SgELA from *S. griseus* DSM 40236<sup>T</sup>.

## 2. Results



## **2.1. Chapter I**

### **A convenient ninhydrin assay in 96-well format for amino acid-releasing enzymes using an air-stable reagent**

Gerrit Haeger, Johannes Bongaerts, Petra Siegert

Analytical Biochemistry 654, 114819 (2022), doi: 10.1016/j.ab.2022.114819

#### Author contributions:

GH: Conceptualization, investigation, data acquisition, visualization, writing - original draft.

JB: Funding acquisition, supervision. PS: Project administration, supervision, writing - review & editing, validation.

Overall contribution GH: 95 %



Contents lists available at ScienceDirect

Analytical Biochemistry

journal homepage: [www.elsevier.com/locate/yabio](http://www.elsevier.com/locate/yabio)

## A convenient ninhydrin assay in 96-well format for amino acid-releasing enzymes using an air-stable reagent

Gerrit Haeger, Johannes Bongaerts, Petra Siegert\*

*Institute of Nano- and Biotechnologies, Aachen University of Applied Sciences, 52428, Jülich, Germany*

### ARTICLE INFO

#### Keywords:

Ninhydrin  
Aminoacylase assay  
Amino acid-releasing enzymes  
96-Well format  
Air-stability  
N<sub>α</sub>-acyl-lysines

### ABSTRACT

An improved and convenient ninhydrin assay for aminoacylase activity measurements was developed using the commercial EZ Nin™ reagent. Alternative reagents from literature were also evaluated and compared. The addition of DMSO to the reagent enhanced the solubility of Ruhemann's purple (RP). Furthermore, we found that the use of a basic, aqueous buffer enhances stability of RP. An acidic protocol for the quantification of lysine was developed by addition of glacial acetic acid. The assay allows for parallel processing in a 96-well format with measurements microtiter plates.

### 1. Introduction

Ninhydrin is widely used for both manual and automated quantification of amino acids. Under acidic, reducing conditions and high temperature, it reacts efficiently with primary amino groups to form diketohydrindylidene-diketohydrindamine, also called Ruhemann's purple. Since its use in automated chromatography has been described, many ready-to-use solutions for HPLC systems have been developed [1]. One well known challenge, however, lies in its insufficient air-stability. During the reaction with amino acids, ninhydrin needs to be reduced, or hydrindantin, a reduced form of ninhydrin, is added directly to the reaction mixture [2]. Since the reducing agents are susceptible to oxidation, the system must be constantly kept under inert gas. For automated systems like HPLC, this can be readily realized, both using commercial reagents or reagents prepared by the researcher. For manual batch analysis, however, the issue remains, as the reagent reservoirs must be opened regularly. With insufficient stability, the reagents need to be gassed out with nitrogen and regular standards must be prepared to ensure accurate measurements. This demonstrates the necessity of an air-stable reagent that facilitates handling.

While dominantly used in automated amino acid analysis, ninhydrin has been used in enzyme assays early on [3]. Especially for aminoacylase (EC 3.5.1.14) activity, quantification of released amino acid with ninhydrin is the method of choice for photometric assays [4–18]. Also, other amino acid-releasing enzymes, like peptidases, have been assayed with ninhydrin [19–22]. Unfortunately, most sources present rather

inconvenient, complicated protocols with volumes in milliliter-scale, not being suitable for high throughput. Furthermore, the presented ninhydrin reagents and protocols differ vastly.

We developed a ninhydrin method for measurement of aminoacylase activity. These enzymes release free amino acids by deacylation. For this purpose, multiple reagents from literature were compared and tested for their suitability. Special respect was paid to ease of reagent preparation, stability of the reagents and to establishing a format that allows for parallel processing of the samples. We found that a commercial reagent intended for use in chromatography systems, trademarked EZ Nin™ (Biochrom Ltd, UK), was most adequate for this purpose. The reagent is air-stable for three years by the use of a temperature-sensitive reducing agent [23]. To enhance the solubility of formed RP, the reagent was diluted with equal volumes of DMSO, henceforth referred to EZ Nin: DMSO. We used the reagent for batch analyses and established its use for aminoacylase activity assays.

### 2. Procedure of the aminoacylase activity assay

An aminoacylase assay was established using the porcine aminoacylase-1 (pAcy1, Sigma Aldrich). A typical aminoacylase reaction is composed of an N-acylamino acid in aqueous buffers. In our case, we used 15 mM N-acetyl-L-methionine in 50 mM Tris-HCl pH 7.0 at 30 °C and 20 µg pAcy1 in a final volume of 200 µl. As Fig. 1 (C) shows, the concentration increased linearly. The specific activity of the lyophilizate, calculated by the slope of the curve, was determined to be 20

\* Corresponding author. Heinrich-Mussmann-Str. 1, 52428, Jülich, Germany.  
E-mail address: [siegert@fh-aachen.de](mailto:siegert@fh-aachen.de) (P. Siegert).

<https://doi.org/10.1016/j.ab.2022.114819>

Received 12 May 2022; Received in revised form 7 July 2022; Accepted 9 July 2022

Available online 15 July 2022

0003-2697/© 2022 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

U per mg enzyme preparation under these conditions. In the following, a stepwise protocol is described:

- (i) Prepare ninhydrin samples by adding 100  $\mu$ l of the ninhydrin reagent to 0.2 ml tubes.
- (ii) Prepare aminoacylase reaction mixtures by adding 190  $\mu$ l substrate solution to fresh 0.2 ml tubes and pre-incubate at the desired temperature (15.79 mM acyl-amino acid in 50 mM Tris-HCl pH 7.0 in a volume of 190  $\mu$ l).
- (iii) Add 10  $\mu$ l enzyme solution and mix to start the assay (e.g. 0.4 mg/ml pAcy1).
- (iv) At suitable time points, withdraw 10  $\mu$ l samples from the aminoacylase reaction and immediately mix with 100  $\mu$ l of the ninhydrin reagent (e.g. sampling intervals of 30 s).
- (v) Heat for 10 min at 99  $^{\circ}$ C and cool the samples down to 4  $^{\circ}$ C.
- (vi) In microtiter plates, fill each well with 225  $\mu$ l 100 mM Na-borate buffer pH 10.0.
- (vii) Add 25  $\mu$ l of the colored ninhydrin samples, mix thoroughly and measure extinction at 570 nm.

Special attention was paid to the scale and format in which the assays were conducted. The goal was to establish a 96-well format in all steps for facile handling using multichannel pipettes and microtiter plates. Thus, both enzyme reactions and ninhydrin reactions were conducted using 0.2 ml reagent tubes. However, all steps can be performed in 1.5 ml-tubes as well. It is recommended to use a dispenser pipette to quickly add 100  $\mu$ l ninhydrin reagent to multiple 0.2 ml tubes. Furthermore, a multichannel pipette can be used to simultaneously start multiple reactions and to simultaneously withdraw samples from multiple reactions and to add to ninhydrin reagent sample tubes. For cysteine and proline, a slight variation needs to be applied. With these amino acids, a yellow to brown color is formed that is best measured at 410 nm.

For color development, a 10  $\mu$ l sample was added to 100  $\mu$ l EZ Nin: DMSO, mixed and heated for 10 min at 99  $^{\circ}$ C and afterwards cooled to 4  $^{\circ}$ C in a thermocycler (Biometra, Analytik Jena). For measurement, the

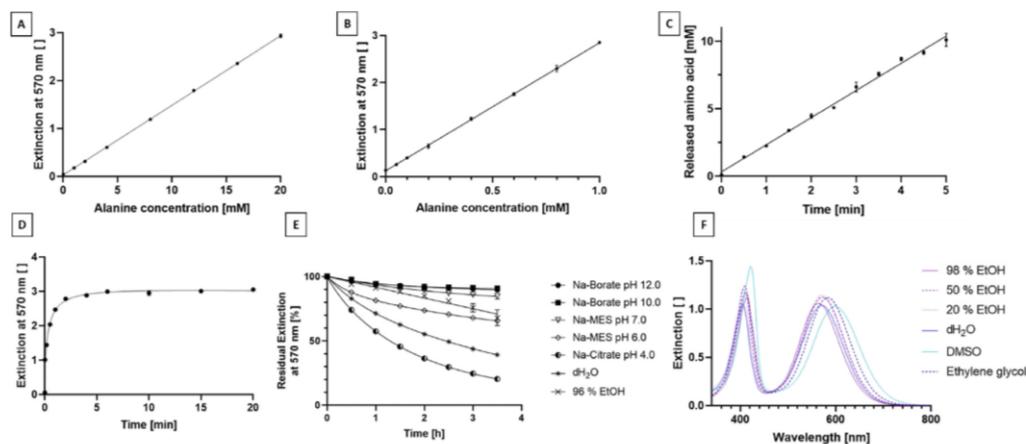
samples were transferred to microtiter plates and diluted with 100 mM Na-borate, pH 10.0. The dilution can be varied at the discretion of the researcher. A suitable dilution would be tenfold, meaning 25  $\mu$ l of the ninhydrin reaction sample was mixed with 225  $\mu$ l of the borate buffer. For cysteine and proline, it is recommended to mix 150  $\mu$ l 100 mM Na-borate buffer pH 10.0 with 100  $\mu$ l of the ninhydrin sample per well. The amino acid concentration in the sample was calculated with a standard of the respective amino acid. An amino acid standard up to 20 mM is shown in Fig. 1 (A) with very good linearity ( $R^2 = 0.99995$ ; measured in triplicates). The heating time for color development was set to 10 min but may even be shortened. As shown in Fig. 1 (D), final color yield is already reached after 5 min.

For higher sensitivity in sub-millimolar amino acid concentrations, the volumes of the ninhydrin reagent and the sample can be adjusted. For example, 200  $\mu$ l of EZ Nin:DMSO can be mixed with a 50  $\mu$ l sample. After heating, 200  $\mu$ l can be transferred into microtiter plates and mixed with 50  $\mu$ l Na-borate buffer pH 10.0. The resulting standard from 15  $\mu$ M–1 mM L-alanine is shown in Fig. 1 (B).

### 3. Choice of ninhydrin reagents

Regarding the ninhydrin reagents used in aminoacylase activity measurements, even recent publications [13,14,16] are based on Moore & Stein (1954) [1], their improvement for batch processing by Rosen (1957) [24] or by Yemm & Cocking (1955) [2]. As every reagent has its (dis)advantages, the most suitable reagent for each application must be chosen from extensive literature study. Besides the choice of organic solvents and acids used in the reagent, a lot of reagents contain different reducing agents. Here, we present a comprehensive comparison of various ninhydrin reagents that addresses these differences.

For qualitative analysis of amino acids, ninhydrin can be solely dissolved in alcohol, e.g. ethanol or isobutanol, mixed and heated with amino acids solutions [25]. For forensic fingerprint techniques or staining in thin layer chromatography, it is sufficient to dissolve ninhydrin in ethanol for a spraying solution. Indeed, amino acids will



**Fig. 1.** (A) Alanine standard in millimolar range measured in a microtiter plate. 10  $\mu$ l amino acid samples were mixed with 100  $\mu$ l EZnin:DMSO (1:1). The samples were diluted 1:10 in 100 mM Na-borate buffer pH 10.0 to a final volume of 250  $\mu$ l. Correlation coefficient  $R^2$  was 0.99995, the slope was 0.1451  $\text{mM}^{-1}$ . (B) Alanine standard in micromolar range measured in a microtiter plate. 50  $\mu$ l amino acid samples were mixed with 200  $\mu$ l EZ Nin:DMSO. 200  $\mu$ l of the samples were transferred to a microtiter plate and mixed with 50  $\mu$ l 100 mM Na-borate buffer pH 10.0. Correlation coefficient  $R^2$  was 0.9998, the slope was 2.719  $\text{mM}^{-1}$ . (C) Aminoacylase reaction: 20  $\mu$ g pAcy1 were added to 15 mM N-acetyl-L-methionine in 50 mM Tris-HCl pH 7.0 at 30  $^{\circ}$ C at a final volume of 200  $\mu$ l. Every 30 s, a 10  $\mu$ l sample was withdrawn and immediately mixed with 100  $\mu$ l EZ Nin:DMSO. The specific activity was determined to be 20 U per mg lyophilisate. (D) Influence of heating time (0 s–20 min) on color development in ninhydrin assay with EZ Nin:DMSO. 20 mM alanine was used as an amino acid sample. Nonlinear fit was applied to visualize the signal increase (GraphPad Prism). (E) Samples from ninhydrin reaction with 16 mM alanine standard measured in a microtiter plate in 30-min intervals. The samples were diluted 1:10 in  $\text{dH}_2\text{O}$ , various 100 mM buffers or in 96% ethanol. Nonlinear fits were applied to visualize the signal decrease of each dataset (GraphPad Prism: two phase decay, least squares fit). (F) UV/Vis-spectra of ninhydrin sample from reaction with 16 mM alanine, diluted tenfold with various solvents. Each spectrum was measured in triplicates. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

react with ninhydrin to form RP when a 2% solution of ninhydrin in absolute ethanol was used as in the protocol described above. However, the efficiency of color formation is very low, and the amino acid standard does not behave linearly, as the sensitivity increases with higher amino acid concentration (data not shown). A linear standard curve is necessary for quantitative analysis. When the reagent is composed of 2% ninhydrin in 50% ethanol and 0.1 M citric acid pH 5.0, the color yield significantly increases, however, the sensitivity still is dependent on the amino acid concentration. Only if the acidic buffer and a reducing agent, in this case 0.8 mg/ml SnCl<sub>2</sub> dihydrate, are added to the reagent, a linear amino acid standard can be obtained (Table 1). In order to find the reagent best suitable for manual batch analysis, various reagents from literature and commercially available reagents that contain acidic buffers and reducing agents were used with our protocol. The reagents published in literature were carefully prepared as described. However, the reagents were only initially purged with nitrogen gas and not kept

under inert gas.

A list of representative ninhydrin reagents, their composition and sensitivity measured with an L-alanine standard is summarized in Table 1. Traditionally, the toxic 2-methoxyethanol (methyl cellosolve) was used as the solvent for ninhydrin [26]. When hydrindantin was used to replace the reducing agent, the solvent was switched to DMSO, which is considered less harmful, to enhance solubility and stability of hydrindantin [27]. The commercial TRIONE® (Pickering Laboratories, USA) contains hydrindantin as well, but uses sulfolane as the organic solvent. The use of ethylene glycol was later described to enhance stability of the reagent compared to 2-methoxyethanol and DMSO, when stannous chloride was used as the reducing agent [28]. Concerning the acidic buffers, it appears that acetic acid pH 5.2 is most widely used, either paired with lithium- or sodium hydroxide, while citric acid is used with stannous chloride, presumably to prevent its precipitation.

More crucial is the choice of the reducing agent itself, which is needed for quantitative analysis. Stannous chloride is widely used but is susceptible to oxidation, and thus requires working under nitrogen gas. Titanium (III) chloride is described to replace stannous (II) chloride without disadvantage [29]. Several other reducing agents have been described, like sodium borohydride [30], sodium cyanide [24] or cadmium (II) chloride [22], which we explicitly avoided because of their toxicity. The reducing agent can be replaced by direct addition of hydrindantin, a reduced form of ninhydrin, to the reagent [27,31]. It is also known that reagents that contain hydrindantin show poor stability in the presence of air due to oxidation of the molecule. This again requires working under nitrogen, making manual handling tedious and frequent amino acid standards necessary. To our best knowledge, the only ninhydrin reagent resistant to oxidation is the EZ Nin™ reagent (invented by JPP Chromatography, UK and distributed by Biochrom, UK). The reagent contains a patent-protected, temperature-sensitive reducing agent, making it stable under air [23]. The reagent was hence investigated for batch processing. When using the unmodified reagent in batch processing of amino acid samples, the formed RP quickly precipitated (8 mM concentration in sample). Presumably, this results from poor solubility of the pigment in ethylene glycol, which is the organic solvent of the reagent. In order to increase solubility of the formed RP, the EZ Nin reagent was mixed in equal parts with anhydrous DMSO, henceforth referred to as EZ Nin:DMSO. The precipitation of RP was also observed with the ethylene glycol reagent described by Starcher [28], which had to be mixed with DMSO as well for batch processing. Furthermore, in view of the intended use as an enzyme assay, DMSO will act as an efficient inactivation agent for the enzyme upon sampling. In conclusion, we recommend using the EZ Nin:DMSO reagent for aminocyclase activity assays or batch analysis in general. In Table 2, the sensitivity of the EZ Nin:DMSO reagent against all canonical amino acids as well as some non-proteinogenic amino acids are shown. As an alternative, non-commercial reagent, we recommend either the reagent described by Zhang et al. [19] or the reagent described by Starcher [28], when mixed 1:1 with DMSO.

#### 4. Acidic ninhydrin reagent for quantification of lysine

With conventional ninhydrin reagents, lysine and N<sub>α</sub>- or N<sub>ε</sub>-acylated lysines cannot be distinguished, as they provide the same color yield (Table 2 or [32]). Therefore, an acidic ninhydrin method had to be developed [33]. The EZ Nin reagent was mixed to equal parts with glacial acetic acid (EZ Nin:GAA). With the EZ Nin:GAA reagent, lysine developed an orange color, while N<sub>α</sub>-acetyl-lysine remained colorless, and N<sub>ε</sub>-acetyl-lysine yielded a slight purple color. The procedure can be used to measure hydrolysis of N<sub>α</sub>-acyl-lysines or to study N<sub>ε</sub>-lysine-acylases [17]. Again, to propose an alternative, non-commercial reagent, the ninhydrin reagent described by Zhang et al. [19] was mixed with glacial acetic acid for comparison. Furthermore, the four reagents described by Work [33] were investigated as well. The results are shown in Table 1.

**Table 1**  
Comparison of various ninhydrin reagents.

Reference	Reagent composition	Sensitivity against L-alanine and comments
JPP, UK [23], modified	EZ Nin™ (Biochrom, UK); Patent protected (contains ninhydrin, ethylene glycol, acetic acid, and a temperature-sensitive reducing agent); mixed 1:1 with DMSO	0.1451 [E <sub>570</sub> /mM] Stable at air and ambient temperature; No need for regular reagent and standard preparation
Zhang et al., 2013 [19]	2.5% ninhydrin, 50% (v/v) DMSO, 100 mM citric acid pH 5.0, 0.80 mg/ml SnCl <sub>2</sub> ·2 H <sub>2</sub> O	0.1531 [E <sub>570</sub> /mM] N <sub>2</sub> mandatory to prevent oxidation
Sun et al., 2006 [31]	2% ninhydrin and 0.3% hydrindantin in 75% DMSO, 1 M sodium acetate pH 5.2	0.1390 [E <sub>570</sub> /mM] Based on [27] with replacement of lithium acetate; N <sub>2</sub> mandatory to prevent oxidation
Starcher 2001 [28], modified	2% ninhydrin in 75% (v/v) ethylene glycol, 1 M sodium acetate pH 5.5 and 2.44 mg/ml SnCl <sub>2</sub> ; mixed 1:1 with DMSO	0.1457 [E <sub>570</sub> /mM] Ethylene glycol is described to enhance stability of the reagent; High viscosity; N <sub>2</sub> mandatory to prevent oxidation
Sigma-Aldrich Corp., USA	Product sheet N7285; based on Moore 1968 [27]; 2% ninhydrin, contains hydrindantin in DMSO and lithium acetate buffer pH 5.2	0.1250 [E <sub>570</sub> /mM] Based on [27]; N <sub>2</sub> mandatory to prevent oxidation
Moore 1968 [27]	2% ninhydrin and 0.3% hydrindantin in 75% DMSO, 1 M lithium acetate pH 5.2	0.1422 [E <sub>570</sub> /mM] N <sub>2</sub> mandatory to prevent oxidation
Moore & Stein, 1948 [26]	2% ninhydrin, 50% (v/v) 2-methoxyethanol, 100 mM citric acid pH 5.0, 0.80 mg/ml SnCl <sub>2</sub> ·2 H <sub>2</sub> O	0.1229 [E <sub>570</sub> /mM] N <sub>2</sub> mandatory to prevent oxidation
This work	2% ninhydrin, 50% (v/v) ethanol, 100 mM citric acid pH 5.0, 0.80 mg/ml SnCl <sub>2</sub> ·2 H <sub>2</sub> O	0.0978 [E <sub>570</sub> /mM] Very low stability of the reagent
<b>Acidic reagents for lysine</b>		<b>Sensitivity against L-lysine</b>
EZ Nin:GAA	EZ Nin™ mixed 1:1 with glacial acetic acid	0.0888 [E <sub>460</sub> /mM]
Zhang:GAA [19], modified	Reagent described by Zhang [19] mixed 1:1 with glacial acetic acid	0.0617 [E <sub>460</sub> /mM]
Work 1957 [33]	250 mg ninhydrin dissolved in (a) 6 ml glacial acetic acid, 4 ml 6 M H <sub>3</sub> PO <sub>4</sub> (b) 6 ml glacial acetic acid, 4 ml 0.6 M H <sub>3</sub> PO <sub>4</sub> (c) 6 ml glacial acetic acid, 4 ml H <sub>2</sub> O (d) 6 ml glacial acetic acid	(a) 0.0255 [E <sub>460</sub> /mM] (b) 0.0557 [E <sub>460</sub> /mM] (c) 0.0490 [E <sub>460</sub> /mM] (d) 0.0498 [E <sub>460</sub> /mM] Low sensitivities

Sensitivity was calculated from amino acid standards; 10 µl amino acid samples were mixed with 100 µl ninhydrin reagents; 25 µl of colored ninhydrin samples were diluted with 225 µl 100 mM Na-borate pH 10.0 and measured at 570 nm (410 nm for cysteine and proline); In case of acidic ninhydrin reagents, 100 µl of colored ninhydrin samples were diluted with 150 µl 100 mM Na-borate pH 10.0 and measured at 460 nm.

**Table 2**

Summary of the three ninhydrin assay variants with regards to sensitivity and color development.

EZ Nin:DMSO reagent		
Compound	Sensitivity [E <sub>570</sub> /mM]	Alanine color equivalent
L-Alanine	0.1451	1.00
D-Alanine	0.1495	1.03
β-Alanine	0.1260	0.87
L-Alanyl-L-phenylalanine	0.1618	1.12
Ammonium chloride	0.1415	0.98
L-Arginine	0.1397	0.96
L-Asparagin	0.1129	0.78
L-Aspartic acid	0.1235	0.85
L-Glutamic acid	0.1451	1.00
L-Glutamine	0.1544	1.06
L-Glycine	0.1392	0.96
L-Histidine	0.1241	0.86
DL-Histidine	0.1275	0.88
DL-Homocystein	0.1245	0.86
L-Isoleucine	0.1463	1.01
L-Leucine	0.1440	0.99
L-Lysine	0.1593	1.10
DL-Lysine	0.1638	1.12
N <sub>ε</sub> -Acetyl-L-lysine	0.1484	1.02
N <sub>α</sub> -Acetyl-L-lysine	0.1559	1.07
L-Serine	0.1427	0.98
L-Methionine	0.1326	0.91
L-Phenylalanine	0.0788	0.54
D-Phenylalanine	0.0796	0.55
L-3,4-Dihydroxyphenylalanine	0.0833	0.57
L-Phenylalanyl-L-alanine	0.1293	0.89
L-Threonine	0.1464	1.01
L-Tryptophan	0.0775	0.53
D-Tryptophan	0.0834	0.58
L-Tyrosine	0.0942	0.65
L-Valine	0.1447	1.00
L-Cysteine <sup>a</sup>	0.1115	–
DL-Cysteine <sup>a</sup>	0.1065	–
L-Proline <sup>a</sup>	0.1641	–
<b>EZ Nin:DMSO reagent, high volume</b>		
L-Alanine <sup>b</sup>	2.7192	1.00
L-Phenylalanine <sup>b</sup>	1.4560	0.54
<b>EZ Nin:GAA reagent</b>		
L-Lysine <sup>c</sup>	0.0888	–
DL-Lysine <sup>c</sup>	0.0826	–
N <sub>ε</sub> -Acetyl-L-lysine <sup>c</sup>	0	–
N <sub>α</sub> -Acetyl-L-lysine <sup>c</sup>	0.0053	–

If not stated otherwise, 25 μl samples were diluted with 225 μl 100 mM Na-borate pH 10.0.

<sup>a</sup> 100 μl ninhydrin reaction sample added to 150 μl 100 mM borate buffer pH 10.0; measured at 410 nm.

<sup>b</sup> Adjusted volumes: 50 μl sample to 200 μl EZ Nin:DMSO; 200 μl ninhydrin reaction sample added to 50 μl borate buffer pH 10.0.

<sup>c</sup> 100 μl ninhydrin reaction sample added to 150 μl borate buffer pH 10.0; measured at 460 nm.

## 5. Dilution in microtiter plates

It has been described that RP undergoes acidic hydrolysis in aqueous solutions [32]. When the colored samples are diluted with water in microtiter plates, the extinction should decrease over time because of the acidic nature of the ninhydrin reagents. To encounter the issue, various 100 mM dilution buffers were investigated, including acidic citric acid buffer pH 4.0 as a negative control. As expected, the extinction signal quickly decreases in water and even faster in citric acid. With borate buffers at pH 10.0 and 12.0, no significant loss in signal was detected. Conventionally, solutions containing RP are still often diluted in ethanol or isopropanol [31,34,35], which may evaporate, posing an inconvenience especially in microtiter plates. The effect of various solvents on the absorption spectrum of RP, formed by the reaction with L-alanine, were analyzed (Fig. 1, E). While the spectrum and observable color varied only slightly between water, ethanol and ethylene glycol,

the color shifted from a purple to a turquoise color when DMSO was used (Fig. 1, F). This indicates that aqueous solutions can be used for dilution of RP with no disadvantage. Hence, we recommend and used 100 mM sodium borate buffer pH 10.0 for dilution of the colored ninhydrin samples.

## CRediT authorship contribution statement

**Gerrit Haeger:** Conceptualization, Investigation, Data acquisition, Visualization, Writing - original draft. **Johannes Bongaerts:** Funding acquisition, Supervision. **Petra Siegert:** Project Administration, Supervision, Writing - review & editing, Validation.

## Acknowledgments

This work was financially supported by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF).

## References

- [1] S. Moore, W.H. Stein, A modified ninhydrin reagent for the photometric determination of amino acids and related compounds, *J. Biol. Chem.* (1954) 907–913.
- [2] E.W. Yemm, E.C. Cocking, The determination of amino-acids with ninhydrin, *Analyst* (1955) 209–214.
- [3] S. Nagai, Enzymatic hydrolysis of N-Palmitoyl-Amino acids by *Mycobacterium avium*, *J. Biochem.* (1961) 428–433.
- [4] J. Matsuno, S. Nagai, Amidohydrolases for N-short and long chain acyl-L-amino acids from *mycobacteria*, *J. Biochem.* (1972) 269–279.
- [5] H. Fukuda, S. Iwade, A. Kimura, A new enzyme: long acyl aminoacylase from *Pseudomonas diminuta*, *J. Biochem.* (1982) 1731–1738.
- [6] H.-Y. Cho, K. Tanizawa, H. Tanaka, K. Soda, Thermostable Aminoacylase from *Bacillus thermoglucosidius*: purification and characterization, *Agric. Biol. Chem.* (1987) 2793–2800.
- [7] M. Dion, F. Loussouarn, N. Batisse, C. Rabiller, V. Sakanyan, Use of the overexpressed *Bacillus stearothermophilus* aminoacylase for the resolution of D,L-amino acids in conventional and non-conventional media, *Biotechnol. Lett.* (1995) 905–910.
- [8] V. Sakanyan, L. Desmarez, C. Legrain, D. Charlier, I. Mett, A. Kochikyan, A. Savchenko, A. Boyen, P. Falmagne, A. Pierard, N. Glansdorff, Gene cloning, sequence analysis, purification, and characterization of a thermostable Aminoacylase from *Bacillus stearothermophilus*, *Appl. Environ. Microbiol.* (1993) 3878–3888.
- [9] E. Wada, M. Handa, K. Imamura, T. Sakiyama, S. Adachib, R. Matsunob, K. Nakanishia, Enzymatic synthesis of N-acyl-L-amino acids in a glycerol-water system using acylase I from pig kidney, *JAACS (J. Am. Oil Chem. Soc.)* (2002) 41–46.
- [10] Z. Liu, Z. Zhen, Z. Zuo, Y. Wu, A. Liu, Q. Yi, W. Li, Probing the catalytic center of porcine aminoacylase 1 by site-directed mutagenesis, homology modeling and substrate docking, *J. Biochem.* 139 (2006) 421–430.
- [11] K. Tanimoto, N. Higashi, M. Nishioka, K. Ishikawa, M. Taya, Characterization of thermostable aminoacylase from hyperthermophilic archaeon *Pyrococcus horikoshii*, *FEBS J.* 275 (2008) 1140–1149.
- [12] Y.B. Yang, K.M. Hsiao, H. Li, H. Yano, A. Tsugita, Y.C. Tsai, Characterization of D-aminoacylase from *Alcaligenes denitrificans* DA181, *Biosci. Biotech. Biochem.* 56 (1992) 1392–1395.
- [13] A. Kołodziejczak-Radzimska, F. Ciesielczyk, T. Jesionowski, A novel biocatalytic system obtained via immobilization of aminoacylase onto sol-gel derived ZrO<sub>2</sub>-SiO<sub>2</sub> binary oxide material: physicochemical characteristic and catalytic activity study, *Adsorption* 25 (2019) 855–864.
- [14] J. Li, Z. Zhao, T. Mo, L. Wang, P. Li, Immobilization of aminoacylase on electrospun nanofibrous membrane for the resolution of dl-theanine, *J. Mol. Catal. B Enzym.* 116 (2015) 24–28.
- [15] J. Kósáry, C.S. Sisak, B. Szajani, L. Boross, Acylation of amino acids by aminoacylase in non-conventional media, *Biocatalysis* 11 (1994) 329–337.
- [16] M. Koreishi, Y. Nakatani, M. Ooi, H. Imanaka, K. Imamura, K. Nakanishi, Purification, characterization, molecular cloning, and expression of a new aminoacylase from *Streptomyces mobaraensis* that can hydrolyze N-(middle/long)-chain-fatty-acyl-L-amino acids as well as N-short-chain-acyl-L-amino acids, *Biosci. Biotech. Biochem.* 73 (2009) 1940–1947.
- [17] M. Koreishi, R. Kawasaki, H. Imanaka, K. Imamura, K. Nakanishi, A novel e-lysine acylase from *Streptomyces mobaraensis* for synthesis of Ne-Acyl-L-lysines, *JAACS (J. Am. Oil Chem. Soc.)* (2005) 631–637.
- [18] M. Koreishi, F. Asayama, H. Imanaka, K. Imamura, M. Kadota, T. Tsuno, K. Nakanishi, Purification and characterization of a novel aminoacylase from *Streptomyces mobaraensis*, *Biosci. Biotech. Biochem.* (2005) 1914–1922.
- [19] Y. Zhang, Y. Fu, S. Zhou, L. Kang, C. Li, A straightforward ninhydrin-based method for collagenase activity and inhibitor screening of collagenase using spectrophotometry, *Anal. Biochem.* 437 (2013) 46–48.

- [20] J. Mukherjee, N. Webster, L.E. Llewellyn, Purification and characterization of a collagenolytic enzyme from a pathogen of the great barrier reef sponge, *Rhopaloeides odorabile*, PLoS One 4 (2009), e7177.
- [21] S.K. Marathe, M.A. Vashistht, A. Prashanth, N. Parveen, S. Chakraborty, S.S. Nair, Isolation, partial purification, biochemical characterization and detergent compatibility of alkaline protease produced by *Bacillus subtilis*, *Alcaligenes faecalis* and *Pseudomonas aeruginosa* obtained from sea water samples, J. Gen. Eng. Biotech. 16 (2018) 39–46.
- [22] E. Doi, D. Shibata, T. Matoba, Modified colorimetric ninhydrin methods for peptidase assay, Anal. Biochem. (1981) 173–184.
- [23] L.J. Pitts, M.G. Pallot, P. Jones, A Method for Analysing Amino Acids and Reagent for Use with the Same - EP 2735876 A1, European Patent Application, 2014.
- [24] H. Rosen, A modified ninhydrin calorimetric analysis for amino acids, Arch. Biochem. Biophys. (1957) 10–15.
- [25] W.H. Fitzpatrick, Spectrophotometric determination of amino acids by the ninhydrin reaction, Science (1949) 469.
- [26] S. Moore, W.H. Stein, Photometric ninhydrin method for use in the chromatography of amino acids, J. Biol. Chem. 176 (1948) 367–388.
- [27] S. Moore, Amino acid analysis: aqueous dimethyl sulfoxide as solvent for the ninhydrin reaction, J. Biol. Chem. 243 (1968) 6281–6283.
- [28] B. Starcher, A ninhydrin-based assay to quantitate the total protein content of tissue samples, Anal. Biochem. 292 (2001) 125–129.
- [29] L.B. James, Amino acid analysis: the reduction of ninhydrin reagent with titanous chloride, J. Chromatogr. (1971) 178–180.
- [30] S. Standara, M. Drdák, M. Veselá, Amino acid analysis: reduction of ninhydrin by sodium borohydride, Nahrung (1999) 410–413.
- [31] S.-W. Sun, Y.-C. Lin, Y.-M. Weng, M.-J. Chen, Efficiency improvements on ninhydrin method for amino acid quantification, J. Food Compos. Anal. 19 (2006) 112–117.
- [32] M. Friedman, Applications of the ninhydrin reaction for analysis of amino acids, peptides, and proteins to agricultural and biomedical sciences, J. Agric. Food Chem. 52 (2004) 385–406.
- [33] E. Work, Reaction of ninhydrin in acid solution with straight-chain amino acids containing two amino groups and its application to the estimation of  $\alpha\epsilon$ -diaminopimelic acid, Biochem. J. (1957) 416–423.
- [34] S. Prochazkova, K.M. Vårum, K. Østgaard, Quantitative determination of chitosans by ninhydrin, Carbohydr. Polym. (1999) 115–122.
- [35] Y.K. Mok, V. Arantes, J.N. Saddler, A  $\text{NaBH}_4$  coupled ninhydrin-based assay for the quantification of protein/enzymes during the enzymatic hydrolysis of pretreated lignocellulosic biomass, Appl. Biochem. Biotechnol. 176 (2015) 1564–1580.

## 2.2. Chapter II

### **Novel recombinant aminoacylase from *Paraburkholderia monticola* capable of N-acyl-amino acid synthesis**

Gerrit Haeger, Tristan Jolmes, Sven Oyen, Karl-Erich Jaeger, Johannes Bongaerts, Ulrich Schörken, and Petra Siegert

Applied Microbiology and Biotechnology 108, 93 (2024)

#### Author contributions:

GH: Data collection and analysis (Investigation, identification, cloning & expression, characterization, initial synthesis screening). TJ: Data collection and analysis (Chemical and biocatalytic synthesis, instrumental analytics, and structure elucidation). SO: Data collection (Supported expression experiments and characterization). GH and TJ wrote the manuscript, JB, KEJ, US and PS revised the manuscript. JB, PS, US did funding acquisition, project design and supervision.

Overall contribution GH: 85 %



## Novel recombinant aminoacylase from *Paraburkholderia monticola* capable of N-acyl-amino acid synthesis

Gerrit Haeger<sup>1</sup> · Tristan Jolmes<sup>2</sup> · Sven Oyen<sup>1</sup> · Karl-Erich Jaeger<sup>3,4</sup> · Johannes Bongaerts<sup>1</sup> · Ulrich Schörken<sup>2</sup> · Petra Siegert<sup>1</sup>

Received: 9 June 2023 / Revised: 12 November 2023 / Accepted: 23 November 2023  
© The Author(s) 2024

### Abstract

N-Acyl-amino acids can act as mild biobased surfactants, which are used, e.g., in baby shampoos. However, their chemical synthesis needs acyl chlorides and does not meet sustainability criteria. Thus, the identification of biocatalysts to develop greener synthesis routes is desirable. We describe a novel aminoacylase from *Paraburkholderia monticola* DSM 100849 (PmAcy) which was identified, cloned, and evaluated for its N-acyl-amino acid synthesis potential. Soluble protein was obtained by expression in lactose autoinduction medium and co-expression of molecular chaperones GroEL/S. Strep-tag affinity purification enriched the enzyme 16-fold and yielded 15 mg pure enzyme from 100 mL of culture. Biochemical characterization revealed that PmAcy possesses beneficial traits for industrial application like high temperature and pH-stability. A heat activation of PmAcy was observed upon incubation at temperatures up to 80 °C. Hydrolytic activity of PmAcy was detected with several N-acyl-amino acids as substrates and exhibited the highest conversion rate of 773 U/mg with N-lauroyl-L-alanine at 75 °C. The enzyme preferred long-chain acyl-amino-acids and displayed hardly any activity with acetyl-amino acids. PmAcy was also capable of N-acyl-amino acid synthesis with good conversion rates. The best synthesis results were obtained with the cationic L-amino acids L-arginine and L-lysine as well as with L-leucine and L-phenylalanine. Exemplarily, L-phenylalanine was acylated with fatty acids of chain lengths from C8 to C18 with conversion rates of up to 75%. N-lauroyl-L-phenylalanine was purified by precipitation, and the structure of the reaction product was verified by LC-MS and NMR.

### Key points

- A novel aminoacylase from *Paraburkholderia monticola* was cloned, expressed in *E. coli* and purified.
- The enzyme PmAcy exhibits exceptional temperature and pH stability and a broad substrate spectrum.
- Synthesis of acyl amino acids was achieved in good yields.

**Keywords** Biosurfactants · Acyl-amino acids · Acylation · Aminoacylase · Biocatalysis · Chaperone

✉ Petra Siegert  
siegert@fh-aachen.de

- <sup>1</sup> Institute of Nano- and Biotechnologies, Aachen University of Applied Sciences, 52428 Jülich, Germany
- <sup>2</sup> Faculty of Applied Natural Sciences, TH Köln University of Applied Sciences-Leverkusen Campus, Leverkusen, Germany
- <sup>3</sup> Institute of Molecular Enzyme Technology, Heinrich Heine University Düsseldorf, 52425 Jülich, Germany
- <sup>4</sup> Institute of Bio- and Geosciences IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

### Introduction

Amino acid-based surfactants, especially acyl-amino acids, are molecules of interest for use in cosmetic products. They act as especially mild detergents and carry little inflammatory potential. In addition, they possess desirable foaming properties, are biodegradable, and have low toxicity (Zhao et al. 2019). The fatty acyl residue mediates lipophilic properties, while the amino acid acts as a polar head group. Variations in both the amino acid and the fatty acid moieties can generate structural variability (Pinheiro and Faustino 2017). Exemplarily, mild surfactant properties are gained with, e.g., lauric acid (C12) or myristic acid (C14) as acyl groups, whereas longer acyl chains like stearic acid (C18) or oleic acid (C18:1) yield emulsifiers. Being a valuable compound

in cosmetics, amino acid-based surfactants are remarkably skin-protective.

While acyl-amino acids are environmentally benign regarding application and biological degradation, the conventional chemical synthesis of these compounds via the Schotten-Baumann reaction is not sustainable (Ibrahim et al. 2020). Chlorinated fatty acids have to be used which are typically produced with toxic phosgene or thionyl chloride. During chlorination and acylation significant amounts of waste products like phosphoric or sulfuric acid and sodium chloride are produced. Additionally, not all acyl-amino acids can be synthesized with the same efficiency. Secondary amino groups have to be protected in order to prevent side reactions (Joondan et al. 2018). Hence, a demand exists to replace the chemical production with a biocatalytic synthesis approach.

Aminoacylases (EC 3.5.1.14) are hydrolytic metallo-enzymes that act on amide bonds of N-acyl-L-amino acids. Enzymes belonging to this diverse class have been described from several organisms. The natural function of these enzymes remains largely unknown, but for mammalian aminoacylases, functions in catabolism of acetyl-peptides or acetyl-amino acids or xenobiotic detoxification have been described (Anders and Dekant 1994). Nonetheless, aminoacylases are enzymes for technical applications, for both hydrolysis and synthesis of N-acyl-amino acids. First process implementations are found using fungal aminoacylases from *Aspergillus* sp. for the production of enantiomerically pure amino acids from racemically acetylated amino acids (Gentzen et al. 1980; Liu et al. 2006). The fungal enzyme has been evaluated for acylation of amino acids; however, yields were below 5% and thus unsatisfactory for commercial application (Kimura et al. 1992). Among all known aminoacylases, the mammalian enzyme pAcy1 from porcine kidney has been described in most detail. The hydrolytic reaction mechanism, in context with dimerization of the enzyme, has been postulated (Lindner et al. 2005; Liu et al. 2006). By directed mutagenesis, the synthesis to hydrolysis-ratio could be altered (Wardenga et al. 2010). With the addition of glycerol to the reaction mixture, yields could be improved, but product concentration remained low (<5 mM) as an excess of amino acid with a low fatty acid concentration was employed (Wada et al. 2002). The stability of the (hog) acylase is poor under favorable synthesis conditions (Henseling and Röhm 1988). In addition, the heterologous expression of soluble, affinity-tagged pAcy1 in *E. coli* has been proven difficult, as the protein is prone to aggregation (Wardenga et al. 2008).

Several bacterial aminoacylases from streptomycetal origin that are capable of N-acyl-amino acid synthesis have been published. Initially, four aminoacylases from *S. mobaraensis* have been investigated. An epsilon-lysine-acylase (SmELA) that synthesizes N<sub>ε</sub>-lauroyl-lysine (Koreishi et al. 2009a) and

an extracellular, multimeric penicillin V acylase (SmPVA; EC 3.5.1.11) capable of synthesis of several lauroyl-amino acids have been described. Also, a broad-spectrum alpha-aminoacylase (SmAA) was cloned and characterized, but no synthesis has been shown. Those enzymes were heterologously expressed in *S. lividans* (Zhang et al. 2007; Koreishi et al. 2009b). Synthetic activity towards acyl-amino acids has also been shown from cell extract of *S. ambofaciens*. The activity could be attributed to homologs of the *S. mobaraensis* enzymes, mainly to SamAA (Bourkaib et al. 2020a; Dettori et al. 2018). Furthermore, homologs of SmPVA were found in *S. lavendulae* (Torres-Bacete et al. 2015) and *Streptomyces* sp. No. 6907 (Ueda et al. 2011). None of the streptomycetal enzymes have been expressed in *E. coli* yet. The aminoacylase MsAA from *Mycolicibacterium smegmatis*, which is homologous to SamAA, has been heterologously expressed in *E. coli* and *V. natriegens* and is capable of lauroyl-amino acid synthesis (Haeger et al. 2023).

Recently, an aminoacylase from *Burkholderia* sp. was characterized that can produce several N-lauroyl-L-amino acids to high yields in aqueous buffers, reaching 51% for N-lauroyl-L-phenylalanine and even 89% for N-lauroyl-L-arginine due to product precipitation (Takakura and Asano 2019). The enzyme was remarkably stable at high temperatures and broad pH values and in the presence of inhibitors and chelators, which is an important trait for production processes. As with the streptomycetal aminoacylases, there was no need for enzyme immobilization or the use of organic solvent to achieve competitive yields. The authors found close homology of the enzyme only to several uncharacterized hydratases and low homology (≤36%) with several proteins belonging to the metal-dependent hydrolase superfamily from various microorganisms. The aminoacylase was attempted to be heterologously produced in *E. coli* BL21 (DE3); however, only traces of recombinant aminoacylase activity could be detected.

Since the enzyme from *Burkholderia* sp. possesses highly interesting traits for industrial biocatalytic production of N-lauroyl-L-amino acids but is hampered in heterologous expression, the search for homologs of the enzyme that can be produced in *E. coli* seems worthwhile. In our study, we cloned an aminoacylase gene from *Paraburkholderia monticola* DSM 100849, designated PmAcy, for heterologous expression in *E. coli* BL21 (DE3). After purification to homogeneity by affinity chromatography, the enzyme was subsequently biochemically characterized and evaluated for synthesis of acyl-amino acids.

## Materials and methods

### Chemicals and reagents

Unless stated otherwise, chemicals were analytical grade and purchased from Sigma-Aldrich (USA). Amino acids, culture

media components, NaCl, and Tris (tris(hydroxymethyl)aminomethane) were from Carl Roth (Germany), and acetyl-amino acids were either from Sigma or Bachem (Switzerland). BSA (bovine serum albumin, fraction V) was from PanReac AppliChem (Germany). *N*-Lauroyl-, *N*-palmitoyl-, and other *N*-acyl-amino acids were synthesized by the Schotten-Baumann-reaction as previously described (Haeger et al. 2023). Reagents for molecular biology were from Thermo Fisher Scientific (USA). Gene synthesis was ordered from GeneArt service by Thermo Fisher Scientific. DNA oligonucleotide synthesis and DNA sequencing were performed by Eurofins Genomics (Germany). Strep-Tactin® columns were from IBA (Germany). The EZ Nin reagent (Biochrom, UK) was used for amino acid quantification.

### Bacterial strains and plasmids

*E. coli* DH5 $\alpha$  (F<sup>-</sup>  $\phi$ 80*lacZ* $\Delta$ M15  $\Delta$ (*lacZYA-argF*)U169 *recA1 endA1 hsdR17*(r<sub>K</sub><sup>-</sup>, m<sub>K</sub><sup>+</sup>) *phoA supE44*  $\lambda^-$  *thi-1 gyrA96 relA1*) was used for cloning. *E. coli* BL21 (F<sup>-</sup> *ompT gal dcm lon hsdS<sub>B</sub>*(r<sub>B</sub><sup>-</sup>m<sub>B</sub><sup>-</sup>)  $\lambda$ (DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB*<sup>+</sup>]<sub>K-12</sub>( $\lambda$ <sup>S</sup>)) was used for protein expression. As expression plasmid, pET28a-based plasmid with recombinant aminoacylase genes was used. The resulting strains were *E. coli* BL21 (DE3) pET28a PmAcy NTag/CTag/noTag. The plasmid pGro7 (Takara Bio Inc.) was used for co-expression of chaperones. Hence, *E. coli* BL21 (DE3) pGro7 pET28a PmAcy NTag was used for recombinant aminoacylase production. The N-terminally tagged enzyme is referred to as PmAcy.

### Sequence analysis and cloning of the aminoacylase gene from *Paraburkholderia monticola*

Homology searches were performed by the program BLASTp provided by NCBI (Altschul et al. 1990). The gene accession number in the NCBI database is KXU85199.1. Multiple sequence alignment was conducted with Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) and displayed with ESPript (<https://espript.ibcp.fr/ESPript/ESPript/>). Upon adding both N- and C-terminal StrepII-Tag sequences with additional serine-glycine-linkers, the protein sequence was reverse translated based on *E. coli* codon usage. The DNA string was ordered as a GeneArt String from Thermo Fisher. The DNA string was amplified with primers carrying BsaI overhangs for golden gate cloning (Table S1). The variants of the gene were cloned, either without affinity tag (noTag, amplified with primer P2 and P4) or with an N- or C-terminal StrepII-tag (NTag or CTag, amplified with primers P1 and P4 or P2 and P3, respectively). The genes were cloned into pET28-eforRED by golden gate assembly using BsaI. Transformation of *E. coli* cells was performed with the heat shock method (Hanahan

1983). The DNA and protein sequences of PmAcy, along with the primer sequences, are included in the supplemental material.

### Production of recombinant aminoacylase in *E. coli* and purification

For protein expression *E. coli* BL21 (DE3) pGro7 pET28a PmAcy (NTag, CTag, or no tag) was grown in Terrific Broth-autoinduction medium (TB-AIM; 2% tryptone from casein, 2.4% yeast extract, 25 mM NaH<sub>2</sub>PO<sub>4</sub>, 25 mM KH<sub>2</sub>PO<sub>4</sub>, 50 mM NH<sub>4</sub>Cl, 2 mM MgSO<sub>4</sub>, 5 mM Na<sub>2</sub>SO<sub>4</sub>, 0.5% glycerol (v/v), 0.5% lactose, and 0.05% glucose). Cells were cultivated for 24 h in autoinduction media. Kanamycin was added at a concentration of 50  $\mu$ g mL<sup>-1</sup>. For co-expression of chaperones from pGro7, 25  $\mu$ g mL<sup>-1</sup> chloramphenicol and 0.5 mg mL<sup>-1</sup> arabinose were added. A 10-mL preculture was grown at 180 rpm and 37 °C overnight. In a 1-L flask, 100 mL of expression medium was inoculated with the preculture to a final OD<sub>600</sub> of 0.2 and grown at 20 °C with shaking at 225 rpm. The cells were harvested by centrifugation at 3000 g and 4 °C for 40 min.

Cells were disrupted by sonication with a sonotrode (Bandelin). The lysis buffer consisted of 100 mM Tris-HCl pH 8.0 supplemented with 0.1% Triton X-100, 1 mM ZnCl<sub>2</sub>, 0.3 mg mL<sup>-1</sup> lysozyme, and 150 mM NaCl. Per gram of harvested cells, 10 mL of lysis buffer was added to the cell pellet. Sonication was conducted for 3  $\times$  4 min at 50% power and 50% cycle. After lysis, the samples were centrifuged at 16,000 g and 4 °C for 40 min. The insoluble fraction was resuspended in 8 M urea.

Recombinant aminoacylases were purified by Strep-tag affinity chromatography using a Strep-Tactin® SuperFlow® high capacity cartridge (IBA) according to the manufacturer's instructions. The buffer used for purification was 100 mM Tris-HCl pH 8.0, 1 mM ZnCl<sub>2</sub>, and 150 mM NaCl and contained 2.5 mM desthiobiotin for elution of PmAcy. The fractions containing the recombinant enzyme were pooled and washed with elution buffer using Vivaspin 6 concentrators (10,000 MWCO; Sartorius) to remove desthiobiotin.

### Determination of protein concentration, purity, and molecular mass

Protein concentrations were determined with the Bradford method (Bradford 1976) using the Roti®-Nanoquant reagent (Carl Roth). Bovine serum albumin (BSA) served as a standard.

The SDS-PAGE analysis of proteins was conducted according to Laemmli (Laemmli 1970) using 8–20% gradient polyacrylamide gels and staining with Roti®Blue quick

(Carl Roth). As a protein ladder, FastGene® BlueEasy Protein Marker (Nippon Genetics) was used.

To determine the native molecular weight, blue native PAGE was performed using reagents from SERVA (Germany) and conducted according to the manufacturer's protocols using SERVAGel™ N4-16% gels.

Matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF/MS) analysis was conducted to determine the molecular mass of the monomeric enzyme. The Axima Confidence device (Shimadzu Europe, Duisburg, Germany) was used in linear positive mode with pulsed extraction optimized for the theoretical molecular weight. Data analysis was conducted with mMass (Strohalm et al. 2008). Concentration of the protein solutions was 1 mg/mL. The samples were diluted tenfold with  $\alpha$ -cyano-4-hydroxycinnamic acid (CHCA) (Sigma-Aldrich, USA). On each target plate spot, 2  $\mu$ L of the sample were applied. Trypsinogen and BSA from LaserBio Labs were used as molecular weight standards.

### Aminoacylase activity assay

Activity of aminoacylases was determined as previously described (Haeger et al. 2022). From 200  $\mu$ L aminoacylase reactions, 10  $\mu$ L samples were mixed with 100  $\mu$ L of EZ Nin:DMSO reagent and heated for 10 min at 99 °C. Afterwards, the samples were diluted 1:10 with 100 mM Na-borate buffer pH 10.0 for measurement in microtiter plates. The reactions consisted of 190  $\mu$ L substrate solution and 10  $\mu$ L enzyme solution. For standard hydrolytic activity measurements, reactions with 15 mM N-lauroyl-L-alanine in 100 mM Na-borate buffer pH 9.0 were performed at 30 °C for 5 min. If not stated otherwise, ZnCl<sub>2</sub> concentration in the reaction was 50  $\mu$ M. One unit of PmAcy was defined as the amount of enzyme that hydrolyzes one  $\mu$ mol of N-lauroyl-L-alanine per minute under the given conditions. For determination of kinetic parameters, volumes of the ninhydrin assay were adjusted to increase sensitivity at low substrate concentrations, so that 50  $\mu$ L sample was mixed with 200  $\mu$ L EZ Nin:DMSO. Then, 200  $\mu$ L of the colored sample was mixed with 50  $\mu$ L 100 mM Na-borate pH 10.0.

### Biochemical characterization of PmAcy

The optimal pH of the hydrolytic reaction was investigated by using following buffers: Na-citrate for pH 5.0, Na-MES (2-(N-morpholino)ethanesulfonic acid) for pH 6.0 and 7.0, Tris-HCl for pH 6.0–9.0, and Na-borate for pH 9.0–12.5. Due to poor solubility at low pH values, only 3 mM N-lauroyl-L-alanine was prepared in 50 mM of buffer and adjusted to the respective pH. The purified enzyme solution was diluted 1:10 in each buffer to a final protein concentration of 2  $\mu$ g/mL. The reactions were conducted for 10 min at 30 °C

with sampling each minute and analysis of initial reaction velocity.

The pH stability was assessed in 100 mM of Na-acetate buffer at pH 4.0–6.0, Na-MES buffer at pH 5.0–7.0, Tris-HCl at pH 6.0–9.0, and Na-borate buffer at pH 9.0–13.0 at 30 °C without agitation. From the incubation solutions containing 240  $\mu$ g/mL enzyme, 10  $\mu$ L was withdrawn and used for activity measurements after 1 h and 24 h. As a substrate, 15 mM N-lauroyl-L-alanine was used in 200 mM Na-borate pH 9.0. The borate buffer concentration in the assay was altered compared to standard conditions to prevent a significant pH shift caused by the incubation buffers upon addition of the incubated enzyme. The reaction was conducted at 30 °C. Final enzyme and ZnCl<sub>2</sub> concentrations were 12  $\mu$ g/mL and 10  $\mu$ M, respectively.

The optimal temperature for hydrolysis was determined by measuring the activity in a range of 20–90 °C. Reactions were conducted with 15 mM N-lauroyl-L-alanine in 100 mM Na-borate pH 9.0 and 1–4  $\mu$ g/mL purified enzyme.

For the assessment of temperature stability or thermal activation, purified PmAcy was incubated at temperatures from 20 to 90 °C in 100 mM Tris-HCl buffer at pH 8 and residual activity was determined after 1 h, 24 h, and 4 days. Reactions were carried out at a PmAcy concentration of 60  $\mu$ g/mL in a volume of 200  $\mu$ L. Residual activity was determined with the standard aminoacylase assay (15 mM N-lauroyl-L-alanine in 100 mM Na-borate buffer pH 9.0 at 30 °C).

### Effect of metal ions and substrate specificity

The effect of various bivalent metal ions on the hydrolytic activity of PmAcy was investigated by incubating 1 mg/mL of the purified enzyme (no ZnCl<sub>2</sub> added to the buffer) with various ions at a final concentration of 0.1 nM to 1 mM (CaCl<sub>2</sub>, CoCl<sub>2</sub>, CuCl<sub>2</sub>, FeSO<sub>4</sub>, MgCl<sub>2</sub>, MnCl<sub>2</sub>, NiCl<sub>2</sub>, and ZnCl<sub>2</sub>) for 1 h at room temperature. After the incubation period, activity was measured with the standard assay. In the absence of divalent metal ions, no activity could be measured. The activity determined by addition of zinc was defined as 100%.

The substrate scope was determined by hydrolysis of various substrates at 15 mM in 100 mM Na-borate, pH 9.0 at 50 °C. The reaction temperature was higher than the standard conditions to solubilize all substrates, because N-palmitoyl-L-amino acids and N-lauroyl-L-cysteine were not soluble at 30 °C. The concentrations of purified enzyme were 6–60  $\mu$ g/mL. The final concentration of bivalent metals (ZnCl<sub>2</sub>, CoCl<sub>2</sub>, or MnCl<sub>2</sub>) in the reactions was 50  $\mu$ M. The activity was determined from the initial reaction velocity.

To determine  $K_M$  and  $V_{max}$  of the hydrolysis reaction towards N-lauroyl-L-alanine and N-lauroyl-L-phenylalanine as substrates, a concentration range of 50  $\mu$ M to 15 mM of

the substrates was prepared in 100 mM Na-borate pH 9.0 at 50 °C. The enzyme concentration was 1.2 µg/mL, and initial reaction velocities were analyzed in sampling intervals of 30 s. The kinetic parameters were determined with Graph-Pad Prism 8 software.

### Thermal shift assay

Thermal shift assay was used as previously described (Falkenberg et al. 2022) to investigate thermal denaturation of PmAcy in real time. After mixing the protein with the fluorogenic dye SYPRO® orange, the sample was heated stepwise. An increase in fluorescence during the heating procedure indicates a denaturation of the protein due to exposed hydrophobic areas. 10 µL protein samples (> 0.1 mg/mL) are mixed with 5 µL 50× SYPRO Orange (Sigma Aldrich) and 20 µL 10 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) pH 8.0. 10 mg/mL lysozyme (Serva) served as a positive control. Measurement was performed with qTower3G and qPCRsoft 4.0 (Analytik Jena) using the TAMRA Channel ( $\lambda_{ex}$  = 535 nm,  $\lambda_{em}$  = 580 nm). The heating program was 25 °C to 95 °C with steps of 2 °C, 120 s hold time per temperature and a heating speed of 4.4 °C/s.

### Biocatalytic synthesis of N-acyl-amino acids

In general, the reaction mixtures consisted of 200 mM amino acid, 100 mM fatty acid, 500 µM ZnCl<sub>2</sub> in 50 mM buffer Tris–HCl buffer pH 8.0 in a volume of 1 mL. The reactions were started by addition of 0.6 Units aminoacylase PmAcy (standard activity assay), and the syntheses were run at 50 °C and 250 rpm for 24 h in 25 mL vessels on an *Infors TM* orbital shaker. Fatty acids, amino acids, and pH were varied in a “one factor at a time” (OFAT) approach. The pH was varied from 7.0 to 13.0 using either 50 mM Tris–HCl buffer (pH 6–9) or 50 mM borate buffer (pH 9.0–13.0). Different fatty acids from C8 to C18 were analyzed with L-phenylalanine as amino acid cosubstrate. For carbon chain length from C16 to C18, 10% (v/v) ethanol was added to the reaction mixture for solubilization purposes. All 20 natural amino acids were tested in the presence of lauric acid as cosubstrate. The reactions were stopped by addition of 990 µL acetonitrile and incubation at 100 °C for 30 min and analyzed by HPLC-ELSD and LC–MS. N-lauroyl-L-phenylalanine was synthesized in 5-mL scale and purified exemplarily according to the method described in the “Scale-up and isolation of N-lauroyl-L-phenylalanine” section.

### Scale-up and isolation of N-lauroyl-L-phenylalanine

200 mM L-phenylalanine, 100 mM sodium laurate, 0.5 mM ZnCl<sub>2</sub>, and 3.0 Units PmAcy (standard activity assay; 60 µg)

were dissolved in 50 mM Tris-buffer pH 8.0 at a total volume of 5 mL. The synthesis reactions were conducted at 50 °C and 250 rpm for 1 h. Afterwards, the pH-value of the reaction mixture was set to 1 by addition of 5 N hydrochloric acid. The precipitate was separated from the liquid phase by centrifugation, dried *in vacuo*, and analyzed.

### Chromatographic and spectroscopic product analysis

The conversion of amino acids to N-lauroyl-L-amino acids was monitored by HPLC-ELSD with a Shimadzu Nexera XR system equipped with a Hitachi LaChrom II column (C18, 5 µm, 4.6×250 mm) and an ELSD 100 (evaporative light scattering detector) from VWR. The products were dissolved in an equal volumetric mixture of acetonitrile and water, and 20 µL was injected to the column. The column temperature was set to 40 °C, and the products were separated in a gradient starting from starting from 80% water (supplemented with 0.1% formic acid):20% acetonitrile to 100% acetonitrile within 10 min. The concentration was held for 6 min, and the gradient was set back to 20% acetonitrile over the course of 2 min. The flow rate was set to 1 mL/min. Amino acids and acyl-amino acid concentrations were calibrated with 1 mM ethyl-L-tryptophanate hydrochloride (TrpOEt-HCl) solution as internal standard.

LC–MS analysis was conducted using a Shimadzu Nexera XR system equipped with a Hitachi LaChrom II column (C18, 5 µm, 4.6×250 mm) and a Shimadzu LCMS-2020-mass spectrometer. Applying the same analysis conditions as in HPLC-ELSD analysis allowed a direct comparison of the chromatograms.

NMR spectroscopy of N-lauroyl-L-phenylalanine was done with a Bruker Avance III HD 400 MHz system. 30 mg N-lauroyl-L-phenylalanine was dissolved in 650 µL CDCl<sub>3</sub>. Proton nuclear magnetic resonance (<sup>1</sup>H-NMR) spectra were analyzed at 400 MHz and chemical shifts are reported in δ units (ppm) with TMS as reference (δ 0.00). Carbon nuclear magnetic resonance (<sup>13</sup>C-NMR) spectra were measured at 100 100 MHz, and chemical shifts are reported in δ units (ppm) with CDCl<sub>3</sub> as reference δ (77.00).

## Results

### Cloning of the *pmAcy* gene from *Paraburkholderia monticola* and expression in *E. coli*

The putative aminoacylase from *P. monticola* DSM 100849 is a homolog of the long-chain aminoacylase from *Burkholderia* sp. Strain LP5\_18B (Takakura and Asano 2019) with 85.5% homology by sequence identity as identified via NCBI BLASTp algorithm (GenBank: KXU85199.1).

The synthetic DNA fragment (GenBank OR188138) was cloned into the pET28a expression vector via Golden Gate cloning. The DNA string contained both N- and C-terminal StrepII-tag sequences. The gene was cloned either with N- or C-terminal Strep-tag or without an affinity tag. The resulting plasmids were designed as pET28a PmAcy (Ntag/Ctag). The theoretical molecular weight of the protein is 47.4 kDa (48.7 kDa with Strep-tag).

Sequence analysis revealed that PmAcy belongs to the conserved protein domain family of metallo-dependent hydrolases (subgroup A). This group of proteins shows a conserved TIM barrel 3D structure (Lu et al. 2020) and can be classified a member of the M38 family from the MJ clan in MEROPS peptidase database classification. The characteristic metal-binding site comprising four histidine residues (H85/87/253/273) and one aspartic acid residue (D340) is conserved in the sequence of PmAcy. In a homologous amidohydrolase structure (PDB-ID: 3MKV; 33% identity), the zinc ions are bridged by a carboxylated lysine residue (Xiang et al. 2010), and the corresponding lysine residue is conserved in the sequence of PmAcy as well (K212). Other conserved residues include H162, which is described as an oxyanion-hole forming residue to stabilize the reaction intermediate, and Y255, which serves to bind the  $\alpha$ -carboxylic group of the substrate (Xiang et al. 2010). A protein sequence alignment of PmAcy with the sequences of the *Burkholderia* aminoacylase (Takakura and Asano 2019) and the referenced amidohydrolase (3MKV) with highlighted active site residues is shown in (Figs. 1 and 2).

The aminoacylase PmAcy was prone to aggregation when heterologously produced in *E. coli* BL21 (DE3). The expression experiments using 1 mM IPTG for induction at 37 °C and 20 °C yielded only inclusion bodies. Lowering the inductor concentration to 0.1 mM IPTG at 20 °C did not yield soluble protein either. Furthermore, switching to autoinduction with lactose did not help to increase the solubility. By co-expression of the chaperonin GroEL/ES in *E. coli* BL21 (DE3) pGro7 at 20 °C using induction with lactose and arabinose, soluble and active protein could be obtained. Still, the majority of PmAcy was found in the insoluble fraction. In comparison, the N-terminally tagged variant of PmAcy was produced best (data not shown) and used for further purification and enzyme characterization.

### Purification of the N-terminally StrepII-tagged aminoacylase PmAcy

The enzyme was typically enriched 16-fold, and 93.8% of the enzyme was recovered with 2288 U, and 3.0 U/mg in the cell-free extract and 2146 U in the elution fractions. The specific hydrolytic activity of purified PmAcy Ntag was determined as 50 U/mg with the substrate N-lauroyl-L-alanine using a standard activity assay (30 °C reaction temperature,

pH 9.0). From 100 mL expression culture, approximately 15 mg purified PmAcy were yielded. As Fig. 2 shows, the enzyme was purified to homogeneity with an apparent molecular mass of approximately 50 kDa. The mass of purified PmAcy NTag was verified by MALDI-TOF analysis to be 48.7 kDa, as expected from its similar theoretical molecular weight. Native PAGE revealed that the aminoacylase is a multimeric enzyme which migrates at a position of  $M_r$  550 kDa, suggesting a dodecameric form (Figs S1, S2).

### Biochemical characterization of PmAcy

The hydrolysis of N-lauroyl-L-alanine at 30 °C was used as a standard reaction for the biochemical characterization of purified PmAcy. To determine the pH optimum, hydrolytic activities were analyzed in various buffers ranging from pH 5–12.5 (Fig. 3A). Experiments at lower pH-values were not possible, because the substrate was insoluble. The highest activity of 278 U/mg was obtained at pH 12, which was set to 100% relative activity (Fig. 3A). Only weak activity was detected in the neutral to acidic pH range. The pH stability of PmAcy was assessed by incubating the enzyme at 30 °C for 1 h and 24 h in buffers ranging from pH 4–13 (Fig. 3B). After incubation, the remaining hydrolytic activity was measured by the standard activity assay (30 °C, pH 9.0). The enzyme was stable in all buffers for 1 h, and still > 80% activity remained at pH 4–12 after 24 h of incubation. Only at pH 13 significant deactivation was observed. Still, the enzyme is sufficiently stable at alkaline conditions, and 92% of its initial activity were retained after 24 h at the optimum pH of 12.

The temperature dependency of the hydrolysis reaction rate was assayed in a range from 20 to 90 °C. The maximum activity was reached at 75 °C with 773 U/mg at pH 9.0 (Fig. 4A). To assess the thermal stability at these high temperatures, a thermal shift assay was performed. The enzyme sample was mixed with the fluorogenic dye Sypro orange. The fluorescence increases upon denaturation of the protein while allowing the dye to bind to newly exposed hydrophobic patches. Hence, heat-induced denaturation of the investigated enzyme can be observed in real time. Thermal denaturation starts at approximately 78 °C, and the melting point of PmAcy is 89 °C. Before the fluorescence signal increases due to denaturation of the enzyme, the signal decreases, starting from an extraordinary high level (Fig. 4B). This might be due to exposed hydrophobic patches of the recombinant enzyme and folding during heating (Boivin et al. 2013).

Motivated by these results, we investigated whether heating of the enzyme may increase its activity permanently (Fig. 4C). After 1-h incubation at various temperatures and thorough cooling of the enzyme samples, hydrolytic activity at 30 °C was measured. Indeed, an enzyme activation was

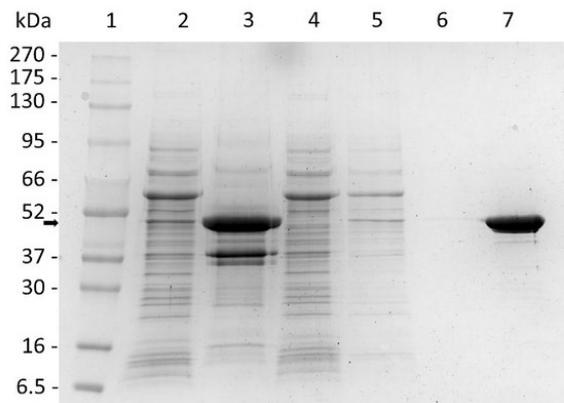


**Fig. 1** Protein sequence alignment. PmAcy (KXU85199.1), the aminoacylase from *Burkholderia* sp. strain LP5\_18B, and an amidohydrolase (PDB: 3MKV) were aligned. The alignment was generated using the Clustal Omega tool and displayed with ESPrnt 3.0. The conserved metal-binding residues (four histidines, one aspartic acid, and one lysine residue) are highlighted by green boxes and asterisks.

The conserved histidine described to be oxyanion-hole forming is highlighted in pink. The conserved tyrosine residue that binds the  $\alpha$ -carboxylic group of amino acid substrates in amidohydrolases is highlighted in orange. The annotation of the secondary structure, based on the structure of 3MKV, is displayed as arrows for  $\beta$ -strands, squiggles for  $\alpha$ -helices, and the letters TT for turns

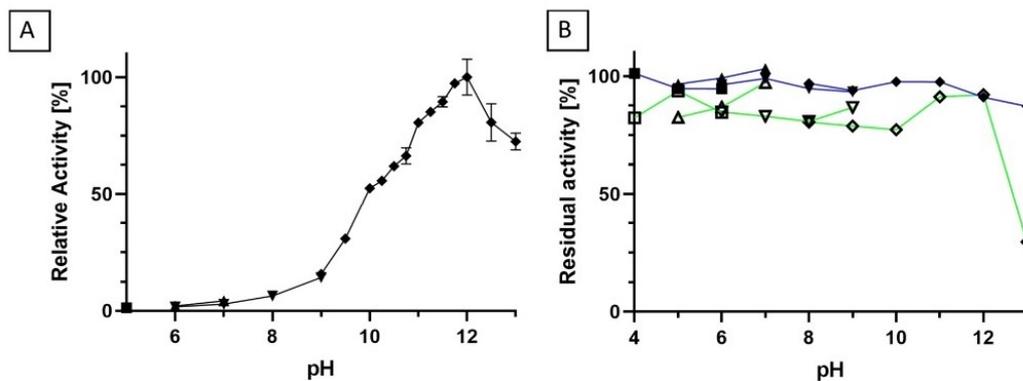
observed upon heating. The strongest effect was observed at 80 °C with a threefold increase of activity. After 24-h incubation at the respective temperatures, a similar behavior was

observed, although counteracted by thermal instability. To assess the stability of the heat-activated PmAcy, the enzyme was first heat-activated at 80 °C for 1 h, cooled down and



**Fig. 2** SDS-PAGE analysis of different fractions obtained during purification of PmAcy. The enzyme was purified from a culture of *E. coli* BL21 pGro7+pET28a PmAcy (TB medium-autoinduction). Lane 1: protein marker (BlueEasy Prestained Protein Marker, Nippon Genetics); lane 2: cell-free extract; lane 3: insoluble fraction; lane 4: flow-through from Strep-tag affinity chromatography; lane 5: first wash fraction; lane 6: second wash fraction; lane 7: Elution of PmAcy. The black arrow indicates the migration height of PmAcy

then further incubated at various temperatures. After 1 h, 24 h and 4 days, the residual hydrolytic activities were measured (Fig. 4D). The enzyme was remarkably heat-stable. After 1 h at 80 °C, no significant loss in activity was found. At 70 °C, residual activities were 83% and 38% after 24 h and 4 days, respectively. At 50 °C, even 81% of activity was retained after 4 days. Additionally, stability after freezing the purified enzyme at -20° or -80 °C and subsequent thawing



**Fig. 3** Influence of pH on activity and stability of PmAcy. **A** pH dependency of hydrolytic activity of PmAcy against N-lauroyl-L-alanine. Reaction conditions: 3 mM N-lauroyl-L-alanine at 30 °C reaction temperature. The following buffers were used at 50 mM: Na-citrate for pH 5.0 (●), Na-MES for pH 6.0 and 7.0 (▲), Tris-HCl for pH 6.0–9.0 (▼), and Na-borate for pH 9.0–13.0 (◆). All reactions were conducted in triplicates. **B** pH dependency of stability of PmAcy after 1 h and 24 h. The following buffers were used at

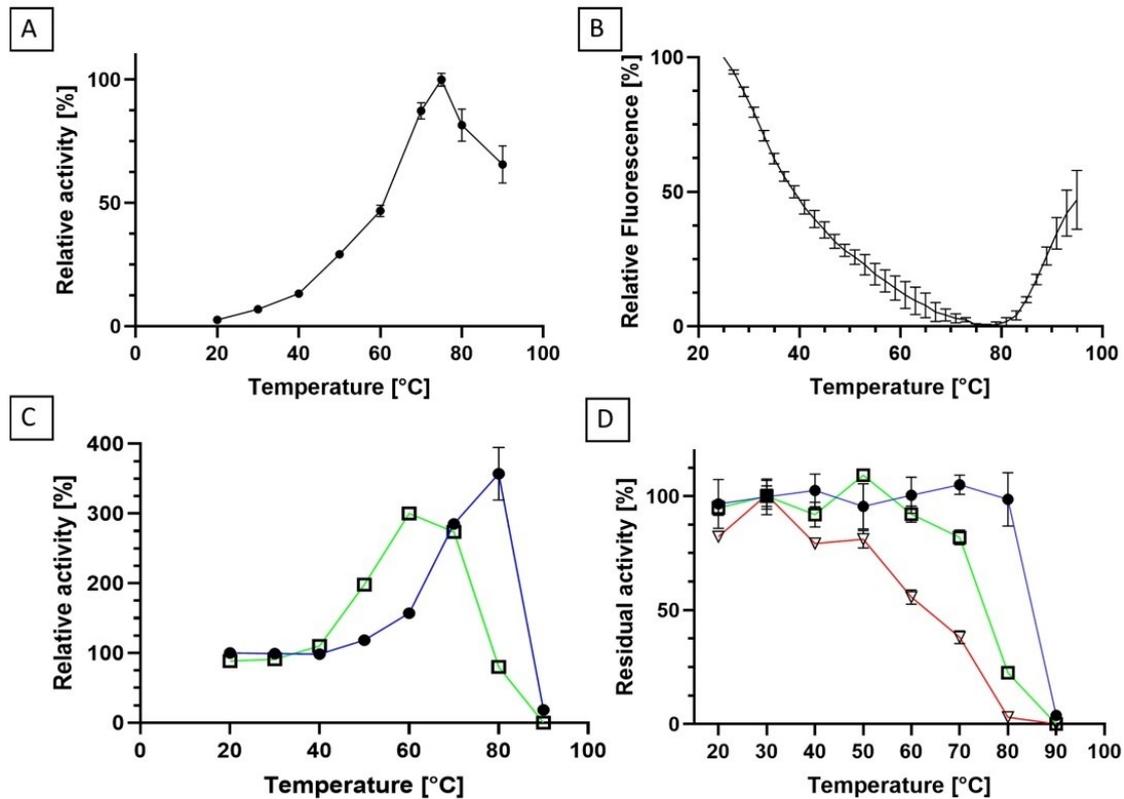
was investigated with and without the addition of 20% and 50% glycerol. No loss of activity was detected compared to the control kept at 4 °C.

### Substrate specificity of PmAcy in hydrolysis reactions and effect of metal ions

After purification of the enzyme without any bivalent metal ions added, no activity was retained. We found that 50 μM CaCl<sub>2</sub>, CuCl<sub>2</sub>, FeSO<sub>4</sub>, MgCl<sub>2</sub>, and NiCl<sub>2</sub> could not restore activity, whereas the addition of ZnCl<sub>2</sub>, CoCl<sub>2</sub>, and MnCl<sub>2</sub> reactivated PmAcy. The hydrolytic activities remained at the same level, when the metal ions were used at 100 μM, but were approximately halved when the enzyme was incubated with 10 μM metal ions. At concentrations of 1 μM hardly any activity was detected. Thus, the substrate specificity of PmAcy was investigated with various acylated L-amino acids in the presence of 50 μM ZnCl<sub>2</sub>, CoCl<sub>2</sub>, or MnCl<sub>2</sub> (Table 1). Reactions were conducted at 15 mM substrate concentration in 100 mM Na-borate pH 9.0 at 50 °C to ensure solubilization of all substrates. L-amino acid released by PmAcy hydrolysis was quantified with the ninhydrin assay.

The enzyme has a bias for lauric acid and non-polar and neutral L-amino acids (Table 1). With zinc, favored substrates were N-lauroyl-L-alanine, N-lauroyl-L-isoleucine, N-lauroyl-L-methionine, and N-lauroyl-L-valine at the chosen conditions. Regarding the fatty acid moiety, N-palmitoyl-L-alanine was hydrolyzed at a rate of 36% compared to N-lauroyl-L-alanine. On the other hand, N-acetyl-L-alanine and various other acetylated amino acids were

100 mM: Na-acetate for pH 4.0–6.0 (■/□), Na-MES for pH 5.0–7.0 (▲/△), Tris-HCl for pH 6.0–9.0 (▼/▽), and Na-borate for pH 9.0–13.0 (◆/◇). Activities measured after 1 h and 24 h are depicted with filled and empty symbols, connected by blue and green lines, respectively. Residual activity was determined with the standard assay conditions and indicated as percental values. All reactions were conducted in triplicate



**Fig. 4** Influence of temperature on activity and stability. **A** Temperature dependency on hydrolytic activity against N-lauroyl-L-alanine. Reaction conditions were 15 mM N-lauroyl-L-alanine in 100 mM Na-borate pH 9.0 at various reaction temperatures. **B** Thermal shift assay with PmAcY. Thermal denaturation was followed via fluorescence measurement of SYPRO Orange. **C** Thermal activation of PmAcY

after 1-h (●, blue line) and 24-h incubation (□, green line) at various incubation temperatures and assayed according to the standard protocol at 30 °C. **D** Thermal stability of PmAcY after 1 h (●, blue line), 24 h (□, green line), and 4 days (▽, red line) incubation at various temperatures in 100 mM Tris-HCl pH 8.0, 150 mM NaCl, and 1 mM ZnCl<sub>2</sub>. All reactions were conducted in triplicate

barely hydrolyzed. Similarly, the activity against N-palmitoyl-L-glutamine and N-capryloyl-L-glutamine was 50% and 13% of N-lauroyl-L-glutamine, respectively, but no activity against N-acetyl-L-glutamine was measured. The aminoacylase also showed no dipeptidase activity against L-Ala-L-Phe and L-Phe-L-Ala. When the enzyme was incubated with 1 mM of bivalent cobalt ions, the substrate scope was altered. However, no additional substrates were accepted. The specific activity against N-lauroyl-L-alanine was altered from 184.4 U/mg with ZnCl<sub>2</sub> to 239.0 U/mg with CoCl<sub>2</sub> and to 63.9 U/mg with MnCl<sub>2</sub>. For N-lauroyl-L-phenylalanine, the differences were smaller (78.0 U/mg, 80.6 U/mg, and 57.4 U/mg, respectively). Remarkably, activity against N-palmitoyl-L-alanine and N-palmitoyl-L-glutamine was increased when cobalt instead of zinc was added to the enzyme (from 60.4 U/mg to 125.3 U/mg and from 38.8 U/mg to 85.6 U/mg, respectively). As the

cobalt-activated enzyme also hydrolyzed N-capryloyl-L-glutamine faster, the specificity for the acyl moiety seems to be loosened. This trend was more pronounced when the enzyme was incubated with manganese, observing almost the same reaction velocity for N-lauroyl- and N-palmitoyl-L-alanine. However, since the activity was generally lower, the activity with manganese never exceeded the activity with zinc or cobalt for any substrate, except for N-lauroyl-L-leucine.

Kinetic values for the hydrolysis of N-lauroyl-L-alanine and N-lauroyl-L-phenylalanine at 30 °C were determined by measuring initial reaction rates at varying concentrations from 50 μM to 15 mM. The  $K_M$  and  $V_{max}$  values against N-lauroyl-L-alanine were found to be 1.95 mM and 55.4 U/mg, respectively. In contrast, for N-lauroyl-L-phenylalanine, the  $K_M$  value was 0.21 mM, and  $V_{max}$  was 25.1 U/mg (Fig. S3).

**Table 1** Substrate specificity of PmAcy. Hydrolytic activity was measured at 15 mM substrate in 100 mM Na-borate pH 9.0. The reaction was conducted at 50 °C. Released amino acid were measured with the ninhydrin assay and activity was determined from the initial reaction velocity. The enzyme and  $\text{Me}^{2+}$  concentrations were 6–60  $\mu\text{g}/\text{mL}$  and 50  $\mu\text{M}$ , respectively. All reactions were conducted in triplicates

Substrate	Specific activity (U/mg)		
	ZnCl <sub>2</sub>	CoCl <sub>2</sub>	MnCl <sub>2</sub>
N-Lauroyl-L-alanine	184.4 ± 1.7	238.9 ± 4.6	63.9 ± 1.3
N-Acetyl-L-alanine	< 1	n.d	n.d
N-Benzoyl-L-alanine	0	n.d	n.d
N-Palmitoyl-L-alanine	60.4 ± 4.8	125.3 ± 7.9	58.6 ± 7.6
L-Phenylalanyl-L-alanine	0	n.d	n.d
N-Lauroyl-L-aspartic acid	0	0	0
N-Lauroyl-L-cysteine	0	0	0
N-Lauroyl-L-glutamic acid	0	0	0
N <sub>α</sub> -Lauroyl-L-glutamine	69.1 ± 3.3	105.8 ± 13.5	16.6 ± 1.2
N <sub>α</sub> -Capryloyl-L-glutamine	10.5 ± 0.4	22.8 ± 0.6	4.3 ± 0.3
N <sub>α</sub> -Palmitoyl-L-glutamine	38.8 ± 4.2	85.6 ± 1.1	28.2 ± 0.7
N-Lauroyl-glycine	123.9 ± 2.1	166.2 ± 5.4	23.6 ± 2.3
N-Lauroyl-L-isoleucine	148.9 ± 4.1	123.1 ± 1.6	39.5 ± 2.0
N-Lauroyl-L-leucine	32.3 ± 3.4	43.4 ± 3.4	47.0 ± 3.2
N-Lauroyl-L-methionine	177.8 ± 13.6	121.5 ± 1.7	55.3 ± 4.1
N-Lauroyl-L-phenylalanine	78.0 ± 3.7	80.6 ± 1.4	57.4 ± 4.1
L-Alanyl-L-phenylalanine	0	n.d	n.d
N-Lauroyl-L-serine	20.4 ± 4.0	31.7 ± 2.1	17.9 ± 2.2
N <sub>α</sub> -Lauroyl-L-tryptophan	4.9 ± 1.3	17.6 ± 0.1	2.1 ± 0.4
N-Lauroyl-L-tyrosine	63.5 ± 6.5	10.1 ± 0.9	30.3 ± 1.2
N-Lauroyl-L-valine	130.6 ± 2.3	168.7 ± 13.6	40.9 ± 1.8

### Biocatalytic N-acyl-L-amino acid synthesis in aqueous media

The synthesis of N-acyl-L-amino acids was initially tested with L-alanine and lauric acid as substrates, since N-lauroyl-L-alanine proved to be the best substrate in the hydrolysis reaction. High substrate concentrations with 100 mM lauric acid and a two-fold stoichiometric excess of L-alanine (200 mM) were utilized. Despite being a favored substrate in hydrolysis, only trace amounts of N-lauroyl-L-alanine were detected in HPLC-ELSD analysis, and thus, synthesis reactions were repeated with L-phenylalanine under the same conditions. Formation of the acylated product was detected in HPLC-ELSD, which corresponded to the retention time of the chemically synthesized N-lauroyl-L-phenylalanine reference substance (Fig. 5A; HPLC ELSD chromatogram overlay). To verify the structure of the new product, it was enriched by acid precipitation according to the standard procedure employed in Schotten-Baumann synthesis and extracted (Fig. 5A). LC-MS analyses (Fig. S4) and NMR

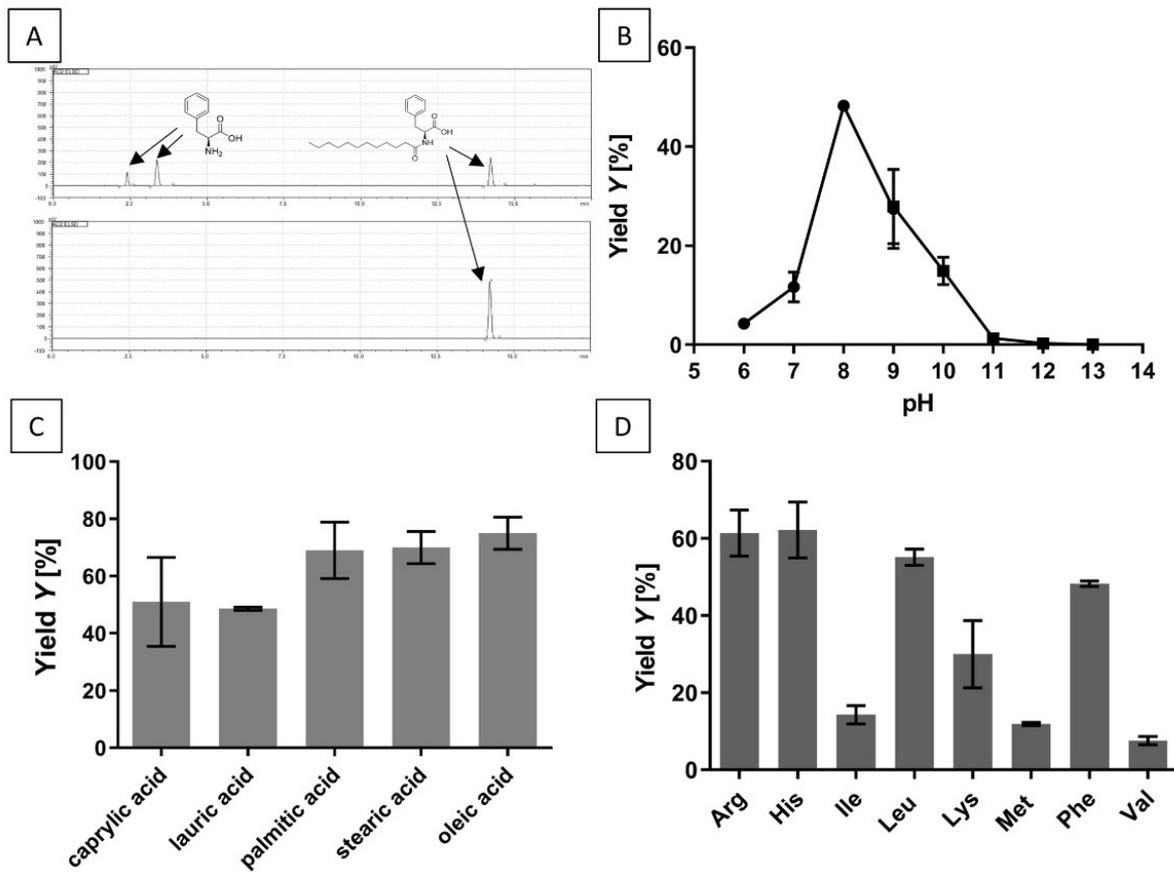
spectroscopy (Fig. S5, S6) unambiguously confirmed the formation of N-lauroyl-L-phenylalanine.

The synthesis potential of the enzyme was then tested in the pH range from 6.0 – 13.0 (Fig. 5B). The highest conversion for synthesis was observed at pH 8.0, whereas the highest initial activity of the enzyme was at pH 12.0 for the hydrolysis reaction. At pH 7.0 and below, the substrates were not completely soluble initially forming a cloudy mixture, whereas raising the pH above 10.0 seems to drastically lower the yield of the reaction. A pH of 8.0 was chosen for all further synthesis experiments. The chain length specificity of fatty acids was tested in the range of C8–C18 with L-phenylalanine as co-substrate (Fig. 5C). For all fatty acids of chain lengths C16 and C18, 10% (v/v) ethanol had to be added to facilitate sufficient solubilization of the hydrophobic substrate in the aqueous reaction system. The synthesis of medium- and long-chain N-acyl-L-phenylalanine showed good conversions between 50 and 75%. Higher conversions of 70–75% were obtained with the C16–C18 fatty acid substrates, while slightly lower conversions of 50–55% were obtained with C8 and C12 fatty acid substrates. The lowest transformation rate was observed with short chain caprylic acid and lauric acid.

Following these results, the conversion of all 20 natural L-amino acids was tested using lauric acid as co-substrate. As Fig. 5D shows, L-histidine, L-arginine, L-leucine, and L-phenylalanine showed the highest conversion rates with 62%, 61%, 55%, and 48% yield after 24 h reaction at 50 °C, respectively, while slightly lower yields between 7 and 30% yield were obtained with L-isoleucine, L-lysine, L-methionine, and L-valine. In the condensation reaction, the enzyme seems to favor highly hydrophobic and basic amino acid side chains, while other polar and weakly non-polar L-amino acids (L-alanine, L-asparagine, L-serine, L-threonine, L-tryptophan, and L-tyrosine) produced only traces of product (not shown). Acidic L-amino acids (L-aspartic acid, L-glutamic acid) as well as L-glutamine, L-cysteine, glycine, and L-proline were not accepted as substrates under the reaction conditions applied.

### Discussion

We described the cloning, expression, characterization, and the evaluation of the biocatalytic potential of the novel aminoacylase PmAcy. The successful expression in recombinant *E. coli* cells described in this study has not been achieved so far; several other aminoacylases were produced in the native host, or heterologous expression was not successful (Takakura and Asano 2019; Bourkaib et al. 2020a). The expression of a mycobacterial aminoacylase could be enhanced through the co-expression of GroEL/S in *E. coli* and *V. natriegens* (Haeger et al. 2023). We developed



**Fig. 5** The synthetic potential of PmAcy for the formation of N-acyl-amino acids. Substrate concentrations were 200 mM amino acid and 100 mM fatty acid. **A** HPLC-ELSD chromatogram of the product N-lauroyl-L-phenylalanine synthesized with PmAcy after 24 h (upper picture) and after purification via acid precipitation (lower picture). **B** pH dependency of N-lauroyl-L-phenylalanine synthesis in 50 mM Tris buffer (●, pH 6.0–9.0) and 50 mM borate buffer (■, pH 10.0–13.0) at 50 °C for 24 h; all reactions were made in tripli-

cates; **C** Synthesis of N-acyl-L-phenylalanine with different fatty acids (C8:0=caprylic acid, C12:0=lauric acid, C16:0=palmitic acid, C18:0=stearic acid, C18:1=oleic acid); all transformations were performed in duplicates. **D** Synthesis of N-lauroyl amino acids with different amino acids (given as 3-letter code); all transformations were done in triplicates; fatty acid and amino acid variations were analyzed at 50 °C in 50 mM Tris buffer pH 8.0

a protocol that enables soluble expression of the recombinant enzyme involving autoinduction with lactose and co-expression of the GroEL/S chaperonine.

The enzyme's stability at alkaline pH values and high temperatures is extraordinary for aminoacylases and has only been described for the homologous aminoacylase from *Burkholderia* sp. (Takakura and Asano 2019). This distinctive trait is especially interesting for use in enzymatic synthesis of N-acyl-L-amino acids, because the disadvantage of other known aminoacylases is often their insufficient stability (Wardenga et al. 2010; Bourkaib et al. 2020b). In this context, we observed an enhanced activity of the purified enzyme after incubation at high temperatures. This permanent heat-activation has not yet been described in literature.

As PmAcy is prone to the formation of multimers, this may hint at a temperature-dependent disintegration of enzyme aggregates.

The aminoacylase was active with the addition of  $Zn^{2+}$ ,  $Co^{2+}$ , and  $Mn^{2+}$ . The dependency of activity on divalent cations has been observed for other aminoacylases as well. The homologous aminoacylase from *Burkholderia* sp. strain LP5\_18B was described to be zinc-dependent (Takakura and Asano 2019). Furthermore, several other aminoacylases were found to be activated by divalent cations, namely, zinc, cobalt, or nickel (Bourkaib et al. 2020a; Koreishi et al. 2009b; Natsch et al. 2003). Regarding hydrolysis of various N-acyl-L-amino acids, PmAcy preferred rather small, hydrophobic L-amino acids among the tested substrates.

The best substrates were N-lauroyl-L-alanine, N-lauroyl-L-isoleucine, N-lauroyl-L-methionine and N-lauroyl-L-valine. Some polar, uncharged acylated amino acids, like N-lauroyl-L-serine, and N-acyl-L-glutamines were accepted to a lesser extent. However, N-lauroyl-L-cysteine and the charged N-lauroyl-L-glutamic acid and N-lauroyl-L-aspartic acid were not accepted at all. Hydrolysis of acylated L-histidine, L-arginine, and L-lysine could not be investigated during this study due to missing reference products or insolubility of the substrates.

The optimal pH for the enzymatic production of N-lauroyl-L-phenylalanine was pH 8.0. This pH value is much lower than the optimal pH for hydrolysis. In chemical synthesis with acyl chlorides, an alkaline pH is favorable for the synthesis of N-acyl-L-amino acids. With free fatty acids, which deprotonate in neutral or alkaline solutions, the optimal pH for synthesis can however be different in biocatalysis. The optimal pH for the reaction is also determined by the micro-environment of the enzyme's active site and could therefore be unique for each enzyme. For the close homolog from *Burkholderia* sp. LP5\_18B, the optimal pH for synthesis of N-lauroyl-L-phenylalanine was pH 9.0, despite having its hydrolytic optimum at pH 12.0. For SamAA, an aminoacylase from *S. ambofaciens*, slightly basic pH of 8.0 was best for maximal yield and pH 8.5 for highest initial velocity of the synthesis of N-undecylenoyl-L-phenylalanine (Bourkaib et al. 2020b). The optimal pH of porcine pAcy1 was around 7.5 in the synthesis of N-lauroyl-L-lysine, N-lauroyl-L-methionine, and N-lauroyl-L-glutamic acid (Wada et al. 2002).

Comparing various aminoacylases from literature, the activity against fatty acids of varying chain length differs. Often, there is a bias towards short-chain acyl residues, e.g., acetic acid, with promiscuous, lower activity to longer acyl chains (Koreishi et al. 2009b, 2005; Haeger et al. 2023). However, in case of the *Burkholderia* aminoacylase and penicillin V acylases, N-lauroyl-L-amino acids are hydrolyzed with high activity, whereas N-acetyl-L-amino acids are not accepted as substrates. The aminoacylase from *P. monticola* described in this study behaves like the latter case, barely hydrolyzing N-acetyl-L-amino acids and favoring N-lauroyl-L-amino acids in hydrolysis. In synthesis, PmAcy even favors long-chain stearic and oleic acids, thus rendering it a long-acyl chain aminoacylase both in hydrolysis and in synthesis. A discrepancy in hydrolytic and synthetic substrate scope is noticeable, most strikingly as N-lauroyl-L-alanine is the favored substrate in hydrolysis but is barely synthesized. A similar behavior has been described for the *Burkholderia* enzyme (Takakura and Asano 2019). Furthermore, PmAcy hydrolyzed N-lauroyl-L-isoleucine faster than N-lauroyl-L-leucine; however, synthesized N-lauroyl-L-leucine is more efficiently. Hence, biocatalytic substrate scope might not only be dependent on substrate binding to the enzyme's active site but also by thermodynamic equilibrium and product solubility.

In summary, we identified a novel aminoacylase from *P. monticola* that is capable of long-chain N-acyl-L-amino acid synthesis. Synthesis of N-acyl-L-amino acids was possible in aqueous conditions in high yield, which is ecologically advantageous to the industrially applied Schotten-Baumann reaction and eventually more economical. Thus, we anticipate the described aminoacylase to be highly attractive for biotechnological applications.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00253-023-12868-8>.

**Author contribution** GH: data collection and analysis (investigation, identification, cloning and expression, characterization, and initial synthesis screening). TJ: data collection and analysis (chemical and biocatalytic synthesis, instrumental analytics, and structure elucidation). SO: data collection (supported expression experiments and characterization). GH and TJ: writing the manuscript. JB, KEJ, US, and PS: manuscript revision. JB, PS, and US: funding acquisition, project design, and supervision. All authors read and approved the final manuscript.

**Funding** Open Access funding enabled and organized by Projekt DEAL. This work was funded by "Bundesministerium für Bildung und Forschung" (BMBF) (FKZ 13FH256PA6 and 13FH256PB6).

**Data availability** Data are available on request.

## Declarations

**Ethics approval** Not applicable.

**Consent to participate** All authors declared their consent to participate.

**Consent for publication** All authors declare their consent to publish their work.

**Competing interests** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Anders MW, Dekant W (1994) Aminoacylases. *Adv Pharmacol* 27:431–448. [https://doi.org/10.1016/S1054-3589\(08\)61042-X](https://doi.org/10.1016/S1054-3589(08)61042-X)
- Boivin S, Kozak S, Meijers R (2013) Optimization of protein purification and characterization using ThermoFluor screens. *Protein Expr Purif* 91:192–206. <https://doi.org/10.1016/j.pep.2013.08.002>

- Bourkaib MC, Delaunay S, Framboisier X, Hôtel L, Aigle B, Humeau C, Guiavarc'h Y, Chevalot I (2020a) N-acylation of L-amino acids in aqueous media: evaluation of the catalytic performances of *Streptomyces ambifaciens* aminoacylases. *Enzyme Microb Technol* 137:109536. <https://doi.org/10.1016/j.enzmictec.2020.109536>
- Bourkaib MC, Delaunay S, Framboisier X, Humeau C, Guilbot J, Bize C, Illous E, Chevalot I, Guiavarc'h Y (2020) Enzymatic synthesis of N-10-undecenoyl-phenylalanine catalysed by aminoacylases from *Streptomyces ambifaciens*. *Process Biochem* 99:307–315. <https://doi.org/10.1016/j.procbio.2020.09.009>
- Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72:248–254. <https://doi.org/10.1006/abio.1976.9999>
- Dettori L, Ferrari F, Framboisier X, Paris C, Guiavarc'h Y, Hôtel L, Aymes A, Leblond P, Humeau C, Kapel R, Chevalot I, Aigle B, Delaunay S (2018) An aminoacylase activity from *Streptomyces ambifaciens* catalyzes the acylation of lysine on  $\alpha$ -position and peptides on N-terminal position. *Eng Life Sci* 18:589–599. <https://doi.org/10.1002/elsc.201700173>
- Falkenberg F, Rahba J, Fischer D, Bott M, Bongaerts J, Siegert P (2022) Biochemical characterization of a novel oxidatively stable, halotolerant, and high-alkaline subtilisin from *Alkalihalobacillus okhensis* Kh10-101<sup>T</sup>. *FEBS Open Bio*. <https://doi.org/10.1002/2211-5463.13457>
- Gentzen I, Löffler H-G, Schneider F (1980) Aminoacylase from *Aspergillus oryzae*. Comparison with the pig kidney enzyme. *Z Naturforsch* 35:544–550. <https://doi.org/10.1515/znc-1980-7-804>
- Haeger G, Bongaerts J, Siegert P (2022) A convenient ninhydrin assay in 96-well format for amino acid-releasing enzymes using an air-stable reagent. *Anal Biochem* 654:114819. <https://doi.org/10.1016/j.ab.2022.114819>
- Haeger G, Wirges J, Tanzmann N, Oyen S, Jolmes T, Jaeger KE, Schörken U, Bongaerts J, Siegert P (2023) Chaperone assisted recombinant expression in *Vibrio natriegens* and *Escherichia coli* and characterization of a mycobacterial aminoacylase capable of N-lauroyl-L-amino acid synthesis. *Microb Cell Fact* 22(1):77
- Hanahan D (1983) Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* 166:557–580. [https://doi.org/10.1016/s0022-2836\(83\)80284-8](https://doi.org/10.1016/s0022-2836(83)80284-8)
- Henseling J, Röhm KH (1988) Aminoacylase I from hog kidney: anion effects and the pH dependence of kinetic parameters. *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism* 959(3):370–377
- Ibrahim TS, Seliem IA, Panda SS, Al-Mahmoudy AMM, Abdel-Samii ZKM, Alhakamy NA, Asfour HZ, Elagawany M (2020) An efficient greener approach for N-acylation of amines in water using benzotriazole chemistry. *Molecules* 25:2501. <https://doi.org/10.3390/molecules25112501>
- Joondan N, Jhaumeer Lauloo S, Caumul P (2018) Amino acids: building blocks for the synthesis of greener amphiphiles. *J Disp Sci Tech* 39:1550–1564. <https://doi.org/10.1080/01932691.2017.1421085>
- Kimura Y, Kobayashi Y, Adachi S, Matsuno R (1992) Aminoacylase-catalyzed synthesis of N-acyl amino acid from fatty acid or its ethyl ester and amino acid. *Biochem Eng* for 2001:109–111. [https://doi.org/10.1007/978-4-431-68180-9\\_27](https://doi.org/10.1007/978-4-431-68180-9_27)
- Koreishi M, Kawasaki R, Imanaka H, Imamura K, Nakanishi K (2005) A novel  $\epsilon$ -lysine acylase from *Streptomyces mobaraensis* for synthesis of N<sub>ε</sub>-acyl-L-lysines. *JAOCS* 82:631–637. <https://doi.org/10.1007/s11746-005-1121-2>
- Koreishi M, Kawasaki R, Imanaka H, Imamura K, Takakura Y, Nakanishi K (2009a) Efficient N<sub>ε</sub>-lauroyl-L-lysine production by recombinant  $\epsilon$ -lysine acylase from *Streptomyces mobaraensis*. *J Biotechnol* 141:160–165. <https://doi.org/10.1016/j.jbiotec.2009.03.008>
- Koreishi M, Nakatani Y, Ooi M, Imanaka H, Imamura K, Nakanishi K (2009b) Purification, characterization, molecular cloning, and expression of a new aminoacylase from *Streptomyces mobaraensis* that can hydrolyze N-(middle/long)-chain-fatty-acyl-L-amino acids as well as N-short-chain-acyl-L-amino acids. *Biosci Biotechnol Biochem* 73:1940–1947. <https://doi.org/10.1271/bbb.90081>
- Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680–685. <https://doi.org/10.1038/227680a0>
- Lindner HA, Alary A, Boju LI, Sulca T, Ménard R (2005) Roles of dimerization domain residues in binding and catalysis by aminoacylase-I. *Biochemistry* 44:15645–15651. <https://doi.org/10.1021/bi051180y>
- Liu Z, Zhen Z, Zuo Z, Wu Y, Liu A, Yi Q, Li W (2006) Probing the catalytic center of porcine aminoacylase 1 by site-directed mutagenesis, homology modeling and substrate docking. *J Biochem* 139:421–430. <https://doi.org/10.1093/jb/mvj047>
- Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Marchler GH, Song JS, Thanki N, Yamashita RA, Yang M, Zhang D, Zheng C, Lanczycki CJ, Marchler-Bauer A (2020) CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res* 48:D265–D268. <https://doi.org/10.1093/nar/gkz2991>
- Natsch A, Gfeller H, Gyax P, Schmid J, Acuna G (2003) A specific bacterial aminoacylase cleaves odorant precursors secreted in the human axilla. *J Biol Chem* 278:5718–5727. <https://doi.org/10.1074/jbc.M210142200>
- Pinheiro L, Faustino C (2017) Amino acid-based surfactants for biomedical applications. In: Najjar R (ed) Application and characterization of surfactants. IntechOpen, London, pp 207–232
- Strohalm M, Hassman M, Kosata B, Kódicck M (2008) mMass data miner: an open source alternative for mass spectrometric data analysis. *Rapid Commun Mass Spectrom* 22:905–908. <https://doi.org/10.1002/rcm.3444>
- Takakura Y, Asano Y (2019) Purification, characterization, and gene cloning of a novel aminoacylase from *Burkholderia* sp. strain LP5\_18B that efficiently catalyzes the synthesis of N-lauroyl-L-amino acids. *Biosci Biotechnol Biochem* 83:1964–1973. <https://doi.org/10.1080/09168451.2019.1630255>
- Torres-Bacete J, Hormigo D, Torres-Gúzman R, Arroyo M, Castellón MP, García LJ, Acebal C, de Matai LA (2015) Overexpression of penicillin V acylase from *Streptomyces lavendulae* and elucidation of its catalytic residues. *Appl Microbiol Biotechnol* 81:1225–1233. <https://doi.org/10.1128/AEM.02352-14>
- Ueda S, Shibata T, Ito K, Oohata N, Yamashita M, Hino M, Yamada M, Isogai Y, Hashimoto S (2011) Cloning and expression of the FR901379 acylase gene from *Streptomyces* sp. no. 6907. *J Antibiot* 64:169–175. <https://doi.org/10.1038/ja.2010.151>
- Wada E, Handa M, Imamura K, Sakiyama T, Adachib S, Matsunob R, Nakanishia K (2002) Enzymatic synthesis of N-acyl-L-amino acids in a glycerol-water system using acylase I from pig kidney. *JAOCS* 79:41–46. <https://doi.org/10.1007/s11746-002-0432-7>
- Wardenga R, Hollmann F, Thum O, Bornscheuer U (2008) Functional expression of porcine aminoacylase 1 in *E. coli* using a codon optimized synthetic gene and molecular chaperones. *Appl Microbiol Biotechnol* 81:721–729. <https://doi.org/10.1007/s00253-008-1716-7>
- Wardenga R, Lindner HA, Hollmann F, Thum O, Bornscheuer U (2010) Increasing the synthesis/hydrolysis ratio of aminoacylase 1 by site-directed mutagenesis. *Biochimie* 92:102–109. <https://doi.org/10.1016/j.biochi.2009.09.017>
- Xiang DF, Patskovsky Y, Xu C, Fedorov AA, Fedorov EV, Sisco AA, Sauder JM, Burley SK, Almo SC, Raushel FM (2010)

- Functional identification and structure determination of two novel prolidases from cog1228 in the amidohydrolase superfamily. *Biochemistry* 49:6791–6803. <https://doi.org/10.1021/bi100897u>
- Zhang D, Koreishi M, Imanaka H, Imamura K, Nakanishi K (2007) Cloning and characterization of penicillin V acylase from *Streptomyces mobaraensis*. *J Biotechnol* 128:788–800. <https://doi.org/10.1016/j.jbiotec.2006.12.017>
- Zhao H, He C, Zhou Y, Yang J, Luo C, Xu B (2019) Study on foaming properties of N-acyl amino acid surfactants: sodium N-acyl glycinate and sodium N-acyl phenylalaninate. *Colloids Surfaces a* 567:240–248. <https://doi.org/10.1016/j.colsurfa.2019.01.073>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## 2.3. Chapter III

### **Chaperone assisted recombinant expression in *Vibrio natriegens* and *Escherichia coli* and characterization of a mycobacterial aminoacylase capable of N-lauroyl-L-amino acid synthesis**

Gerrit Haeger, Jessika Wirges, Nicole Tanzmann, Sven Oyen, Tristan Jolmes, Karl-Erich Jaeger, Ulrich Schörken, Johannes Bongaerts, Petra Siegert

Microbial Cell Factories 22:77 (2023), doi: 10.1186/s12934-023-02079-1

#### Author contributions:

GH designed the study. GH conducted cloning and bioinformatic analysis. GH, JW, NT, SO performed the experiments and analyzed the data for protein expression, purification, and biochemical characterization. TJ synthesized acyl-amino acids chemically and performed HPLC-MS analysis. GH wrote the manuscript. GH, US, JB, KEJ and PS edited the manuscript. PS and JB supervised the work of GH, JW, NT and SO and US supervised the work of TJ. PS, JB and US did funding acquisition. All authors read and approved the final manuscript.

Overall contribution GH: 85 %

## RESEARCH

## Open Access



# Chaperone assisted recombinant expression of a mycobacterial aminoacylase in *Vibrio natriegens* and *Escherichia coli* capable of N-lauroyl-L-amino acid synthesis

Gerrit Haeger<sup>1</sup>, Jessika Wirges<sup>1</sup>, Nicole Tanzmann<sup>1</sup>, Sven Oyen<sup>1</sup>, Tristan Jolmes<sup>2</sup>, Karl-Erich Jaeger<sup>3,4</sup>, Ulrich Schörken<sup>2</sup>, Johannes Bongaerts<sup>1</sup> and Petra Siegert<sup>1\*</sup>

## Abstract

**Background** Aminoacylases are highly promising enzymes for the green synthesis of acyl-amino acids, potentially replacing the environmentally harmful Schotten-Baumann reaction. Long-chain acyl-amino acids can serve as strong surfactants and emulsifiers, with application in cosmetic industries. Heterologous expression of these enzymes, however, is often hampered, limiting their use in industrial processes.

**Results** We identified a novel mycobacterial aminoacylase gene from *Mycobacterium smegmatis* MKD 8, cloned and expressed it in *Escherichia coli* and *Vibrio natriegens* using the T7 overexpression system. The recombinant enzyme was prone to aggregate as inclusion bodies, and while *V. natriegens* Vmax™ could produce soluble aminoacylase upon induction with isopropyl β-d-1-thiogalactopyranoside (IPTG), *E. coli* BL21 (DE3) needed autoinduction with lactose to produce soluble recombinant protein. We successfully conducted a chaperone co-expression study in both organisms to further enhance aminoacylase production and found that overexpression of chaperones GroEL/S enhanced aminoacylase activity in the cell-free extract 1.8-fold in *V. natriegens* and *E. coli*. Eventually, *E. coli* ArcticExpress™ (DE3), which co-expresses cold-adapted chaperonins Cpn60/10 from *Oleispira antarctica*, cultivated at 12 °C, rendered the most suitable expression system for this aminoacylase and exhibited twice the aminoacylase activity in the cell-free extract compared to *E. coli* BL21 (DE3) with GroEL/S co-expression at 20 °C. The purified aminoacylase was characterized based on hydrolytic activities, being most stable and active at pH 7.0, with a maximum activity at 70 °C, and stability at 40 °C and pH 7.0 for 5 days. The aminoacylase strongly prefers short-chain acyl-amino acids with smaller, hydrophobic amino acid residues. Several long-chain amino acids were fairly accepted in hydrolysis as well, especially N-lauroyl-L-methionine. To initially evaluate the relevance of this aminoacylase for the synthesis of N-acyl-amino acids, we demonstrated that lauroyl-methionine can be synthesized from lauric acid and methionine in an aqueous system.

\*Correspondence:

Petra Siegert  
siegert@fh-aachen.de

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

**Conclusion** Our results suggest that the recombinant enzyme is well suited for synthesis reactions and will thus be further investigated.

**Keywords** Aminoacylase, *Vibrio natriegens*, Chaperone co-expression, Inclusion bodies, Acyl-amino acids

## Background

L-Aminoacylases catalyze the hydrolysis of N-acyl-L-amino acids. The reverse reaction, namely the synthesis of acyl-amino acids from amino acids and fatty acids, is also possible. Conventionally, these molecules are produced by chemical synthesis using the Schotten-Baumann reaction, which needs to employ fatty acyl chlorides for acylation. Besides the chlorination with toxic chemicals like thionyl chloride or phosgene, sodium chloride is stoichiometrically produced as waste. Therefore, aminoacylases are highly interesting enzymes for an alternative biocatalytic, green synthesis. Among aminoacylases (EC 3.5.1.14) mostly mammalian enzymes have been studied. The porcine enzyme pAcy1 was extensively characterized regarding mechanism and mode of action [1, 2], recombinant expression [3] and synthesis of acyl-amino acids [4–6]. The industrial application is hampered by substrate specificity, difficult expression and insufficient stability. The fungal aminoacylases, sourced from *Aspergillus melleus* or *A. oryzae*, in contrary, have a long-established biocatalytic application. However, they are primarily used for the production of enantiomerically pure L-amino acids from a racemic mixture of acetylated amino acids, rather than for the synthesis of acyl-amino acids. Aminoacylases for the synthesis of acyl-amino acids should accept long acyl chains as substrates, so that amino acid-based surfactants or emulsifiers can be produced. Furthermore, the biocatalysts should be easily produced and purified and should be sufficiently stable.

Bacterial aminoacylases show more favorable characteristics for the synthesis of acyl-amino acids and have been found in various bacterial genera, among them *Bacillus* [7, 8], *Burkholderia* [9], *Corynebacterium* [9], *Pseudomonas* [10], or *Streptomyces* [11–14]. Aminoacylases from *Streptomyces* species seem promising candidates, as multiple strains have been identified that produce enzymes with interesting properties. An alpha-aminoacylase (AA) and an epsilon-lysine-acylase (ELA) have been described in *S. mobaraensis* (SmAA [12], SmELA [11, 15]) and *S. ambofaciens* (SamAA, SamELA [14, 16]). Especially SamAA from *S. ambofaciens* showed good catalytic potential and a broad substrate specificity [14, 17]. Besides these bacterial sources, aminoacylases have also been described from *Mycobacterium smegmatis* [18, 19] in 1964 and 1972, respectively. Protein sequences for these enzymes were not specified, and no further data regarding mycobacterial aminoacylases have been published yet. The gram-positive bacteria *Streptomyces* and *Mycobacterium* are distantly related, as they belong

to the class of Actinomycetia, characterized by a high G+C content of their genomes. At first glance, members of those two genera are clearly distinct, but similarities can be found on a cellular and a molecular level [20]. A unique characteristic of *Mycobacteria* is the occurrence of extremely long-chain mycolic acids that form a waxy layer in the bacterial cell wall suggesting that these organisms may possess aminoacylases which accept long-chain acyl-amino acids.

For characterization and evaluation of the potential in synthesis of acyl-amino acids, it is advantageous to heterologously overexpress the relevant aminoacylase. This way, enough protein of interest can be obtained. By fusing the aminoacylase gene to an affinity tag, purification is straightforward. In most cases, *E. coli* is used as an expression host strain [21], however, problems may occur including inclusion body (IB) formation. This is particularly relevant if the T7 expression system is used [22]. To prevent this, the *E. coli* Tuner™ strain (Novagen), a *lacY* deletion mutant that allows to adjust the cellular IPTG concentration, has been constructed [23]. Another option is the usage of autoinduction media that contain glucose and lactose in which bacteria first grow to high densities before expression is induced [24].

Another approach to increase the solubility of aggregation-prone recombinant proteins is the co-expression of molecular chaperones, which can help to fold otherwise misfolded proteins [25]. Most commonly, combinations of so-called heat-shock proteins like GroEL/GroES and DnaK/DnaJ/GrpE, or foldases like Trigger Factor (Tf) can be co-expressed. Since these chaperones function differently, positive influence on recombinant proteins must be tested individually or from combinations thereof. Trigger Factor is an ATP-independent chaperone with additional peptidylprolyl isomerase activity that binds to nascent polypeptide chains at ribosomes and is thus the first chaperone to interact with newly synthesized proteins. It mediates first folding steps, can prevent premature folding and restricts access of downstream chaperones or can prevent premature degradation [26]. The chaperones DnaK, DnaJ and GrpE function in concert and should thus be co-expressed, as overexpression of DnaK or DnaJ alone can impair cell viability. The ATP-dependent DnaK is responsible for main folding activity and can work both co- and post-translationally, and DnaJ and GrpE modulate ATP hydrolysis and conformational changes [25]. The chaperonine GroEL/GroES is a multimeric protein that forms a barrel-like structure composed of two heptameric rings of GroEL which can be closed by heptameric

rings of GroES on each side, forming a lid. Upon entering of the misfolded protein into the cavity formed by the GroEL ring, the lid formed by GroES closes and the protein substrate gets folded by conformational change directed by hydrolysis of ATP [27]. As an alternative expression host, we have investigated *Vibrio natriegens*, a marine bacterium which has recently gathered attention for its use in biotechnology due to its extraordinary high growth rates. Its close genetic proximity to *E. coli* renders many molecular biology tools applicable also for *V. natriegens*, as shown for the plasmid pET-based expression system, making this strain an interesting alternative to *E. coli* for heterologous protein expression [28, 29].

Here we present the characterization of a novel mycobacterial aminoacylase and show its general applicability in the synthesis of acyl-amino acids. The recombinant expression by classical IPTG-induction in *E. coli* BL21 (DE3) led to inclusion body formation without obtaining soluble protein. We improved the expression by changing the mode of induction, co-expression of molecular chaperones, and adjusting expression temperatures. Furthermore, expression in *V. natriegens* with chaperone co-expression was established.

## Results and discussion

### Cloning and sequence analysis of the *msAA* gene from *Mycobacterium smegmatis* MKD 8

The motivation of this study was to identify a novel aminoacylase gene and facilitate its heterologous expression. The aminoacylase should have the ability to synthesize acyl-amino acids with long acyl chains and was searched in the sequence scope of *Mycobacteria*. The aminoacylase MsAA was identified by NCBI BLASTp homology search. The streptomycetal protein sequences for SmAA (Accession No. BAI44523.1) and SamAA (Accession No. AKZ54783.1) were used as baits in a BLASTp search for homologous sequences in *Mycobacterium smegmatis*. The search revealed a protein sequence which shows 58% similarity to both SmAA and SamAA and was designated MsAA. The encoding nucleotide sequence was deduced. The obtained DNA sequence was optimized regarding the codon usage of *E. coli* and the synthetic DNA strand was ordered. By Golden Gate cloning, the gene was cloned into pET28a expression vectors with either N- or C-terminal Strep-tag or without an affinity tag. The calculated molecular weight of MsAA is 48.5 kDa and 49.9 kDa with the affinity tag.

### Protein sequence analysis and structure prediction

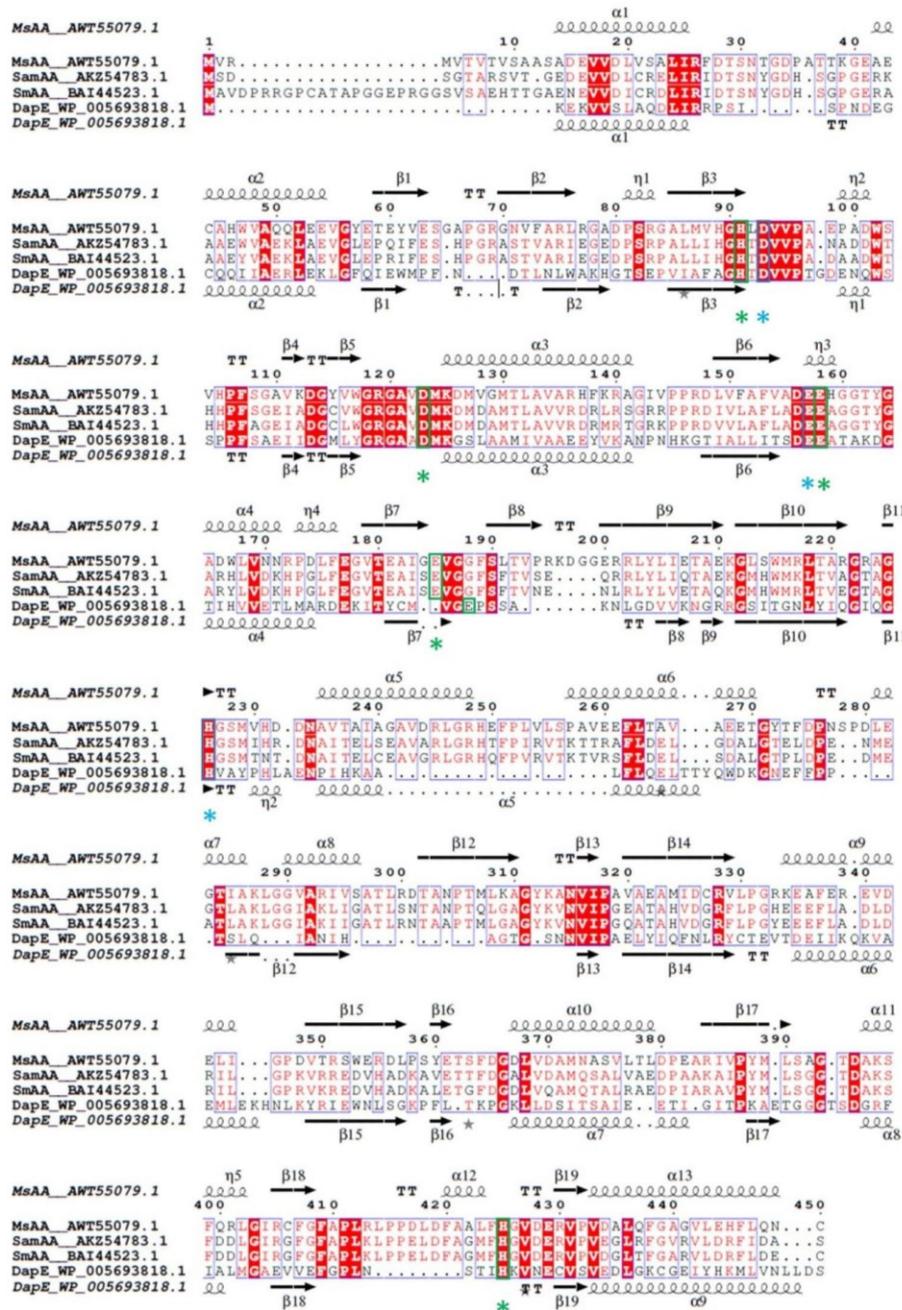
Analysis of the protein sequence comprising 450 amino acids revealed conserved catalytic (D93, E157, H226) and metal-binding (H91, D123, E158, E185, H425) residues as well as signature motifs for the M20 peptidase family [12, 30, 31]. Members of the M20 peptidase family also

include non-peptidase homologues and bind two cocatalytic zinc ions, usually bound by histidine, glutamic acid or aspartic acid [32]. The three-dimensional protein structure of dimeric MsAA was predicted with the ColabFold algorithm. In general, the generated structure showed very good per-residue confidence metric with a pLDDT of 95.2 (local accuracy) and a pTM score of 0.92 (global accuracy). The multiple sequence alignment of MsAA, SamAA, SmAA, and DapE is shown in Fig. 1, along with the secondary structural elements taken from the predicted structure of MsAA and a crystal structure from DapE. We used the MIB metal ion-binding site prediction and docking server to predict zinc binding sites based on the AlphaFold-generated structure. Two zinc-binding sites were predicted in which the zinc atoms are bound by residues which are conserved in the M20 peptidase family. The distance of metal-binding atoms and zinc ions was approximately 2 Å. The predicted structure of MsAA with its bound zinc ions is shown in Fig. 2. Structurally, members of the M20 family can either be monomeric or dimeric proteins. One domain of the structure is referred to as the catalytic domain, as it contains catalytic and metal-binding residues, while the other domain is called lid domain for monomeric or dimerization domain for dimeric enzymes. The two domains form an internal cavity harboring the active site [32]. In the case of the homologous dimeric DapE, a succinyl-diaminopimelate desuccinylase, a histidine (H194) from the dimerization domain enters the active site of the opposing monomer [31]. This histidine is also conserved in MsAA as H226 (Fig. 1). From the predicted dimeric protein structure of MsAA, the characteristic dimerization domain and catalytic domain can be distinguished (Fig. 2(A)). The active site is located in the cavity formed between the two domains. The H226 residue from one dimer reaches into the active site of the opposing dimer (Fig. 2(B)).

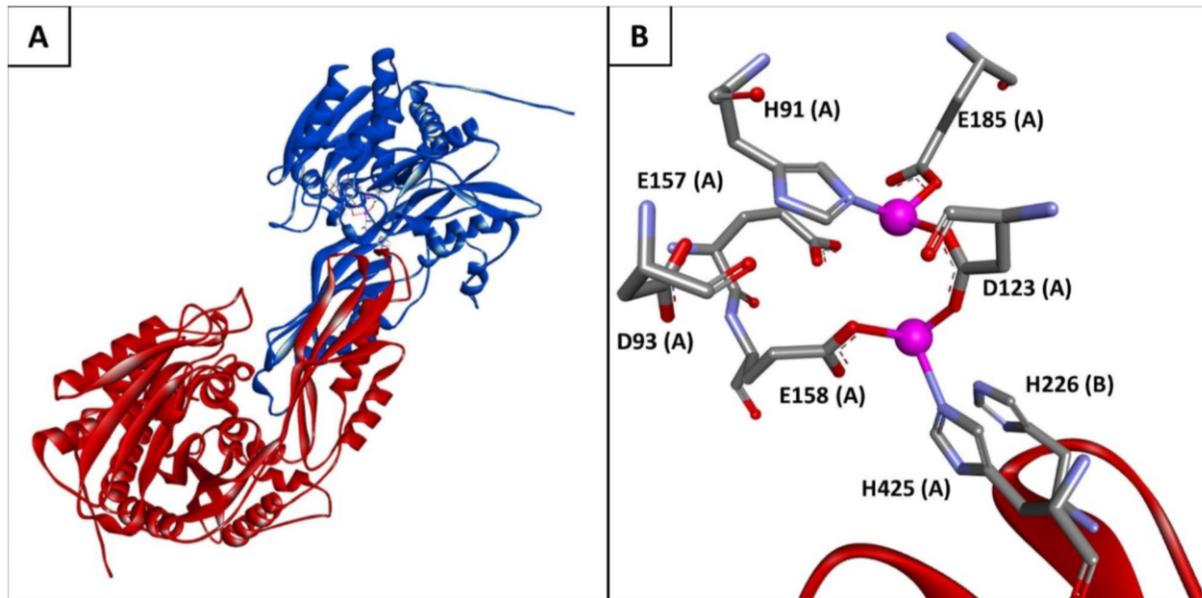
### Comparative expression of MsAA in *E. coli* and *V. natriegens* with co-expression of molecular chaperones and affinity purification

Expression of aminoacylase MsAA in *E. coli* BL21 (DE3) with IPTG resulted in the formation of inclusion bodies. The issue could be solved by growing the cells in autoinduction medium. This way, a fraction of soluble MsAA was obtained; however, a significant amount of insoluble enzyme still prevailed. Hence, we systematically optimized the expression conditions in *E. coli* and *V. natriegens* by varying induction mechanism, cultivation temperature and duration as well as by co-expression of molecular chaperones. To our best knowledge, this is the first study on chaperone co-expression in *V. natriegens*.

Since *V. natriegens* does not carry the genes for lactose import and catabolism, autoinduction with lactose was



**Fig. 1** Multiple sequence alignment of MsAA and homologous proteins  
 MsAA from *M. smegmatis* (AWT55079.1), the aminoacylases SamAA from *S. ambofaciens* (AKZ54783.1), SmAA from *S. mobaraensis* (BAI44523.1), and DapE from *Haemophilus influenzae* (WP\_005693818.1) are shown. The alignment generated using T-Coffee and displayed with ESPrift 3.0. The conserved metal-binding residues (H91, D123, E158, E185, H425) and catalytic residues (D93, E157, H226) are highlighted by green and blue boxes with asterisks. The annotation of the secondary structural elements shown above is based on the ColabFold-generated structure of MsAA; the secondary structural elements shown below belong to DapE (imported from PDB-entry 5vo3). The secondary structure is displayed with arrows for β-strands, squiggles for α-helices or  $3_{10}$ (η)-helices and the letters TT for turns



**Fig. 2** Predicted protein structure of MsAA using ColabFold2.

(A) 3D structure with monomers shown in red and blue. (B) Active site of MsAA with two bound zinc ions (magenta balls) and metal-binding and catalytic residues from both dimers (A and B) are shown

not possible; instead, expression was induced with IPTG. To better compare the expression potential of *E. coli* and *V. natriegens*, the *lacY*-deletion mutant *E. coli* Tuner (DE3), was investigated as well.

#### Expression in *Vibrio natriegens* Vmax™

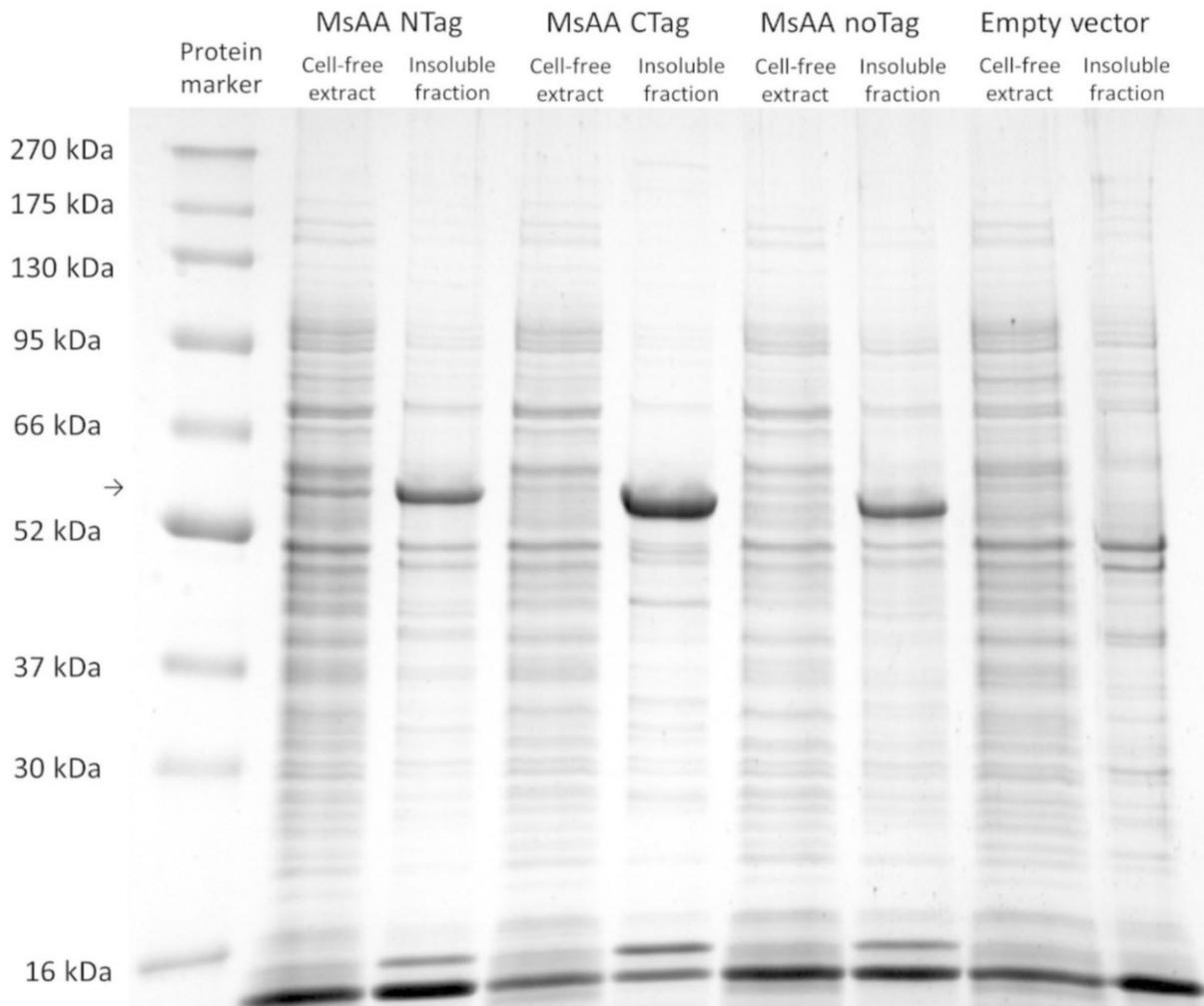
As an alternative system for T7-based, IPTG induced expression, *V. natriegens* Vmax™ was tested for expression of MsAA. In a first step, MsAA was expressed with all Strep-tag variants, namely N- or C-terminally or without affinity tag and at 37 °C, 30 °C and 20 °C to determine the optimal expression temperature. As the N-terminally tagged MsAA was best produced in soluble form and yielded less inclusion bodies, MsAA NTag was further used in this study. Regarding temperature, expression 30 °C (Fig. 3) was found to be most suitable, as IB formation was reduced compared to 37 °C (Additional file 1: Fig. S1). At 20 °C, growth and total protein production was hampered (Additional file 1: Fig. S2). Strikingly, in contrast to *E. coli* BL21 (DE3), *V. natriegens* Vmax could form soluble and active MsAA upon induction with 1 mM IPTG. Therefore, we conducted the chaperone co-expression experiments at 30 °C cultivation temperature for 4 h.

We transformed *V. natriegens* Vmax with the chaperone plasmid set from Takara Bio containing pGKJE8 (GroEL/S, DnaK/J/GrpE), pKJE7 (DnaK/J/GrpE), pGTf2 (GroEL/S, Tf), pTf16 (Tf) or pGro7 (GroEL/S) and pET28a MsAA NTag to generate chaperone overexpression strains. Without chaperones, activity and specific

activity in the cell-free extract was 8.7 ( $\pm 0.5$ ) U/ml and 1.9 ( $\pm 0.2$ ) U/mg, respectively. The co-expression with pGro7 lead to an almost 1.8-fold increase of activity to 15.5 ( $\pm 0.6$ ) U/ml. Specific activity also increased 1.8-fold to 3.4 ( $\pm 0.2$ ) U/mg. The co-expression from pGTf2 and pTf16 also raised activity, to 12.2 ( $\pm 0.9$ ) U/ml and 10.3 ( $\pm 0.4$ ) U/ml, respectively, while having a non-significant effect on specific activity. In contrast, pGKJE8 and pKJE7 had detrimental effects on soluble protein expression, abolishing aminoacylase activity in the cell-free extract (Fig. 4(A); gels from SDS-PAGE in Additional file 1: Fig. S3 and S4). Cultivation for further 20 h led to reduction of activity (1.6 U/ml for Vmax and 7.6 U/ml for Vmax pGro7; data not shown). We purified recombinant MsAA NTag from *V. natriegens* Vmax and *V. natriegens* Vmax pGro7 and obtained specific activities of 122.1 ( $\pm 0.5$ ) U/mg and 129.2 ( $\pm 3.4$ ) U/mg (results of SDS-PAGE analysis in Additional file 1: Fig. S5 and S6).

#### Expression in *E. coli* BL21 (DE3)

With *E. coli* BL21 (DE3), *E. coli* BL21 (DE3) pGro7 and *E. coli* Tuner (*lacZY*<sup>-</sup>), no aminoacylase activity could be detected in the cell-free extracts upon IPTG induction, and MsAA was found only in the insoluble fractions (Additional file 1: Fig. S7 and S8). Hence, autoinduction with 0.2% lactose additionally to glucose in TB-AIM medium was performed. Growth of *E. coli* was much slower compared to *V. natriegens* Vmax, so the cells were harvested after 24 h cultivation at 30 °C. *E. coli* BL21



**Fig. 3** SDS-PAGE analysis of cell extracts obtained from *V. natriegens* Vmax™ after expression of MsAA variants

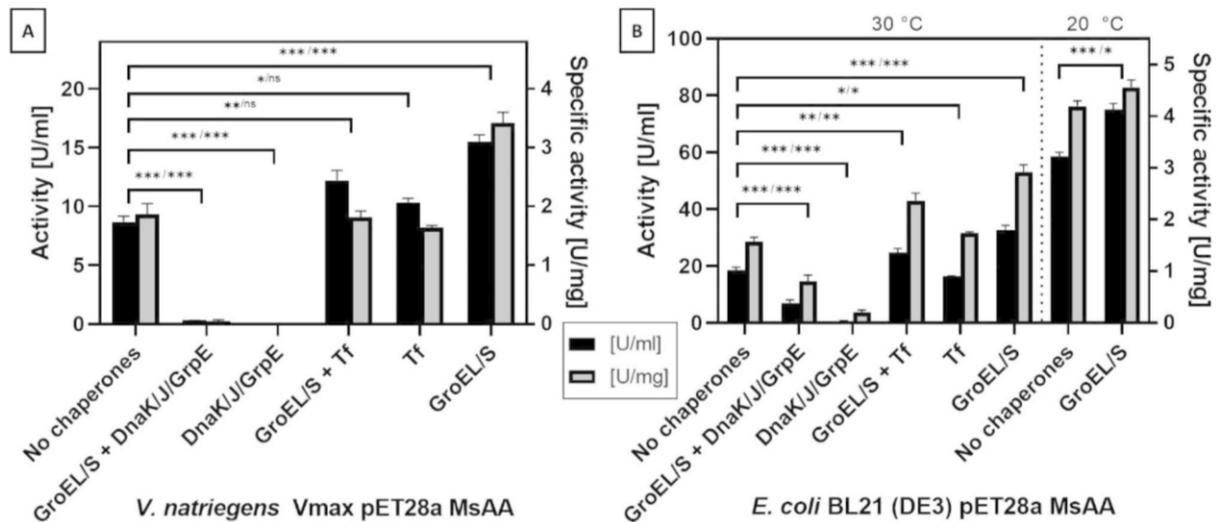
Cells were grown in BHLv2 medium for 5 h at 30 °C. The arrow indicates the position of MsAA variants

Lane 1: protein marker (BlueEasy Prestained Protein Marker, Nippon Genetics); lane 2: cell-free extract with MsAA NTag; lane 3: insoluble fraction with MsAA NTag; lane 4: cell-free extract with MsAA CTag; lane 5: insoluble fraction with MsAA CTag; lane 6: cell-free extract with MsAA noTag; lane 7: insoluble fraction with MsAA noTag; lane 8: cell-free extract of empty vector control; lane 9: insoluble fraction of empty vector control

(DE3) was transformed with the five chaperone plasmids and all strains were used for the expression study.

The co-expression of chaperones in *E. coli* revealed results comparable to *V. natriegens* Vmax™. The co-expression of GroEL/S from pGro7 had the greatest influence on aminoacylase activity, raising activity and specific activity 1.8-fold from 18.5 ( $\pm 1.0$ ) U/ml to 32.7 ( $\pm 1.6$ ) U/ml and from 1.6 ( $\pm 0.1$ ) U/mg to 2.9 ( $\pm 0.1$ ) U/mg, respectively. Furthermore, expression of chaperones GroEL/GroES and Tf from plasmid pGTF2 was beneficial with respect to formation of soluble MsAA. Again, the co-expression of chaperones from plasmids pGKJE8 and pKJE7 resulted in lowered enzymatic activity (Fig. 4(B); SDS-PAGEs in Additional file 1: Fig. S9 and S10).

Additionally, we compared the expression with and without pGro7 at 20 °C. We found enhanced aminoacylase activity compared to expression at 30 °C in both strains with 58.6 ( $\pm 1.4$ ) U/ml in *E. coli* BL21 (DE3) and 75.0 ( $\pm 2.2$ ) U/ml in *E. coli* BL21 (DE3) pGro7. Hence, activity could be more than doubled by lowering the expression temperature. The influence of the chaperone co-expression on activity was less pronounced by lowering the expression temperature and led to approximately 1.3-fold enhancement. Protein formation is slower at 20 °C giving proteins more time to fold correctly; the refolding activity of GroEL/S may thus not be required. On the other hand, the chaperonine GroEL/S has its temperature optimum at 30 °C and might thus not work



**Fig. 4** Effect of chaperone co-expression on MsAA NTag overexpression on aminoacylase activity

Aminoacylase activity and specific activity were measured from the cell-free extract from *V. natriegens* Vmax™ and *E. coli* BL21 (DE3). Activity was measured in triplicates and expression without and with chaperones were repeated as three biological replicates. Statistical significance for the differences in activity and specific activity calculated by unpaired students t-test was indicated as followed: not significant (ns) for  $p > 0.05$ , \* for  $p \leq 0.05$ , \*\* for  $p \leq 0.01$ , and \*\*\* for  $p \leq 0.001$

as efficiently at 20 °C [33]. Nonetheless, for *E. coli* BL21 (DE3), the optimized expression conditions for MsAA include co-expression of GroEL/S from pGro7 at 20 °C cultivation temperature, induction of chaperone expression with 0.5 mg/ml arabinose at the start of the cultivation in TB-AIM growth medium, autoinduction of the pET-system with 0.2% lctose and harvesting after 24 h expression. We purified recombinant MsAA from *E. coli* BL21 (DE3) and *E. coli* BL21 (DE3) pGro7 and found specific activities of 129.4 ( $\pm 0.9$ ) U/mg and 131.0 ( $\pm 6.0$ ) U/mg (Additional file 1: Fig. S11 and S12).

In summary, the co-expression of GroEL/S from pGro7 in *V. natriegens* and *E. coli* had the most beneficial effect on soluble recombinant protein formation and hence on aminoacylase activity in the cell-free extract. This is in line with data from literature which also indicate that a chaperone often enhances solubility of the protein of interest, while others have no or even negative effects [3, 34–36].

#### Expression in *E. coli* ArcticExpress

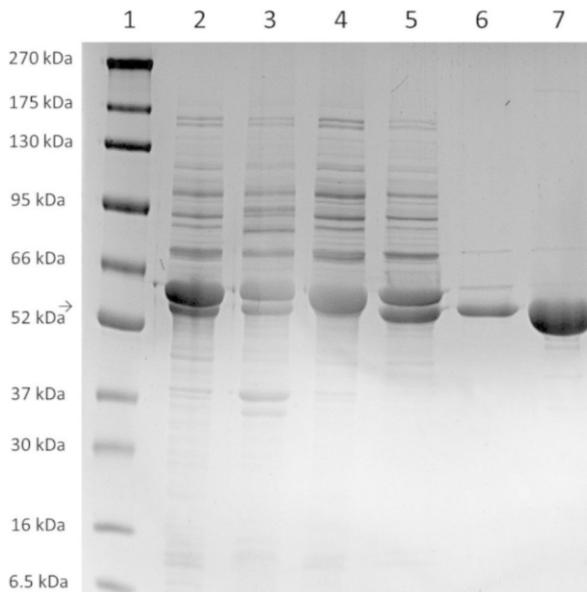
Since in both *V. natriegens* Vmax and *E. coli* BL21 (DE3), GroEL/S had a positive effect on the production of soluble heterologous aminoacylase and lowering the temperature to 20 °C increased the yield of recombinant protein, we combined these findings by using another chaperone co-expressing strain, *E. coli* ArcticExpress (DE3), for production of MsAA. *E. coli* ArcticExpress carries a plasmid for expression of a cold-adapted chaperonin Cpn60/10, a homolog of GroEL/S isolated from *Oleispira antarctica*, a psychrophilic bacterium. This chaperonin has its

temperature optimum in a range of 4 to 12 °C [33] allowing expression at a temperature of 12 °C. Aminoacylase activity and specific activity in the cell-free extract was measured as 139.1 ( $\pm 6.9$ ) U/ml and 9.2 ( $\pm 0.5$ ) U/mg, respectively. This were the highest activities observed among the strains used for MsAA overexpression and thus, *E. coli* ArcticExpress was used for production of MsAA for biochemical characterization. The enzyme was purified to homogeneity and specific activity of purified MsAA from *E. coli* ArcticExpress was 127.2 ( $\pm 4.8$ ) U/mg. The SDS-PAGE analysis of cell extracts purification is shown in Fig. 5. The theoretical mass of purified MsAA containing the N-terminal Strep-tag was verified by MALDI-TOF analysis as 49.9 kDa (Additional file 1: Fig. S13). Since members of the M20 peptidase family can be both monomeric or multimeric enzymes, native PAGE was performed with MsAA and reference proteins. We found that the recombinant MsAA is a dimeric protein (Additional file 1: Fig. S14). By isoelectric focusing, it was shown that the enzyme has a pI of approximately 4.3 (Additional file 1: Fig. S15).

#### Biochemical characterization of MsAA

##### Effect of pH and temperature on activity and stability of recombinant MsAA

The purified recombinant aminoacylase MsAA was characterized regarding its properties in hydrolytic reactions. The optimal pH for the hydrolytic reaction was pH 7.0 with 127.2 ( $\pm 4.8$ ) U/mg (Fig. 6(A)). Investigation of the pH-dependent stability revealed good stability at neutral pH values, as the enzyme was most stable at pH 6.0

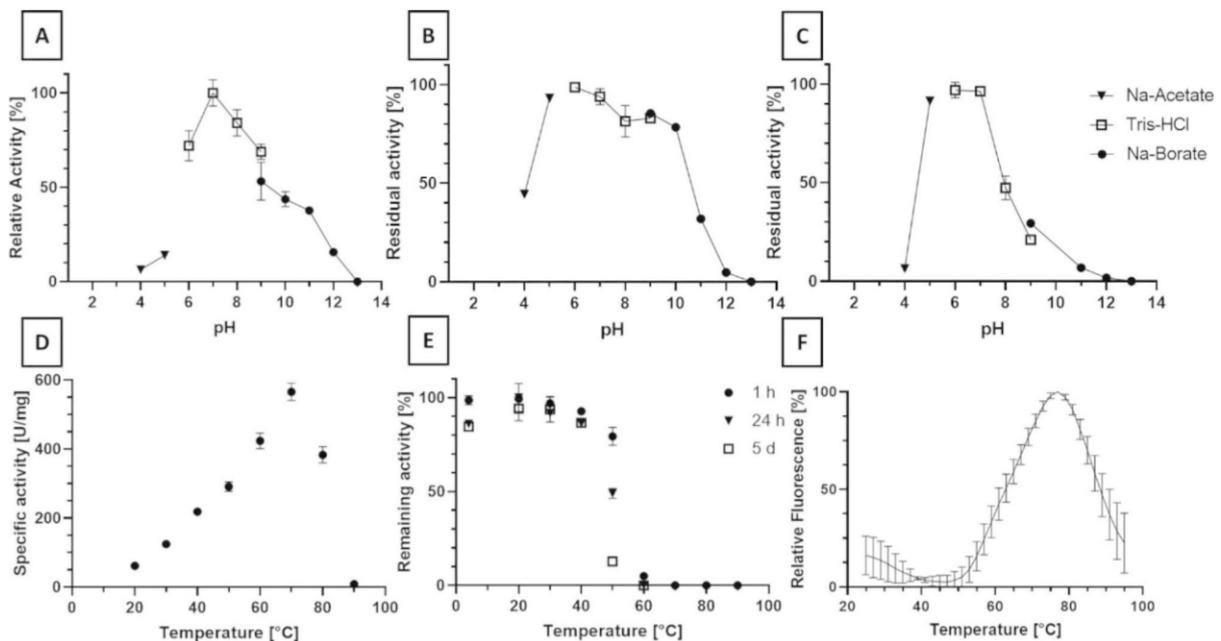


**Fig. 5** SDS-PAGE of MsAA NTag overexpression and Strep-tag purification from *E. coli* ArcticExpress (DE3).

Lane 1: Protein marker (BlueEasy Prestained Protein Marker, Nippon Genetics); lane 2: cell-free extract; lane 3: insoluble fraction; lane 4: Flow-through; lane 5/6: wash fractions; lane 7: Elution of MsAA NTag.

and 7.0. Stability of the enzyme quickly decreased at pH values of 4.0 and 11.0 (Fig. 6(B and C)). After 24 h incubation, the enzyme was still relatively stable in alkaline buffers up to pH 10.0. After 5 days, however, 47.3% activity remained at pH 8.0. Thus, the recombinant aminoacylase is best stored at pH 7.0 and storage buffer was 100 mM Tris-HCl pH 7.0 150 mM NaCl 1 mM ZnCl<sub>2</sub>.

With the optimal hydrolytic reaction temperature being 70 °C with 565.8 (±25.3) U/mg, the enzyme is more active at higher temperatures (Fig. 6(D)). However, the stability starts to decrease at 50 °C. At 40 °C, the enzyme is stable even for five days (Fig. 6(E)). The thermal shift assay also revealed that thermal denaturation starts at approximately 50 °C (Fig. 6(F)). Considering its stability, reaction temperatures for the biocatalytic production of acyl-amino acids should not exceed 40 °C for a long time. The optimal pH for the condensation reaction might differ from the hydrolytic optimum. Still, the enzyme is stable at rather neutral or slightly acidic pH values of 6.0–7.0.



**Fig. 6** Biochemical characterization of MsAA

(A) pH dependency of hydrolytic activity. Reaction conditions: 15 mM acetyl-alanine at 30 °C reaction temperature. The following buffers were used at 100 mM: Na-acetate for pH 4.0 and 5.0, Tris-HCl for pH 6.0–9.0, Na-borate for pH 9.0–13.0. (B) pH dependency of stability after 24 h and (C) after 5 d at 30 °C. Buffers and corresponding symbols used are identical to the pH activity optimum. Residual activity was determined with the standard assay conditions and indicated as percental values. (D) Temperature dependency on hydrolytic activity. Reaction conditions were 15 mM acetyl-alanine in 100 mM Tris-HCl pH 7.0 at various reaction temperatures. (E) Thermal stability after 1 h, 24 h and 5 days incubation at various temperatures in 100 mM Tris-HCl buffer pH 7.0. (F) Thermal shift assay. Thermal denaturation of MsAA was followed via fluorescence measurement of SYPRO Orange. All reactions were conducted in triplicates

### Effect on metal ions and chelating agent EDTA on hydrolytic activity

We investigated the metal ion dependency by incubation with various divalent metal cations on MsAA. When the enzyme was purified without the addition of metal ions in the purification buffers, activity was 51% ( $65.2 \pm 0.6$  U/mg) of the activity with zinc added during purification. Addition of 1 mM EDTA to this enzyme preparation decreased its activity from 51 to 36% ( $45.7 \pm 3.6$  U/mg). When 1 mM EDTA was added to the enzyme saturated with 1 mM  $ZnCl_2$ , activity barely decreased. The results suggest that the zinc ions bound to the enzyme can hardly be removed via complexation using the chelating agent EDTA.

The addition of  $MgSO_4$ ,  $MnCl_2$ ,  $CaCl_2$ , or  $CoCl_2$  at a concentration of 1 mM to the enzyme purified without the addition of metal ions had no significant influence on hydrolytic activity, while 1 mM of  $NiSO_4$ ,  $FeSO_4$ , and  $CuSO_4$  lowered the specific activity to 42% ( $54.4 \pm 5.5$  U/mg), 41% ( $52.9 \pm 1.1$  U/mg), and 40% ( $51.2 \pm 0.2$  U/mg), respectively. The addition of 1 mM  $ZnCl_2$  increased the activity to 84% ( $106.8 \pm 8.6$  U/mg) referring to activity when MsAA was purified with 1 mM zinc ions in each buffer. The results clearly indicate that MsAA is a zinc-dependent enzyme, as it is typical for members or homologues of the M20 peptidase family. Incubating the enzyme at  $ZnCl_2$  concentrations from 1 nM to 5 mM revealed that a minimal concentration of 10  $\mu M$   $ZnCl_2$  is needed in the enzyme storage buffer for full activity of MsAA. The results are summarized in Table 1.

### Hydrolytic substrate specificity

The hydrolytic substrate specificity was determined against various acyl-amino acids. The enzyme generally prefers amino acids acylated with acetic acid compared to longer or bulkier acyl chains (Table 2). In the case of alanine, valine, isoleucine and leucine, the ratio of hydrolytic activity against lauroyl-amino acids and the respective acetyl-amino acids is only 7.7%, 2.2%, 0.8% and 0.3%, respectively. In contrast, the activity against lauroyl-methionine is 39.0% of the activity against acetyl-methionine. Hence, lauroyl-methionine is accepted well by MsAA and hydrolytic activity against this substrate is comparable to other acetyl-amino acids. Comparing the substrate specificity of MsAA with its homologue SmAA from *S. mobaraensis*, similarities exist in that SmAA also prefers shorter acyl residues. Regarding the fatty acyl chain length, SmAA prefers octanoic acid [7]. However, we found in the case of glutamine, that acetyl-glutamine is strongly preferred over octanoyl-glutamine, which was barely hydrolyzed, and lauroyl- and palmitoyl-glutamine were not accepted at all by MsAA. Furthermore, MsAA is specific to substrates acylated at the alpha-amino position, as  $N_\alpha$ -acetyl-lysine was hydrolyzed with 18.1 ( $\pm 0.6$ ) U/mg, whereas no activity against  $N_\epsilon$ -acetyl-lysine could be detected. The substrate specificity of MsAA is in approximate accordance with the substrate specificity of short acyl aminoacylase isolated from *M. smegmatis* ATCC 607 [19].

**Table 1** Effect on metal ions and chelating agent EDTA on hydrolytic activity. The enzyme purified without divalent cations added served as the control. The divalent metal salts or EDTA were added to this sample. The highest activity was observed when 1 mM  $ZnCl_2$  was added in all purification buffers and was set to 100%

Metal ion or chelating agent	Relative activity [%]	Specific activity [U/mg]
Control	51	65.2 ( $\pm 0.6$ )
1 mM EDTA	36	45.7 ( $\pm 3.6$ )
Purified with 1 mM $ZnCl_2$	100	127.2 ( $\pm 0.6$ )
Purified with 1 mM $ZnCl_2$ , 1 mM EDTA added	96	122.7 ( $\pm 1.6$ )
1 mM $ZnCl_2$	84	106.8 ( $\pm 8.6$ )
0.5 mM $ZnCl_2$	81	103.5 ( $\pm 6.6$ )
0.1 mM $ZnCl_2$	84	107.3 ( $\pm 5.8$ )
0.05 mM $ZnCl_2$	85	107.9 ( $\pm 6.2$ )
0.01 mM $ZnCl_2$	86	109.7 ( $\pm 4.0$ )
0.001 mM $ZnCl_2$	65	82.3 ( $\pm 7.0$ )
1 mM $MgSO_4$	47	60.3 ( $\pm 4.2$ )
1 mM $MnCl_2$	48	60.9 ( $\pm 4.4$ )
1 mM $CaCl_2$	50	63.8 ( $\pm 2.1$ )
1 mM $CoCl_2$	52	66.0 ( $\pm 8.7$ )
1 mM $NiSO_4$	43	54.4 ( $\pm 5.5$ )
1 mM $FeSO_4$	42	52.9 ( $\pm 1.1$ )
1 mM $CuSO_4$	40	51.2 ( $\pm 0.1$ )

**Table 2** Hydrolytic substrate specificity of MsAA. Substrates were used at 15 mM concentration in 100 mM Tris-HCl pH 7.0. Reaction temperature was elevated to 50 °C to ensure solubilization of all substrates

Substrate	Activity [U/mg]
Acetyl-alanine	265.9 (± 7.8)
Benzoyl-alanine	2.7 ± (0.1)
Lauroyl-alanine	20.6 (± 1.4)
Palmitoyl-alanine	20.7 (± 0.8)
Acetyl-methionine	140.9 (± 6.6)
Lauroyl-methionine	54.9 (± 0.7)
Acetyl-valine	326.1 (± 5.7)
Lauroyl-valine	7.1 (± 0.8)
Acetyl-isoleucine	284.2 (± 13.4)
Lauroyl-isoleucine	2.2 (± 0.3)
Acetyl-glutamine	44.6 (± 1.7)
Octanoyl-glutamine	0.8 (± 0.1)
Lauroyl-glutamine	0
Palmitoyl-glutamine	0
Acetyl-leucine	147.1 (± 6.8)
Lauroyl-leucine	0.5 (± 0.1)
Acetyl-threonine	63.5 (± 2.1)
Acetyl-aspartic acid	41.0 (± 0.6)
Lauroyl-aspartic acid	0
Acetyl-glycine	89.6 (± 1.8)
Lauroyl-glycine	2.7 (± 0.1)
Acetyl-glutamic acid	27.5 (± 1.4)
Acetyl-arginine	17.5 (± 0.6)
Acetyl-asparagine	20.3 (± 0.7)
Acetyl-cysteine	39.4 (± 2.5)
Lauroyl-cysteine	0
Acetyl-phenylalanine	3.4 (± 0.2)
Lauroyl-phenylalanine	0
Acetyl-tyrosine	1.2 (± 0.1)
Lauroyl-tyrosine	0
N <sub>α</sub> -acetyl-lysine	18.1 (± 0.6)
N <sub>ε</sub> -acetyl-lysine	0
Acetyl-tryptophane	0
Lauroyl-tryptophane	0
Acetyl-proline	0

#### Biocatalytic synthesis of lauroyl-amino acids in aqueous media

The hydrolytic substrate specificity of MsAA revealed that the aminoacylase strongly favors acyl-amino acids with short-chain acyl residues, while most lauroyl-amino acids were barely accepted as substrates under the chosen conditions. An exception was observed for the hydrolysis of lauroyl-methionine, which was hydrolyzed with almost half of the activity compared to hydrolysis of acetyl-methionine. To evaluate the synthetic potential of MsAA for lauroyl-amino acids, an excess of amino acids and lauric acid was deployed in an aqueous reaction system. Initial reaction conditions were chosen according to the results obtained from the hydrolytic reactions, namely

L-amino acid and lauric acid in equimolar concentration at 100 mM in 100 mM Tris-HCl pH 7.0 at 40 °C for 3 days in triplicates. The reaction mixtures were analyzed with HPLC-UV/-ELSD and product concentrations were measured with the respective lauroyl-amino acid standard. Among all proteinogenic amino acids, acylation was only observed for alanine, isoleucine, leucine, methionine, phenylalanine, and valine. The measured product concentrations were 1.1, 5.8, 5.1, 7.4, 0.1, and 4.5 mM, respectively. Hence, a bias for smaller, hydrophobic amino acids can be observed for the synthesis reaction. In accordance with the hydrolytic substrate specificity, lauroyl-methionine was found to be the preferred substrate in the synthesis reaction as well. To verify the identity of the synthesized molecule, the reaction mixture was analyzed with LC-MS, which confirmed the production of lauroyl-methionine (Additional File 1: Fig. S16).

#### Conclusions

In this work, a chaperone co-expression approach is described for the functional expression of a novel mycobacterial aminoacylase that is prone to inclusion body formation in *E. coli* and *V. natriegens*. We showed that *V. natriegens* was advantageous compared to *E. coli* when IPTG induction was used, as *V. natriegens* Vmax™ readily produced soluble recombinant aminoacylase. *V. natriegens* was superior to *E. coli* for short expression times of 4 h. We described chaperone co-expression in *V. natriegens* for the first time and showed that it can be used to optimize heterologous expression. In both *V. natriegens* and *E. coli*, co-expression of GroEL/S had the most pronounced influence on MsAA activity in the cell-free extract. However, *E. coli* BL21 (DE3) could only produce soluble MsAA by lactose autoinduction. Furthermore, we found that the use of *E. coli* ArcticExpress, which expresses Cpn60/10, combined with low cultivation temperatures, was best for expression of MsAA.

It has been described that the genetic elements of the Takara chaperone plasmid set, the p15A origin of replication and the *araBAD* promoter, function in *V. natriegens* [28]. We confirmed that the chaperone plasmids can be introduced to *V. natriegens* and that co-expression of the chaperones can influence the solubility of the protein of interest and hence activity obtained from recombinant expression. Since pGTf2 and pGKJE8 use the tetracycline inducible promoter *Pzt-1*, regulated by TetR, these genetic elements also seem to function in *V. natriegens*.

From all tested hosts, specific activity of purified MsAA was similar. We biochemically characterized the aminoacylase based on hydrolytic activities. Hydrolytic substrate specificity indicated favoring of short-chain acyl residues, except for lauroyl-methionine. Lastly, we demonstrated that the enzyme is suitable for the biocatalytic synthesis of N-lauroyl-L-methionine from lauric acid and

methionine in an aqueous system. The enzyme is thus highly interesting for further investigation and biocatalytic process intensification to optimize production of acyl-amino acids.

## Materials and methods

### Chemicals and reagents

Amino acids, cultivation media components, Tris (tris(hydroxymethyl)aminomethane), metal salts and solvents were from Carl Roth (Germany), acetyl amino acids were obtained from Sigma-Aldrich. Reagents for molecular biology were from Thermo Fisher Scientific (USA). DNA oligonucleotide synthesis and DNA sequencing were performed by Eurofins Genomics (Germany). Strep-Tactin columns were from IBA (Germany). Products for isoelectric focusing and native PAGE, as well as lysozyme were purchased from SERVA (Germany). Other chemicals were purchased from Sigma-Aldrich (USA). The EZ Nin reagent for amino acid quantification was from Biochrom (UK).

### Bacterial strains and plasmids

The microbial strains and plasmids used in this study are listed and described in Table 3. For cloning and plasmid maintenance, *E. coli* DH5 $\alpha$  was used. For expression experiments, various *E. coli* strains were used. *E. coli* BL21 was used for most expression experiments. As an alternative strain deficient of the *lac*-operon, *E. coli* Tuner (DE3) was used. Furthermore, *E. coli* ArcticExpress was used for expression at low temperatures. As an alternative expression host, *Vibrio natriegens* Vmax<sup>™</sup> was investigated. The co-expression with molecular chaperones was realized by co-transforming the strains with the chaperone plasmid set. Expression of MsAA was realized with pET28a MsAA NTag/CTag/noTag.

### Cloning of the aminoacylase gene from *Mycobacterium smegmatis* MKD 8

The *msAA* gene was commercially ordered and synthesized by GeneArt (Thermo Fisher Scientific, USA). The gene was codon-optimized for *E. coli* with added sequences coding for additional StrepII-tag (WSH-PQFEK) sequences at both ends, each connected with a linker (SG). The gene was amplified with primers carrying BsaI overhangs for Golden Gate cloning into pET28-eforRED, which derived from pET28a(+) (Novagen) [37]. Depending on the primer sequence, the aminoacylase gene could thus be cloned either with an N-terminal or a C-terminal Strep-tag, or without any affinity tag (Primers P1 & P4, P2 & P3 and P2 & P4, respectively; sequences for protein, gene and primers in Additional file 1: Table S1).

### Database searches and sequence analysis

Database searches were performed by using the BLASTp service from NCBI (National Center for Biotechnology Information) (<https://blast.ncbi.nlm.nih.gov/>) [38]. The alignment of multiple homologous sequences was conducted using the T-Coffee algorithm from EMBL-EBI (<https://www.ebi.ac.uk/Tools/msa/tcoffee/>) [39]. The alignment was displayed using ESPript 3.0 (<https://esprpt.ibcp.fr/>) [40]. ColabFold was used to predict a putative three-dimensional protein structure from the protein sequence of MsAA, based on the AlphaFold2 algorithm [41] and was used to add secondary structure elements to sequence alignments. For analysis of metal-binding sites and generation of a protein structure with added zinc ions, the MIB: Metal Ion-Binding Site Prediction and Docking Server was used [42].

**Table 3** Strains and plasmids used in this study

Strain/Plasmid	Genotype/Description	Source
<i>E. coli</i> DH5 $\alpha$	F <sup>-</sup> $\phi$ 80lacZ $\Delta$ M15 $\Delta$ ( <i>lacZYA-argF</i> )U169 <i>recA1 endA1 hsdR17</i> (r <sub>K</sub> <sup>-</sup> , m <sub>K</sub> <sup>+</sup> ) <i>phoA supE44 <math>\lambda</math><sup>-</sup> thi-1 gyrA96 relA1</i>	Thermo Fisher Scientific (USA)
<i>E. coli</i> BL21 (DE3)	F <sup>-</sup> <i>ompT gal dcm lon hsdS<sub>B</sub></i> (r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> ) $\lambda$ (DE3 [ <i>lacI lacUV5-T7p07 ind1 sam7 nin5</i> ]) [ <i>malB<sup>+</sup></i> ] <sub>K-12</sub> ( $\lambda$ <sup>S</sup> )	Thermo Fisher Scientific (USA)
<i>E. coli</i> Tuner (DE3)	F <sup>-</sup> <i>ompT hsdS<sub>B</sub></i> (r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> ) <i>gal dcm lacY1</i> (DE3)	Novagen <sup>®</sup> , Merck (Germany)
<i>E. coli</i> ArcticExpress	F <sup>-</sup> <i>ompT hsdS</i> (r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> ) <i>dcm<sup>+</sup> Tet'gal <math>\lambda</math></i> (DE3) <i>endA Hte [cpn10 cpn60 Gent<sup>r</sup>]</i>	Agilent Technologies (USA)
<i>Vibrio natriegens</i> Vmax <sup>™</sup>	<i>Vibrio natriegens</i> 14,048 <i>dns::LacI-T7-RNAP</i>	Telesis Bio (USA)
pET28a MsAA NTag/CTag/noTag	Derived from pET28a-eforRED[37], with coding sequence of MsAA with N- or C-terminal Strep-Tag, or without affinity tag	This study
pGKJE8	Arabinose-induced expression of <i>dnaK/J/grpE</i> ; tetracycline-induced expression of <i>groEL/S</i>	Takara Bio Europe (France)
pKJE7	Arabinose-induced expression of <i>dnaK/J/grpE</i>	
pGTf2	Tetracycline-induced expression of <i>dnaK/J/grpE</i> and <i>tig</i> (Trigger factor)	
pTf16	Arabinose-induced expression of <i>tig</i>	
pGro7	Arabinose-induced expression of <i>groEL/S</i>	

### Generation of chemically competent bacterial cells and transformation

For the preparation of chemically competent *E. coli* according to [43] with TMF buffer (100 mM CaCl<sub>2</sub>, 50 mM RbCl, 40 mM MnCl<sub>2</sub>) for final storage of the competent cells. The generation of chemically competent *V. natriegens* Vmax™ and subsequent transformation was conducted according to [28].

### Production of recombinant aminoacylase in *E. coli* or *V. natriegens* and purification

Various media and protocols were used for heterologous aminoacylase expression depending on the organisms and strains. Furthermore, the T7-based expression was either induced by addition of IPTG or by autoinduction with lactose. For *E. coli*, precultures were grown in 10 ml LB medium overnight at 37 °C supplemented with the necessary antibiotics. For *E. coli* BL21 (DE3) and Tuner (DE3), 100 ml Terrific Broth medium was used (TB; 2% tryptone from casein, 2.4% yeast extract, 25 mM NaH<sub>2</sub>PO<sub>4</sub>, 25 mM KH<sub>2</sub>PO<sub>4</sub>, 50 mM NH<sub>4</sub>Cl, 2 mM MgSO<sub>4</sub>, 5 mM Na<sub>2</sub>SO<sub>4</sub>, 0.5% glycerol (v/v) and 0.05% glucose). The expression cultures were inoculated with an OD<sub>600</sub> of 0.05. The cultures were induced with 1 mM IPTG once an OD<sub>600</sub> of 0.5 was reached. The autoinduction with lactose was realized by adding 0.2% lactose to the medium (TB-AIM). The cells were cultivated for 4 h upon induction with IPTG, while the bacteria were cultivated for 24 h after inoculation when using autoinduction. Temperatures during the expression were either 20 or 30 °C. *E. coli* ArcticExpress (DE3) was cultivated for 24 h at 12 °C in TB autoinduction medium, with prior growth for 6 h at 30 °C. Furthermore, no antibiotics were added in the expression cultures of *E. coli* ArcticExpress (DE3). For co-expression of chaperones with *E. coli*, 0.5 mg/ml arabinose and/or 5 ng/ml tetracycline were added.

For *V. natriegens*, additional v2 salts (204 mM NaCl, 4.2 mM KCl, 23.4 mM MgCl<sub>2</sub>) were added to the media. Pre-cultures were grown overnight at 37 °C in 10 ml LBv2 medium. For expression in *V. natriegens* Vmax™, BHlv2 medium (37 g/l Brain-Heart-Infusion Broth, Carl Roth, Germany; with v2 salts) or TBv2 medium was used. Like for *E. coli*, the cultures were inoculated to an OD<sub>600</sub> of 0.05, induced with 1 mM IPTG at OD<sub>600</sub> of 0.5 and then cultured for 4 h at 30 °C (if not stated otherwise). For co-expression of chaperones, 4 mg/ml arabinose and/or 10 ng/ml tetracycline were added.

The cells were harvested by centrifugation at 3000 g and 4 °C for 40 min. To each cell pellet, harvested from 50 ml expression culture, 10 ml of lysis buffer was added (100 mM Tris-HCl pH 7.0 supplemented with 0.1% Triton X-100, 1 mM ZnCl<sub>2</sub>, 0.3 mg ml<sup>-1</sup> lysozyme and 150 mM NaCl) and the cells were disrupted by sonication

(sonotrode MS 73, Bandelin, Germany). After lysis, the samples were centrifuged at 16,000 g and 4 °C for 60 min. The supernatant containing soluble protein was separated and the insoluble pellet was resuspended in 5 ml of 5 M urea for SDS-PAGE analysis.

Strep-tag affinity purification of recombinant MsAA NTag from the soluble fraction was conducted using a 5 ml Strep-Tactin® SuperFlow® high capacity cartridge according to manufacturer's instructions. The buffer system consisted of 100 mM Tris-HCl pH 7.0, 1 mM ZnCl<sub>2</sub> and 150 mM NaCl with 2.5 mM desthiobiotin for the elution buffer. The fractions containing sufficient enzyme concentrations were pooled and rebuffered to the same buffer without desthiobiotin in Vivaspin™ 6 concentrators (10000 MWCO; Sartorius, Germany).

### Determination of protein concentration, purity, and molecular mass

Protein concentrations were determined with the method of Bradford [44] using the Roti®-Nanoquant reagent (Carl Roth). Bovine serum albumin (BSA) served as a standard. The SDS-polyacrylamide gel electrophoresis of proteins was performed according to Laemmli [45] by using 8–20% gradient gels and staining with Roti®Blue quick (Carl Roth). As a protein marker, FastGene® BlueE-asy Protein Marker (Nippon Genetics) was used.

Native molecular weight was determined by performing blue native PAGE using reagents from SERVA (Germany). SERVAGel™ N4-16% gels were used for electrophoresis, conducted as described by manufacturer's protocols. The isoelectric point of the purified, desalinated enzyme was determined by isoelectric focusing (IEF) using the Servalyt™ Precotes™ gel and the IEF marker pH 3–10 by SERVA (Germany) according to manufacturer's protocols. Desalination of the purified protein was conducted by buffer exchange to 10 mM Tris-HCl pH 7.0 using Spin Columns (3 kDa cut-off, VWR).

Matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF/MS) analysis was used to determine the molecular mass of MsAA NTag. The device used for analysis was Axima Confidence (Shimadzu Europe, Duisburg, Germany), which was operated in linear positive mode with pulsed extraction optimized for the theoretical molecular weight. The mMass software was used for data analysis [46]. Protein solutions were diluted to 1 mg/ml in enzyme storage buffer. The samples were then diluted tenfold with α-Cyano-4-hydroxycinnamic acid (CHCA) (Sigma-Aldrich, USA). Per spot on the MALDI target plate, 2 μl of the sample were applied. Trypsinogen and BSA (Laserbio, France) were used as molecular weight standards.

### Aminoacylase activity assay

Activity of aminoacylases was assayed by quantification of released amino acids with a ninhydrin-based assay as previously described [47]. Briefly, 10  $\mu$ l sample from amino acid solutions or aminoacylase reactions were mixed with 100  $\mu$ l of EZ Nin:DMSO reagent, heated for 10 min at 99 °C and diluted with 100 mM Na-borate buffer pH 10.0 for measurement. In general, 200  $\mu$ l reactions consisted of 190  $\mu$ l substrate solution and 10  $\mu$ l enzyme solution. For standard hydrolysis activity measurement, reaction with 15 mM N-acetyl-L-alanine in 100 mM Tris-HCl buffer pH 7.0, 50  $\mu$ M ZnCl<sub>2</sub>, were performed at 30 °C for 5 min. At 1 min sampling intervals, 10  $\mu$ l samples were withdrawn for the ninhydrin reaction. One unit of MsAA was defined as the amount of enzyme that hydrolyzes one  $\mu$ mol of N-acetyl-L-alanine per minute under the given conditions.

### Biochemical properties of aminoacylase

For the determination of the pH dependency of the hydrolytic reaction rate, reactions were carried out in the following buffers: Na-acetate for pH 4.0 and 5.0, Tris-HCl for pH 6.0–9.0 and Na-borate for pH 9.0–13.0. As substrate solution, 15 mM N-acetyl-L-alanine was prepared in 100 mM of the respective buffer and adjusted to respective pH at 30 °C. The purified enzyme concentrations in the assay were 10–100  $\mu$ g/ml.

For investigating pH stability, buffers used were Na-acetate for pH 4.0 and 5.0, Tris-HCl for pH 6.0–9.0 and Na-borate for pH 9.0–13.0, all at concentrations of 100 mM. The enzyme solution was diluted 10-fold with the incubation buffers to a concentration of 200  $\mu$ g/ml and incubated at 30 °C. From the incubation solutions, 10  $\mu$ l was withdrawn for standard aminoacylase activity measurements after 24 h and 3 days. The final enzyme concentration was 10  $\mu$ g/ml.

The optimal temperature for the hydrolysis was determined in a range of 20–90 °C with standard aminoacylase activity. The pH of the solutions was set at the respective temperatures. For the assessment of temperature stability, purified MsAA was incubated at temperatures from 20 to 90 °C and residual activity was determined after 1 h, 24 h, 5 days and 7 days. As incubation buffers, 100 mM Tris-HCl was set to pH 7.0 at the respective temperature. The purified enzyme was incubated at 380  $\mu$ g/ml and 1 mM ZnCl<sub>2</sub>. Residual activity was determined with standard aminoacylase assay.

Recombinant MsAA was purified without any metal ions added to the buffers to investigate the effect of metal ions or chelating agents added to the purified enzyme. Various bivalent metal ions (CaCl<sub>2</sub>, CoCl<sub>2</sub>, CuCl<sub>2</sub>, FeSO<sub>4</sub>, MgCl<sub>2</sub>, MnCl<sub>2</sub>, NiCl<sub>2</sub>, ZnCl<sub>2</sub>) were added to a concentration of 1 mM to the enzyme solution and incubated for 1 h at room temperature before measuring standard

aminoacylase activity. For ZnCl<sub>2</sub>, concentrations were set from 1  $\mu$ M to 5 mM. The influence of 1 mM ethylenediaminetetraacetic acid (EDTA) as a chelating agent of bivalent ions was also investigated and residual activity was measured after 1 h incubation at room temperature.

The substrate specificity was determined by hydrolysis of various substrates at 15 mM concentration in 100 mM Tris-HCl pH 7.0 at 50 °C and 50  $\mu$ M ZnCl<sub>2</sub>. The reaction temperature was higher than the standard conditions to solubilize all substrates. The concentrations of purified enzyme were 5–100  $\mu$ g/ml.

### Determination of thermal denaturation

With the thermal shift assay, thermal denaturation of the investigated protein was assessed as previously described [48]. The fluorogenic dye Sypro orange was mixed with the protein and heated stepwise. When the protein is folded correctly, the dye does not bind efficiently on the hydrophilic surface. When hydrophobic stretches are exposed due to denaturation of the protein, binding of the dye leads to an increase in fluorescence. For the assay, 10  $\mu$ l protein samples (>0.1 mg/ml) are mixed with 5  $\mu$ l 50x SYPRO Orange (Sigma Aldrich) and 20  $\mu$ l 10 mM HEPES pH 8.0. As a positive control, 10 mg/ml lysozyme (Serva) was used. Measurement was done with qTower3G and qPCRsoft 4.0 (Analytik Jena) using the TAMRA Channel ( $\lambda_{ex}$ =535 nm,  $\lambda_{em}$ =580 nm). The heating program was 25 to 95 °C with steps of 2 °C, 120 s hold time per temperature and a heating speed of 4.4 °C/s.

### Chemical synthesis of N-acyl-amino acids

N-lauroyl-, N-palmitoyl- and other N-acyl-amino acids were synthesized by the Schotten-Baumann-reaction following the protocols of Takehara [49]. Amino acids (70 mmol) and NaOH (1.40 g, 70 mmol, 1.0 eq) were dissolved in 49 ml H<sub>2</sub>O and 35 ml acetone. With glutamic or aspartic acid as substrates, 2.80 g NaOH (140 mmol, 2.0 eq.) were used for the initial dissolution in the water/acetone mixture. To the cooled mixtures (4 °C, 2000 rpm) 4.2 g (105 mmol, 1.5 eq) NaOH in 15 ml water and acyl chloride (84 mmol, 1.1 eq.) were added dropwise over a period of 30 min. The reactions were run for a total time of 4 h and allowed to warm to room temperature. Acidification with 12 N hydrochloric acid led to the precipitation of N-acyl-amino acids. The crude products were separated by filtration, carefully washed with water (40 ml) and washed twice with 100 ml of petroleum ether (bp=40–60 °C). Remaining solvent was removed *in vacuo* and the white products were dried by lyophilization afterwards. The products were analyzed by HPLC and LC-MS. Purity data and HPLC retention times of the N-acyl-amino acids are provided in Additional file 1: Table S2.

### Biocatalytic synthesis of lauroyl-amino acids

Initial biocatalytic synthesis of lauroyl-amino acids was investigated from 100 mM L-amino acid and 100 mM lauric acid in 100 mM Tris-HCl pH 7.0 at 40 °C for 72 h at a reaction volume of 0.5 ml without agitation. Reactions were started by adding 10 µg of MsAA (1.3 U, according to standard activity assay). For analysis, a 100 µl sample was withdrawn and immediately mixed with 100 µl of a mixture of 80% acetonitrile and 20% water containing 0.1% trifluoroacetic acid (TFA).

### HPLC-UV/ELSD analysis

Analysis of biocatalytic reactions to investigate the synthesis of lauroyl-amino acids and to quantify product concentration, an HPLC system (S5200 and S2100; Sykam, Germany) equipped with a ISAspher 100-5 C18 BDS column (C18, 5 µm, 4.0 \* 250 mm; Isera, Germany) coupled with UV (UV Detector 2500 Sykam, Germany)- and evaporative light scattering detectors (ELSD; ZAM 4000, Schambeck SFD, Germany) was used. The column was heated to 40 °C and isocratic separation was performed with a flow rate of 1 ml/min and a solvent mixture of 80% acetonitrile and 20% water containing 0.1 TFA. Concentration of lauroyl-amino acids was calculated with a respective external standard, analyzed by both UV (210 nm) and ELSD. As values obtained by both detectors were in good accordance, only concentrations obtained from UV measurement are shown.

### HPLC-MS analysis

A Shimadzu Nexera XR system equipped with a Hitachi LaChrom II column (C18, 5 µm, 4.6 \* 250 mm) and a Shimadzu LCMS-2020-mass spectrometer was used to verify the correct mass of the product N-lauroyl-methionine from biocatalytic synthesis. The column was operated at 40 °C with a flow of 1 ml/min and the analytes were separated by applying a gradient run starting from a mixture of 20% acetonitrile and 80% water containing 0.1% formic acid, going to 100% acetonitrile in 10 min. The concentration was held for 6 min before returning to 20% acetonitrile over 2 min.

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12934-023-02079-1>.

Additional file 1: MsAA protein sequence, primer sequences, gel electrophoresis results, chemically synthesized acyl-amino acids, MS spectrum for lauroyl-methionine.

### Author contributions

GH designed the study. GH conducted cloning and bioinformatic analysis. GH, JW, NT, SO performed the experiments and analyzed the data for protein expression, purification, and biochemical characterization. TJ synthesized acyl-amino acids chemically and performed HPLC-MS analysis. GH wrote the manuscript. GH, US, JB, KEJ and PS edited the manuscript. PS and JB

supervised the work of GH, JW, NT and SO and US supervised the work of TJ. PS, JB and US did funding acquisition. All authors read and approved the final manuscript.

### Funding

Open Access funding enabled and organized by Projekt DEAL.

### Data Availability

The datasets supporting the conclusions of this article are included within the article and its additional files.

### Declarations

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

All authors have approved the manuscript.

#### Competing interests

The authors declare no competing interests.

#### Author details

<sup>1</sup>Institute of Nano- and Biotechnologies, Aachen University of Applied Sciences, 52428 Jülich, Germany

<sup>2</sup>TH Köln - Campus Leverkusen, 51379 Leverkusen, Germany

<sup>3</sup>Institute of Molecular Enzyme Technology, Heinrich Heine University Düsseldorf, 52425 Jülich, Germany

<sup>4</sup>Institute of Bio- and Geosciences IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, Jülich 52425, Germany

Received: 28 February 2023 / Accepted: 4 April 2023

Published online: 21 April 2023

### References

- Liu Z, Zhen Z, Zuo Z, Wu Y, Liu A, Yi Q, Li W. Probing the catalytic center of porcine aminoacylase 1 by site-directed mutagenesis, homology modeling and substrate docking. *J Biochem.* 2006;139:421–30. <https://doi.org/10.1093/jb/mvj047>.
- Lindner HA, Alary A, Boju LI, Sulea T, Ménard R. Roles of dimerization domain residues in binding and catalysis by aminoacylase-1. *Biochemistry.* 2005;44:15645–51. <https://doi.org/10.1021/bi051180y>.
- Wardenga R, Hollmann F, Thum O, Bornscheuer U. Functional expression of porcine aminoacylase 1 in *E. coli* using a codon optimized synthetic gene and molecular chaperones. *Appl Microbiol Biotechnol.* 2008;81:721–9. <https://doi.org/10.1007/s00253-008-1716-7>.
- Wada E, Handa M, Imamura K, Sakiyama T, Adachib S, Matsunob R, Nakanishia K. Enzymatic synthesis of N-acyl-L-amino acids in a glycerol-water system using acylase I from pig kidney. *JAOCS.* 2002;79:41–6. <https://doi.org/10.1007/s11746-002-0432-7>.
- Wardenga R, Lindner HA, Hollmann F, Thum O, Bornscheuer U. Increasing the synthesis/hydrolysis ratio of aminoacylase 1 by site-directed mutagenesis. *Biochimie.* 2010;92:102–9. <https://doi.org/10.1016/j.biochi.2009.09.017>.
- Yokoigawa K, Sato E, Esaki N, Soda K. Enantioselective synthesis of N-acetyl-L-methionine with aminoacylase in organic solvent. *Appl Microbiol Biotechnol.* 1994;287–9. <https://doi.org/10.1007/BF00902730>.
- Cho H-Y, Tanizawa K, Tanaka H, Soda K. Thermostable aminoacylase from *Bacillus thermoglucosidius*: purification and characterization. *Agric Biol Chem.* 1987;51:2793–800.
- Sakanyan V, Desmarez L, Legrain C, Charlier D, Mett I, Kochikyan A, et al. Gene cloning, sequence analysis, purification, and characterization of a thermostable aminoacylase from *Bacillus stearothermophilus*. *Appl Environ Microbiol.* 1993;59:3878–88.
- Takakura Y, Asano Y. Purification, characterization, and gene cloning of a novel aminoacylase from *Burkholderia* sp. strain LP5\_188 that efficiently catalyzes the synthesis of N-lauroyl-L-amino acids. *Biosci Biotechnol Biochem.* 2019;83:1964–73.

10. Shintani Y, Fukuda H, Okamoto N, Murata K, Kimura A. Isolation and characterization of N-long chain acyl aminoacylase from *Pseudomonas diminuta*. *Biochem*. 1984;637–43.
11. Koreishi M, Kawasaki R, Imanaka H, Imamura K, Nakanishi K. A novel  $\epsilon$ -lysine acylase from *Streptomyces mobaraensis* for synthesis of  $N_{\epsilon}$ -acyl-L-lysines. *JAOCs*. 2005;82:631–7. <https://doi.org/10.1007/s11746-005-1121-2>.
12. Koreishi M, Nakatani Y, Ooi M, Imanaka H, Imamura K, Nakanishi K. Purification, characterization, molecular cloning, and expression of a new aminoacylase from *Streptomyces mobaraensis* that can hydrolyze N-(middle/long)-chain-fatty-acyl-L-amino acids as well as N-short-chain-acyl-L-amino acids. *Biosci Biotechnol Biochem*. 2009;73:1940–7. <https://doi.org/10.1271/bbb.90081>.
13. Koreishi M, Asayama F, Imanaka H, Imamura K, Kadota M, Tsuno T, Nakanishi K. Purification and characterization of a novel aminoacylase from *Streptomyces mobaraensis*. *Biosci Biotechnol Biochem*. 2005;69:1914–22. <https://doi.org/10.1271/bbb.69.1914>.
14. Bourkaib MC, Delaunay S, Framboisier X, Hôtel L, Aigle B, Humeau C, et al. N-acylation of L-amino acids in aqueous media: evaluation of the catalytic performances of *Streptomyces ambofaciens* aminoacylases. *Enzyme Microb Technol*. 2020;137:109536. <https://doi.org/10.1016/j.enzmictec.2020.109536>.
15. Koreishi M, Kawasaki R, Imanaka H, Imamura K, Takakura Y, Nakanishi K. Efficient  $N_{\epsilon}$ -lauroyl-L-lysine production by recombinant  $\epsilon$ -lysine acylase from *Streptomyces mobaraensis*. *J Biotechnol*. 2009;141:160–5. <https://doi.org/10.1016/j.jbiotec.2009.03.008>.
16. Dettori L, Ferrari F, Framboisier X, Paris C, Guivarc'h Y, Hôtel L, et al. An aminoacylase activity from *Streptomyces ambofaciens* catalyzes the acylation of lysine on  $\alpha$ -position and peptides on N-terminal position. *Eng Life Sci*. 2018;18:589–99. <https://doi.org/10.1002/elsc.201700173>.
17. Bourkaib MC, Delaunay S, Framboisier X, Humeau C, Guilbot J, Bize C, et al. Enzymatic synthesis of N-10-undecenoyl-phenylalanine catalysed by aminoacylases from *Streptomyces ambofaciens*. *Process Biochem*. 2020;99:307–15. <https://doi.org/10.1016/j.procbio.2020.09.009>.
18. Nagai S, Matsuno J. Enzyme hydrolyzing N-long chain acyl-L-aspartic acids from *Mycobacterium smegmatis*: purification and specificity of the enzyme, and the effect of alkaline metal ions on its activity. *J Biochem*. 1964;56:465–76. <https://doi.org/10.1093/oxfordjournals/jbchem.a128018>.
19. Matsuno J, Nagai S. Amidohydrolases for N-short and long chain acyl-L-amino acids from *Mycobacteria*. *J Biochem*. 1972;269–79. <https://doi.org/10.1093/oxfordjournals/jbchem.a129906>.
20. Scherr N, Nguyen L. *Mycobacterium* versus *Streptomyces* we are different, we are the same. *Curr Opin Microbiol*. 2009;12:699–707. <https://doi.org/10.1016/j.mib.2009.10.003>.
21. Sørensen HP. Towards universal systems for recombinant gene expression. *Microb Cell Fact*. 2010;9:27. <https://doi.org/10.1186/1475-2859-9-27>.
22. Rosano GL, Ceccarelli EA. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front Microbiol*. 2014;5:172. <https://doi.org/10.3389/fmicb.2014.00172>.
23. Turner P, Holst O, Karlsson EN. Optimized expression of soluble cyclomalto-dextrinase of thermophilic origin in *Escherichia coli* by using a soluble fusion-tag and by tuning of inducer concentration. *Protein Expr Purif*. 2005;39:54–60. <https://doi.org/10.1016/j.pep.2004.09.012>.
24. Studier FW. Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif*. 2005;41:207–34. <https://doi.org/10.1016/j.pep.2005.01.016>.
25. Hoffmann F, Rinas U. Roles of heat-shock chaperones in the production of recombinant proteins in *Escherichia coli*. *Adv Biochem Eng Biotechnol*. 2004;89:143–61. <https://doi.org/10.1007/b93996>.
26. Anja Hoffmann B, Bukau G. Structure and function of the molecular chaperone trigger factor. *Biochim Biophys Acta*. 2010;650–61. [https://doi.org/10.1016/0006-3002\(62\)90265-2](https://doi.org/10.1016/0006-3002(62)90265-2).
27. Sakikawa C, Taguchi H, Makino Y, Yoshida M. On the maximum size of proteins to stay and fold in the cavity of GroEL underneath GroES. *J Biol Chem*. 1999;274:21251–6. <https://doi.org/10.1074/jbc.274.30.21251>.
28. Weinstock MT, Hesk ED, Wilson CM, Gibson DG. *Vibrio natriegens* as a fast-growing host for molecular biology. *Nat Methods*. 2016;13:849–51. <https://doi.org/10.1038/nmeth.3970>.
29. Xu J, Dong F, Wu M, Tao R, Yang J, Wu M, et al. *Vibrio natriegens* as a pET-compatible expression host complementary to *Escherichia coli*. *Front Microbiol*. 2021;12:627181. <https://doi.org/10.3389/fmicb.2021.627181>.
30. Biagini A, Puigserver A. Sequence analysis of the aminoacylase-1 family. A new proposed signature for metalloexopeptidases. *Comp Biochem Physiol - B Biochem Mol Biol*. 2001;469–81.
31. Nocek B, Reidl C, Starus A, Heath T, Bienvenue D, Osipiuk J, et al. Structural evidence of a major conformational change triggered by substrate binding in DapE enzymes: impact on the catalytic mechanism. *Biochemistry*. 2018;57:4–84. <https://doi.org/10.1021/acs.biochem.7b01151.s001>.
32. Rawlings ND, Salvesen G, editors. *Handbook of proteolytic enzymes*. Academic Press; 2013.
33. Ferrer M, Chernikova TN, Yakimov MM, Golyshin PN, Timmis K. N. Chaperonins govern growth of *Escherichia coli* at low temperatures. *Nat Biotechnol*. 2003;21:66–7. <https://doi.org/10.1038/nbt1103-1266>.
34. Eberhardt F, Aguirre A, Menzella HG, Peiru S. Strain engineering and process optimization for enhancing the production of a thermostable steryl glucosidase in *Escherichia coli*. *J Ind Microbiol Biotechnol*. 2017;44:141–7. <https://doi.org/10.1007/s10295-016-1866-z>.
35. Tong Y, Feng S, Xin Y, Yang H, Zhang L, Wang W, Chen W. Enhancement of soluble expression of codon-optimized *Thermomicrobium roseum* sarcosine oxidase in *Escherichia coli* via chaperone co-expression. *J Biotechnol*. 2016;218:75–84. <https://doi.org/10.1016/j.jbiotec.2015.11.018>.
36. Lu X, He S, Zong H, Song J, Chen W, Zhuge B. Improved 1, 2, 4-butanetriol production from an engineered *Escherichia coli* by co-expression of different chaperone proteins. *World J Microbiol Biotechnol*. 2016;32:149. <https://doi.org/10.1007/s11274-016-2085-5>.
37. Muschallik L, Molinnus D, Jablonski M, Kipp CR, Bongaerts J, Pohl M, et al. Synthesis of  $\alpha$ -hydroxy ketones and vicinal (*R,R*)-diols by *Bacillus clausii* DSM 8716<sup>T</sup> butanediol dehydrogenase. *RSC Adv*. 2020;10:12206–16. <https://doi.org/10.1039/d0ra02066d>.
38. Johnson M, Zaretskaya I, Raytselis Y, Merezukh Y, McGinnis S, Madden TL. NCBI BLAST: a better web interface. *Nucleic Acids Res*. 2008;36:W5–9. <https://doi.org/10.1093/nar/gkn201>.
39. Notredame C, Higgins DG, Heringa J. T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J Mol Biol*. 2000;302:205–17. <https://doi.org/10.1006/jmbi.2000.4042>.
40. Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res*. 2014;42:W320–4. <https://doi.org/10.1093/nar/gku316>.
41. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. *Nat Methods*. 2022. <https://doi.org/10.1038/s41592-022-01488-1>.
42. Lin Y-F, Cheng C-W, Shih C-S, Hwang J-K, Yu C-S, Lu C-H. Metal ion-binding site prediction and docking server. *J Chem Inf Model*. 2016;56:2287–91. <https://doi.org/10.1021/acs.jcim.6b00407>.
43. Hanahan D. Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol*. 1983;166:557–80. [https://doi.org/10.1016/s0022-2836\(83\)80284-8](https://doi.org/10.1016/s0022-2836(83)80284-8).
44. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem*. 1976;72:248–54. <https://doi.org/10.1006/abio.1976.9999>.
45. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. 1970;227:680–5. <https://doi.org/10.1038/227680a0>.
46. Strohal M, Hassman M, Kosata B, Kodicek M. Rapid Commun Mass Spectrom. 2008;22:905–8. <https://doi.org/10.1002/rcm.3444>. mMass data miner: an open source alternative for mass spectrometric data analysis.
47. Haeger G, Bongaerts J, Siebert P. A convenient ninhydrin assay in 96-well format for amino acid-releasing enzymes using an air-stable reagent. *Anal Biochem*. 2022;654:114819. <https://doi.org/10.1016/j.ab.2022.114819>.
48. Falkenberg F, Rahba J, Fischer D, Bott M, Bongaerts J, Siebert P. Biochemical characterization of a novel oxidatively stable, halotolerant, and high-alkaline subtilisin from *Alkalihalobacillus okhensis* Kh10-101<sup>T</sup>. *FEBS Open Bio*. 2022. <https://doi.org/10.1002/2211-5463.13457>.
49. Takehara M, Yoshimura I, Takizawa K, Yoshida R. Surface active N-acylglutamate: I. Preparation of long chain N-acylglutamic acid. *J Am Oil Chem Soc*. 1972;49:157. <https://doi.org/10.1007/BF02633785>.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## 2.4. Chapter IV

### **Biocatalytic potential of MsAA aminoacylase for synthesis of N-acyl-L-amino acids in aqueous media**

Gerrit Haeger, Jessika Wirges, Laureline Genesseeux, Yann Guiavarc'h, Catherine Humeau, Cédric Paris, Isabelle Chevalot, Karl-Erich Jaeger, Johannes Bongaerts, Petra Siegert

Manuscript.

Author contributions:

GH designed the study and performed biocatalytic reactions and molecular docking. JW performed the biocatalytic experiments. CP performed and analyzed HPLC-MS analysis. CH supervised and LG supported the molecular docking experiments. YG, IC, and PS supervised the biocatalytic experiments. GH wrote the manuscript. GH, YG, CP, CH, IC, JB, KEJ and PS edited the manuscript. PS and JB did funding acquisition. All authors read and approved the final manuscript.

Overall contribution GH: 85 %

## Biocatalytic Potential of Mycobacterial Aminoacylase for Synthesis of *N*-Acyl-L-Amino Acids in Aqueous Media

Gerrit Haeger<sup>a</sup>, Jessika Wirges<sup>a</sup>, Laureline Gennesseaux<sup>b</sup>, Yann Guiavarc'h<sup>b</sup>, Catherine Humeau<sup>b</sup>, Cédric Paris<sup>c</sup>, Isabelle Chevalot<sup>b</sup>, Karl-Erich Jaeger<sup>d,e</sup>, Johannes Bongaerts<sup>a</sup>, Petra Siegert<sup>a</sup>

<sup>a</sup>Institute of Nano- and Biotechnologies, Aachen University of Applied Sciences, D-52428 Jülich, Germany

<sup>b</sup>Université de Lorraine, CNRS, Laboratoire Réactions et Génie des Procédés (LRGP), F-54000, Nancy, France

<sup>c</sup>Université de Lorraine, CNRS, Laboratoire d'Ingénierie des Biomolécules (LIBio), F-54000, Nancy, France

<sup>d</sup>Institute of Molecular Enzyme Technology, Heinrich Heine University Düsseldorf, D-52425 Jülich, Germany

<sup>e</sup>Institute of Bio- and Geosciences IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, D-52425 Jülich, Germany

Keywords:

Acylation, amino acids, aminoacylase, biocatalysis, biosurfactants, docking

### Abstract

In this study, we present an investigation of the recently identified aminoacylase MsAA for the synthesis of *N*-acyl-L-amino acids, focusing on lauroyl-methionine. We found optimal reaction conditions at pH 8.0 and a temperature of 40-45 °C with substrate concentrations of 400 mM methionine and 150 mM lauric acid. The highest product concentration of 100 mM means 67 % substrate conversion. The reaction could be upscaled with a nearly identical reaction course. Several other fatty acids were also accepted by the enzyme. Besides methionine, only hydrophobic amino acids were accepted for acylation. We performed a two-factorial design of experiments (DoE), varying the temperature along with various glycerol contents, which revealed a positive influence of glycerol on the product formation, especially at higher temperatures. For a detailed analysis of the catalytic reaction mechanism of MsAA, we performed *in silico* protein modeling studies. Molecular docking of lauric acid and methionine to the predicted MsAA structure resulted in a similar mode of substrate binding as described for the related *N*-succinyl-L,L-diaminopimelic acid desuccinylase from *Haemophilus influenzae*. Differences in amino acid sequence of structurally conserved substrate-binding residues explain the distinct substrate scope of the enzymes.

## Introduction

Acyl-amino acids are valuable compounds for cosmetic products. They consist of an amino acid linked to a fatty acid, most commonly through *N*-acylation under the formation of an amide bond, but also through *O*-acylation by ester formation. Most commonly lauric acid is used, or fatty acids derived from coconut oil, which is rich in lauric acid. Amino acid-based surfactants are considered as remarkably skin-friendly due to their low inflammatory potential. In high-grade cosmetic products, *N*-lauroyl-L-glutamic acid is commonly used as a surfactant<sup>[1]</sup>.

Even though the products are of plant or other biological origin, and can thus be classified as biosurfactants, chemical synthesis of these compounds is typically achieved by the environmentally harmful Schotten-Baumann synthesis. For this acylation method, acyl chlorides are necessary, rendering the synthesis not compliant with green chemistry principles<sup>[2]</sup>. First, fatty acid chlorides must be synthesized, which can be done using phosgene, thionyl chloride, or phosphoryl trichloride. All of these are especially harmful chemicals. The fatty acyl chlorides are particularly reactive and can hydrolyze, resulting in the formation of hydrogen chloride, which poses an additional hazard. Furthermore, during the acylation reaction, sodium chloride is stoichiometrically produced as waste.

Biocatalytic synthesis using enzymes is therefore a promising alternative because it can be performed with unmodified, free fatty acids. This eliminates the need for chlorination and prevents the formation of waste products. Aminoacylases have been described and investigated for this purpose but are widely unexplored and underexploited for commercial applications. They catalyze the reversible hydrolysis of *N*-acylated amino acids, mainly proteinogenic L-amino acids at the  $\alpha$ -position.  $\epsilon$ -Lysine aminoacylases have been described as well<sup>[3]</sup>. Aminoacylases have been identified from various organisms<sup>[4]</sup>, and especially in *Streptomyces*. A variety of aminoacylases have been shown to be applicable for biocatalytic synthesis<sup>[5,6,7]</sup>. Recently, the broad-spectrum  $\alpha$ -aminoacylase SamAA from crude extract of *S. ambifaciens* has been used for acylation of a variety of amino acids, with conversion rates reaching 30 % for  $\alpha$ -lauroyl-lysine<sup>[6]</sup>. The acylation reaction with SamAA has been optimized for 10-undecenoyl-phenylalanine, and a strong influence of pH and substrate concentration was determined, with optimal conditions at pH 8.0 and 100 mM equimolar substrate concentration. However, there are a number of obstacles to an economical process, in particular the difficulty of recombinant expression of the enzymes and their low stability<sup>[7]</sup>. Recently, we identified the mycobacterial aminoacylase MsAA from *Mycobacterium smegmatis* MKD 8, and optimized its expression in *E. coli* and *V. natriegens*<sup>[8]</sup>. In a first screening for acylation of proteinogenic L-amino acids with MsAA, lauroyl-methionine, lauroyl-isoleucine, lauroyl-leucine, lauroyl-valine, lauroyl-alanine, and lauroyl-phenylalanine were synthesized with product concentrations of 7.4 mM, 5.8 mM, 5.1 mM, 4.5 mM, 1.1 mM, and 0.1 mM, respectively<sup>[8]</sup>.

The aminoacylase MsAA belongs to the M20A family of metallopeptidases, characterized by a cocatalytic active site, most often with two zinc ions, ligated by histidine and glutamic acid residues, with an aspartic acid residue bridging the zinc ions<sup>[9]</sup>. Some enzymes from this family have been

thoroughly characterized and their crystal structure have been published. A homodimeric structure has been found for *N*-succinyl-L,L-diaminopimelic acid desuccinylase HiDapE from *Haemophilus influenzae*<sup>[10]</sup> and carboxypeptidase G2 from *Pseudomonas* sp.<sup>[11]</sup>. The structural studies enabled not only the proposal of a catalytic mechanism for hydrolysis, but also revealed dynamic movement of the dimeric units during substrate binding and release. The crystal structure with bound substrates of HiDapE has been solved<sup>[12]</sup>. The tertiary structure can be divided into a catalytic domain and a dimerization domain linked by a flexible hinge, which closes and opens with substrate binding and release, respectively. Closing the structure brings the dimers closer together at each catalytic site, playing a role in catalytic mechanism and substrate binding. During hydrolysis, the non-metal binding, catalytic glutamate residue acts as a general base and can deprotonate a zinc-bound water molecule. The formed hydroxide ion can attack the substrate's carbonyl carbon atoms, leading to a tetrahedral intermediate. This negatively charged transition state is stabilized by a histidine from the opposing dimer, forming an oxyanion hole<sup>[12]</sup>. This suggests that the monomeric protein is not catalytically active. Upon decomposition to the hydrolyzed products, the substrate is released.

MsAA was recombinantly expressed in *E. coli*, initially characterized, and first used in biocatalytic synthesis with the highest production yield of lauroyl-methionine. A conversion of 7.4 % after 72 h starting from 100 mM lauric acid and 100 mM methionine in 100 mM Tris-HCl at pH 7.0 and 40 °C<sup>[8]</sup> was initially achieved. Here, we present an optimization of the acylation conditions resulting in a final concentration of 100 mM lauroyl-methionine and 67 % substrate conversion. The optimization of reaction parameters resulted in a 13.5-fold increase of product concentration and a 9.0-fold increase in conversion rate of the substrate lauric acid.

To gain further insight into the biocatalytic acylation of methionine on a molecular level, analysis of the protein structure of MsAA predicted by AlphaFold<sup>[8,13]</sup>, and docking experiments were conducted. Comparison to HiDapE revealed several similarities, such as domain architecture and relative positioning of the active site. Therefore, we performed molecular docking of lauric acid and methionine. The docking results show major differences in substrate-binding residues that can explain the distinctive substrate scope of these enzymes.

## Results and Discussion

### Enzyme production, hydrolytic and synthetic activity

The aminoacylase MsAA was recombinantly produced in *E. coli* and purified to homogeneity as described previously<sup>[8]</sup>. Hydrolytic activity against 15 mM acetyl-alanine was 130 U/mg (standard activity assay, in 100 mM Tris-HCl pH 7.0, 30 °C). As described in a previous study, hydrolytic pH optimum was at pH 7.0 and the enzyme was stable up to 40 °C. These conditions were chosen to initially investigate the catalytic potential of acylation of L-amino acids. As a screening to probe the acceptance of proteinogenic amino acids, equimolar substrate concentrations of 100 mM lauric acid and amino acid

were used. Only small, hydrophobic amino acids were accepted. Lauroyl-methionine was produced with the highest yield<sup>[8]</sup>. Therefore, we choose this amino acid for optimization of the acylation reaction.

#### Determination of pH and temperature optimum

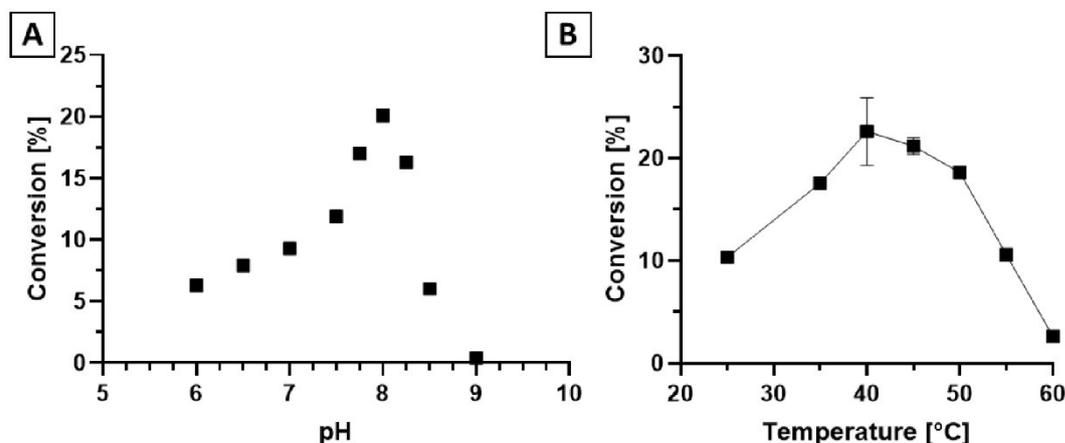


Figure 1: pH and temperature dependency of biocatalytic lauroyl-methionine synthesis with the aminoacylase MsAA. (A) pH dependency of lauroyl-methionine synthesis. The reactions were conducted at 40 °C for 72 h with 100 mM methionine and 100 mM lauric acid in 100 mM Tris-HCl at various pH. All reactions were conducted in triplicates (B) Temperature dependency of lauroyl-methionine synthesis. The reactions were conducted for 72 h with 100 mM methionine and 100 mM lauric acid in 100 mM Tris-HCl at pH 8.0 at various temperatures. All reactions were conducted in triplicates. Error bars under 0.7 % are not displayed.

The influence of pH on the conversion rate was investigated to find optimal conditions for the synthesis of lauroyl-methionine. In contrast to the hydrolytic pH profile, optimal synthesis reaction was observed at pH 8.0 with 20.1 % conversion (Fig. 1A). The synthesis is highly sensitive to changes of pH. At pH 7.0, only half of the conversion was detected, and at pH 9.0, no synthetic activity was detected at all. The latter can be explained by insufficient stability of the enzyme at pH 9.0<sup>[8]</sup>. Hence, further optimization of the synthesis reaction was conducted at pH 8.0. The slight change of pH optimum to more basic pH values from hydrolysis to synthesis may be explained by higher nucleophilicity of the amino group of methionine at higher pH values. An optimal pH of 8.0 for the acylation of phenylalanine was also determined for the homologous aminoacylase SamAA from *S. ambofaciens*<sup>[7]</sup>.

The optimal temperature for hydrolysis was measured to be 70 °C<sup>[8]</sup>. However, as the synthesis reaction is much slower and may take several days to reach equilibrium, insufficient stability can counteract activity. Thus, the temperature optimum was again determined for synthesis in a range of 25 °C to 60 °C and conversion was measured after 72 h. After 24 h, the optimal reaction temperature was 45 °C with a conversion of 21.3 % and did not further increase after 72 h (data not shown). At 40 °C, conversion after 24 h was 14.7 % and 22.6 % after 72 h (Fig. 1B). Conversion rates at 40 °C and 45 °C are roughly in

## 2. Results - 2.4. Chapter IV

the same range after 72 h. Hence, productivity might also be the same due to reached equilibrium conditions. The optimal temperature for synthesis with SamAA was 45 °C, and the enzyme quickly lost its activity at 55 °C<sup>[7]</sup>.

### Determination of optimal substrate concentrations

In general, a substrate excess of one of the substrates is present in excess towards the other to perform acylation by aminoacylases. After determination of the pH and temperature profiles of MsAA with both substrates used equimolarly at 100 mM, the optimal substrate concentrations for the synthesis were determined. The concentration of one substrate was kept constant at 100 mM, while the other substrate was set to 25, 50, 100, 150 or 200 mM. Furthermore, the reaction of 300 mM and 400 mM methionine was tested with lauric acid at 100 mM. Reactions were conducted at equimolar ratios at the same concentration increments as well. At 400 mM methionine, various concentrations of lauric acid were investigated. An increasing methionine concentration had the strongest positive effect on product formation. With 100 mM lauric acid, increasing the methionine concentration steadily increased the formation of lauroyl-methionine. An excess of lauric acid inhibited the reaction; the methionine concentration should thus be adjusted to the lauric acid concentration. The highest concentrations of lauroyl-methionine were determined at 400 mM methionine and 150 mM lauric acid with 84.9 mM product and 56.6 % conversion (Fig. 2).

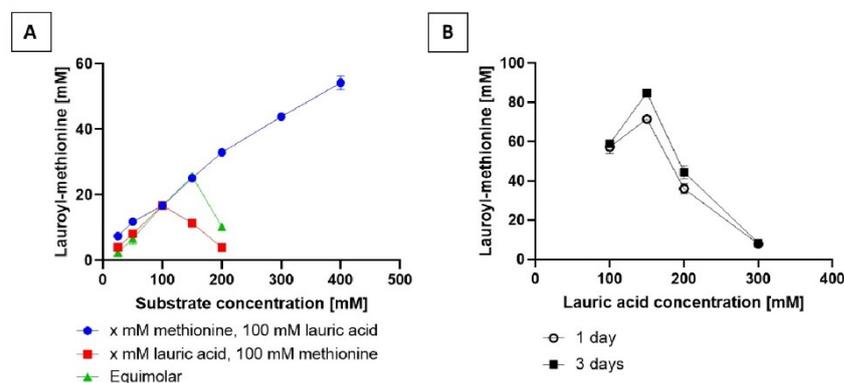


Figure 2: Lauroyl-methionine synthesis using the aminoacylase MsAA at varying substrate conditions. (A) Synthesis of lauroyl-methionine with varying concentration of methionine and lauric acid. Product formation was tested with 100 mM lauric acid and 25 – 400 mM methionine (blue), 100 mM methionine and 25 – 200 mM lauric acid (red), and at equimolar conditions from 25 – 200 mM (green). Reactions were conducted at pH 8.0 and 45 °C for 24 h. (B) Synthesis with 400 mM methionine and 100 – 300 mM lauric acid at pH 8.0 and 40 °C after 24 h and 72 h. All reactions were conducted in triplicates. Error bars under 2 mM are not displayed.

### Upscaling of the enzymatic acylation

In order to test if the biocatalytic conversion and the reaction course remained consistent between different reactor types, an upscaling to a stirred enzyme reactor was performed. The small-scale reaction

was conducted with 1 mL total reaction volume in 1.5 mL reaction tubes at 500 rpm. The larger scale consisted of 20 mL reaction volume in a 100 mL stirred reactor. The reaction was conducted at 40 °C with 400 mM methionine and 150 mM lauric acid. In both scales, product formation over the reaction course was almost identical, indicating that results obtained from small-scale reaction tube screening can be transferred to the dimensions of a lab-scale stirred tank reactor. Maximal product formation was observed after 72 h with 100.2 mM lauroyl-methionine, which is a conversion rate of 67 % compared to lauric acid (Fig. 3). Compared to the initial product concentration and conversion rate of 7.4 mM and 7.4 %<sup>[8]</sup>, the optimization of reaction parameters resulted in a 13.5-fold increase of product concentration and a 9.0-fold increase in conversion rate.

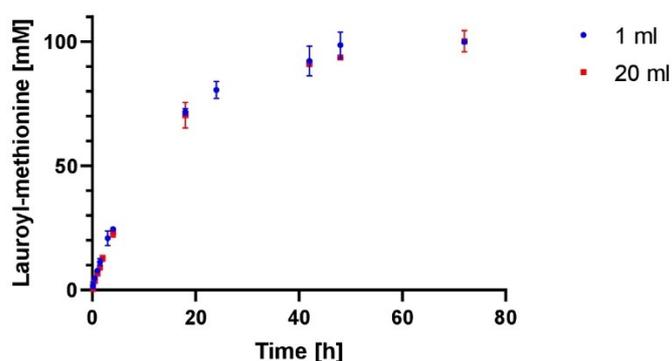


Figure 3: Time course of the biocatalytic synthesis reaction in 1 ml and 20 ml scale

The reaction was conducted with 400 mM methionine and 150 mM lauric acid in 100 mM Tris-HCl pH 8.0 at 40 °C in either 1.5 mL reaction tubes with 1 mL filling volume or in Wheaton Cellstir® reactors with 20 mL filling volumes. The tubes were shaken on an orbital shaker with 500 rpm, and the Cellstir® reactors were stirred with 500 rpm. Error bars under 1 mM are not displayed.

### Optimization of acylation with proteinogenic amino acids

After finding the optimized acylation conditions for methionine, other acyl acceptors were investigated. In contrast to the initial substrate scope, which was conducted at pH 7.0, the pH of the reactions was set to pH 8.0. The reaction was conducted at 45 °C for 24 h. No further acylated amino acid products were detected for acylation by altering the pH-value<sup>[8]</sup>. Furthermore, because it was observed that a high excess of methionine increased the product concentration, it was investigated if the conversion of the other amino acids, namely alanine, isoleucine, leucine, phenylalanine, and valine, could be increased by varying the amino acid concentration as well. The amino acids were used in concentrations from 25 mM to 200 mM, which was the limit of solubility for these substrates in aqueous buffer. In all cases, the conversion rate could be increased by higher amino acid concentration up to 200 mM. The highest product concentrations for lauroyl-alanine, -isoleucine, -leucine, -phenylalanine, and -valine were 2.8 mM, 1.9 mM, 1.4 mM, 0.26 mM, and 9.7 mM respectively (Fig. 4). Hence, with 9.7 mM product

concentration, valine was the second-best substrate for acylation, after methionine, which yielded 35.0 mM product under similar conditions.

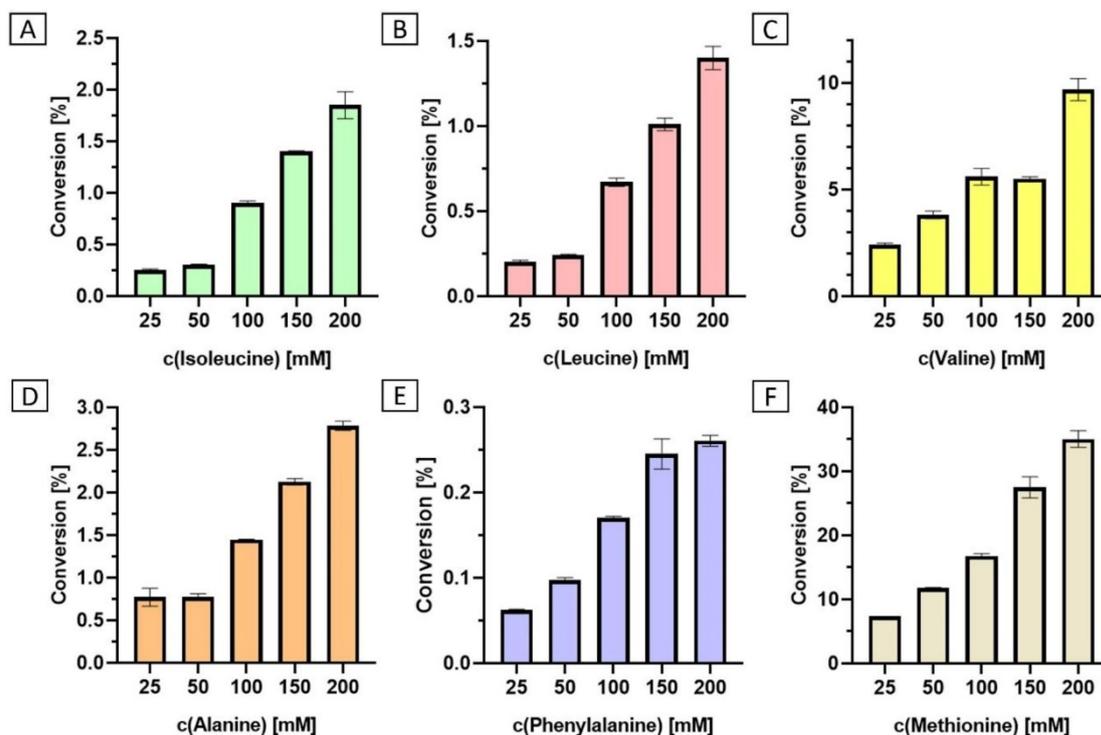


Figure 4: Lauroyl-amino acid synthesis with the aminoacylase MsAA at varying substrate conditions. Synthesis of lauroyl-isoleucine (A), and -leucine (B), -valine (C), -alanine (D), -phenylalanine (E), and -methionine (F) with varying concentration of the respective amino acid and 100 mM lauric acid at pH 8.0 and 45 °C for 24 h. All reactions were conducted in triplicates.

#### Acyl donor specificity for acylation of methionine

The acyl donor specificity for the acylation of methionine was investigated with caprylic acid, decanoic acid, 10-undecenoic acid, myristic acid, palmitic acid, oleic acid and the aromatic acids cinnamic acid and ferulic acid. The acylation with lauric acid served as a reference to estimate product concentrations for the remaining aliphatic fatty acids. With all investigated fatty acids, product could be detected by HPLC-MS. Cinnamic acid and ferulic acid were not accepted as substrates by MsAA. Palmitic acid and oleic acid products could not be quantified via UV detection. The acylation efficiency of the remaining acyl-amino acids was estimated by using the area below the product peak obtained by UV detection of lauroyl-methionine as a reference. With lauroyl-methionine set as 100 %, the acylation of methionine with caprylic acid, decanoic acid, 10-undecenoic acid, and myristic acid were 54 %, 72 %, 87 %, and 28 % compared to acylation with lauric acid (Table 1). The results are in line with the results obtained with the homologous aminoacylase from *S. ambofaciens*, where lauric acid and undecenoic acid were the preferred acyl donors, while acylation with caprylic acid and oleic acid yielded lower conversions<sup>[6]</sup>.

Table 1: Acyl donor specificity for the acylation of methionine with aminoacylase MsAA. The reactions were conducted with 400 mM methionine and 150 mM fatty acid (<sup>[a]</sup>myristic acid was used at 100 mM and <sup>[b]</sup>palmitic and oleic acid were used at 50 mM due to low solubility) at pH 8.0 and 40 °C for 72 h. n.d. = not detected. 100 % peak area for lauroyl-methionine corresponds to 84.9 mM product.

Fatty acid	Peak area (UV <sub>210 nm</sub> ) relative to lauroyl-methionine product [%]
Caprylic acid (C8:0)	54
Decanoic acid (C10:0)	72
10-Undecenoic acid (C11:1)	87
Lauric acid (C12:0)	100
Myristic acid (C14:0) <sup>[a]</sup>	28
Palmitic acid (C16:0) <sup>[b]</sup>	n.d., traces detected by MS
Oleic acid (C18:1) <sup>[b]</sup>	n.d., traces detected by MS
Cinnamic acid	n.d.
Ferulic acid	n.d.

#### Influence on glycerol and temperature on synthesis reaction

The use of glycerol can have an influence on the conversion rate for two reasons. First, the replacement of water by glycerol can shift the reaction equilibrium since water is a product of the condensation reaction. Second, glycerol can have a positive impact on enzyme stability, protecting its activity during longer periods of heat-exposure in the synthesis reaction, potentially improving yield. Therefore, the influence of glycerol on the synthesis of lauroyl-methionine at pH 8.0 for 72 h was investigated by varying reaction temperature as well, so that both effects can be evaluated. For this, a two-factor, face-centered central composite design of experiments was used. The experimental design limited the highest glycerol content to 30 %, as 400 mM methionine was not soluble at a higher glycerol concentration. As Fig. 5 shows, a positive effect of glycerol addition could be observed at all temperatures. The effect was less pronounced at 40 °C, but yielded substantial improvement of conversion at 60 °C. Without glycerol, almost no product was formed at 60 °C (2.4 mM), but with 15 % and 30 % glycerol, 14.4 mM and 37.1 mM lauroyl-methionine was formed at the same temperature, respectively. The positive effect of glycerol on product formation at higher temperatures demonstrates the action of glycerol on improving thermostability. At a reaction temperature of 40 °C, a glycerol content of 0 %, 15 %, and 30 % resulted in 65.1 mM, 75.1 mM, and 76.1 mM lauroyl-methionine, respectively. At a higher reaction temperature of 50 °C, product concentrations of 24.6 mM, 73.2 mM, and 78.3 mM were measured for a glycerol content of 0 %, 15 %, and 30 %, respectively. The formula resulting from statistical analysis and estimation of coefficients is:

$$y = a * T + b * G + c * T * G + d * T^2 + e * G^2 + const$$

with y = product [mM], T = temperature [°C] and G = glycerol-content [%], const = constant.

## 2. Results - 2.4. Chapter IV

Table 2: Estimation of the coefficients for the surface response model of the experimental design. The values are coded experimental based on conditions between -1 and +1, meaning -1 for 40 °C and 0 %, and +1 for 60 °C and 30 %, for temperature and glycerol content, respectively.

Term	Coefficient	Value	Standard error	p-value
Temperature (40 °C, 60°C)	A	-27.049	3.009	< 0.001
Glycerol content (0 %, 30 %)	B	16.54	3.009	<0.001
Temperature x Glycerol	C	5.959	3.685	0.1282
Temperature <sup>2</sup>	D	-16.612	4.824	0.0040
Glycerol <sup>2</sup>	E	-9.915	4.824	0.0590
Constant value	const	68.248	4.404	< 0.001

The estimated values for the formula of the model are shown in Table 2. The parameters a, b, and const were highly significant with a p value < 0.001, and parameter d was very significant with a p value of 0.0040. Parameter e (for G<sup>2</sup>) was close to being significant (p = 0.0590). Parameter c was not significant (p = 0.1282). The latter was kept in the model to improve accuracy in the response curve. According to the model, a maximal product concentration will be achieved at 42.97 °C and 23.34 % glycerol with 82.9 mM lauroyl-methionine. This concentration has been surpassed in previous experiments, for example from the determination of optimal substrate concentrations, where the formation of 84.9 mM lauroyl-methionine from 400 mM methionine and 150 mM lauric acid at pH 8.0 and 40 °C was observed after 72 h reaction (Fig. 2(B)).

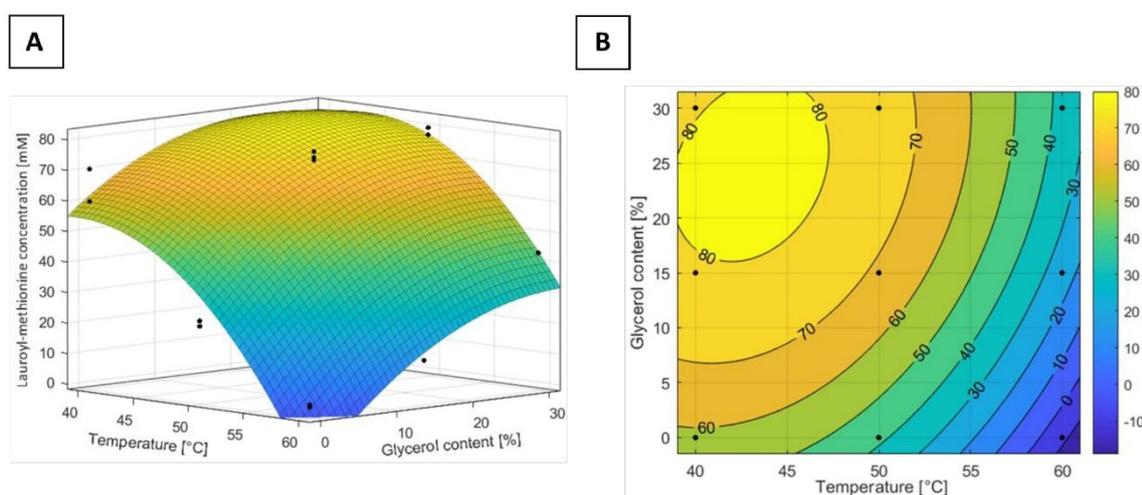


Figure 5: Influence of temperature and glycerol on synthesis reaction

Surface response (A) and isoresponse (B) curves of lauroyl-methionine concentration after 72 h synthesis at pH 8.0 versus temperature and glycerol content of the reaction mixture. The measured

experimental points are indicated as black dots (2 duplicate corner points each, 2 duplicate star points each, quadruplicate central point). The numbers of the color scale for the isoresponse curve indicate product concentrations in [mM]. The substrate concentration was 400 mM methionine and 150 mM lauric acid under all conditions.

### **Molecular modeling of substrate docking and proposed mechanism for the synthesis reaction**

The protein sequence analysis of MsAA, along with its AlphaFold-predicted structure with added zinc ions, has recently been published<sup>[8]</sup>. The enzyme can be assigned to the M20A metallopeptidase family and its metal-binding residues (H91, D123, E158, E185, H425) for the cocatalytic zinc site and catalytic residues (D93, E157, H226) are conserved, as in HiDapE, or in SamAA, the homologous aminoacylase from *S. mobaraensis* (SmAA)<sup>[14]</sup>, and the porcine aminoacylase-1 (pAcy1)<sup>[15]</sup> (Table 3). The residue D93 is considered as catalytic, because it can form interactions between the N<sub>δ</sub>-proton of H91, assisting in orientation and influencing the electronegativity of the N<sub>ε</sub>-nitrogen, leading to a decreased Lewis acidity of the zinc ion<sup>[10,16]</sup>. Furthermore, the dimeric protein structure has been described for MsAA<sup>[8]</sup>, HiDapE<sup>[12]</sup>, and pAcy1<sup>[17]</sup>. In contrast, SmAA has been described as a monomeric enzyme<sup>[14]</sup> and the potential oligomeric state of SamAA has not yet been determined. The aminoacylase MsAA shows 21.4 % sequence identity to HiDapE, 25.5 % identity to pAcy1, 25.2 % identity to the human aminoacylase-1 (hAcy1)<sup>[18]</sup>, and 56.3 % and 54.9 % identity to SamAA and SmAA, respectively. Despite sharing similarities with the aminoacylases SmAA and pAcy1, HiDapE does not hydrolyze *N*-acetyl-amino acids like acetyl-glutamate, acetyl-glutamine, acetyl-arginine, acetyl-lysine, and acetyl-ornithine, but can catalyze the hydrolysis of *N*-succinyl-*L,L*-diaminopimelic acid to succinate and *L,L*-diaminopimelic acid<sup>[19]</sup>.

### **Docking of lauric acid and methionine and reaction intermediate and product to MsAA**

The experimental investigation of the acylation of amino acids with MsAA revealed that the enzyme has a strong preference for methionine and accepts only hydrophobic amino acids as substrates for acylation. With the protein structure predicted by AlphaFold, we performed docking experiments to comprehend this behavior and understand which amino acid residues favor binding of hydrophobic amino acid substrates, especially methionine, in the acylation reaction. In contrast, MsAA also accepts some hydrophilic acetyl-amino acids in hydrolysis<sup>[8]</sup>, for example acetyl-glutamic acid or acetyl-arginine. Furthermore, by comparison with the structurally and mechanistically well-studied HiDapE, which can hydrolyze *N*-succinyl-*L,L*-diaminopimelic acid, a putative mechanism for aminoacylase MsAA can be proposed. Due to the conserved catalytic and metal-binding residues between MsAA and HiDapE, it seemed worthwhile to compare these enzymes. The structure of HiDapE with its ligands (PDB 5VO3) has been solved by X-ray crystallography<sup>[12]</sup>.

Both substrates, lauric acid and methionine, could be docked to the active site of dimeric MsAA (Fig. 6A). The carbonyl-oxygen of lauric acid is positioned between the two zinc ions. The methionine

## 2. Results - 2.4. Chapter IV

molecule is positioned with its amino group facing the carbonyl group of lauric acid. In the docking experiment, interactions were established between the catalytic residue E157<sub>MsAA</sub> with both methionine and lauric acid and distances between these groups are  $< 2.9 \text{ \AA}$ . The glutamic acid residue might act in deprotonation of the amino group of methionine and protonation of the carboxyl group of lauric acid, as it is suggested for HiDapE. The interaction with E157<sub>MsAA</sub> with both the amino group of methionine and the carboxylic group of lauric acid suggests and affirms the role as a general acid/base in catalysis<sup>[20]</sup>.

In MsAA, N315(B)<sub>MsAA</sub> of the second dimer binds the amino acid at the  $\alpha$ -carboxylic group (Fig. 6), and this residue is conserved in HiDapE as N245(B)<sub>HiDapE</sub> and also in the human aminoacylase hAcy1 as N263(B)<sub>hAcy1</sub><sup>[18]</sup>. The presence of  $\alpha$ -carboxylic groups in both acyl-amino acids as well as in diamino pimelic acid explains the similarities of substrate binding. In HiDapE, adjacent to N245(B)<sub>HiDapE</sub>, a second N244(B)<sub>HiDapE</sub> binds the  $\epsilon$ -carboxylic group of diamino pimelic acid. This N244(B) is not conserved in MsAA but exchanged to a hydrophobic A314(B)<sub>MsAA</sub>. This might explain why MsAA does not accept e.g. glutamic acid as a substrate for acylation.

Furthermore, in HiDapE, Y197(B)<sub>HiDapE</sub> of the second dimer, oriented in proximity to H194(B)<sub>HiDapE</sub>, participates in substrate binding, namely the non-amide forming carboxylic group of the succinic acid moiety<sup>[12]</sup>. This tyrosine residue is exchanged to M229<sub>MsAA</sub> in MsAA and participates in binding of lauric acid by hydrophobic interactions (Fig. 6B). In HiDapE, the same carboxylic group of succinic acid interacts with R178<sub>HiDapE</sub> but is exchanged to E210<sub>MsAA</sub> in MsAA. This glutamic acid residue is however not structurally conserved and does not participate in substrate binding in MsAA docking.

In HiDapE, R258<sub>HiDapE</sub> binds the  $\alpha$ -carboxylic group of diaminopimelic acid, both this residue and its function are conserved in MsAA as R328<sub>MsAA</sub> (Fig. 6A). For hAcy1, this residue is also conserved as R276<sub>hAcy1</sub>. Furthermore, for hAcy1, it is described that D274<sub>hAcy1</sub> assists in orientation of R276<sub>hAcy1</sub>, possibly through a salt bridge, and a D274A<sub>hAcy1</sub> variant exhibits strongly reduced activity<sup>[18]</sup>. This aspartic residue is conserved in MsAA as D326<sub>MsAA</sub> and might similarly interact with R328<sub>MsAA</sub> (Fig. 6A). In HiDapE, the aspartic acid residue is exchanged to an asparagine. There is some discrepancy for the substrate binding role of R276<sub>hAcy1</sub> in hAcy1, as it is shown to not be essential, and R348<sub>hAcy1</sub> has been postulated for substrate binding<sup>[15]</sup>. This residue is not conserved in MsAA, where it is exchanged to K397<sub>MsAA</sub>. However, this residue is conserved in HiDapE as R329<sub>HiDapE</sub> and interacts with the  $\epsilon$ -carboxylic group of diaminopimelic acid. For hAcy1, a R348A<sub>hAcy1</sub> variant was not hydrolytically active, hence it was suspected to be involved in substrate binding<sup>[15]</sup>. Liu *et al.* have published the docking of *N*-acetyl-L-methionine to a homology model of hAcy1, which indicated hydrogen bonding of R348<sub>hAcy1</sub> to the  $\alpha$ -carboxylic acid and amide-carbonyl groups<sup>[15]</sup>, suggesting a similar role to R258<sub>HiDapE</sub> in HiDapE and R328<sub>MsAA</sub> in MsAA.

For the docking of lauric acid and methionine to MsAA (Fig. 6A), the substrate pocket and residues of amino acids were visualized (Fig. 6B). A clear partitioning of the substrate pocket into hydrophilic and hydrophobic regions for amino acid and fatty acid, respectively, can be observed. Furthermore, oriented below the side chain of methionine, two structurally adjacent leucine residues, L213<sub>MsAA</sub> and L357<sub>MsAA</sub>,

participate in substrate binding, which may be responsible for the specificity of MsAA for methionine and hydrophobic amino acids in general. These two leucine residues are not conserved in the homologous enzymes SamAA, SmAA, HiDapE, or hAcy1. These enzymes do not exhibit a bias towards hydrophobic amino acids. In the HiDapE protein structure, in place of these two hydrophobic leucine residues, two serine residues can be found, which can interact with the amino- and carboxylic groups of diaminopimelic acid.

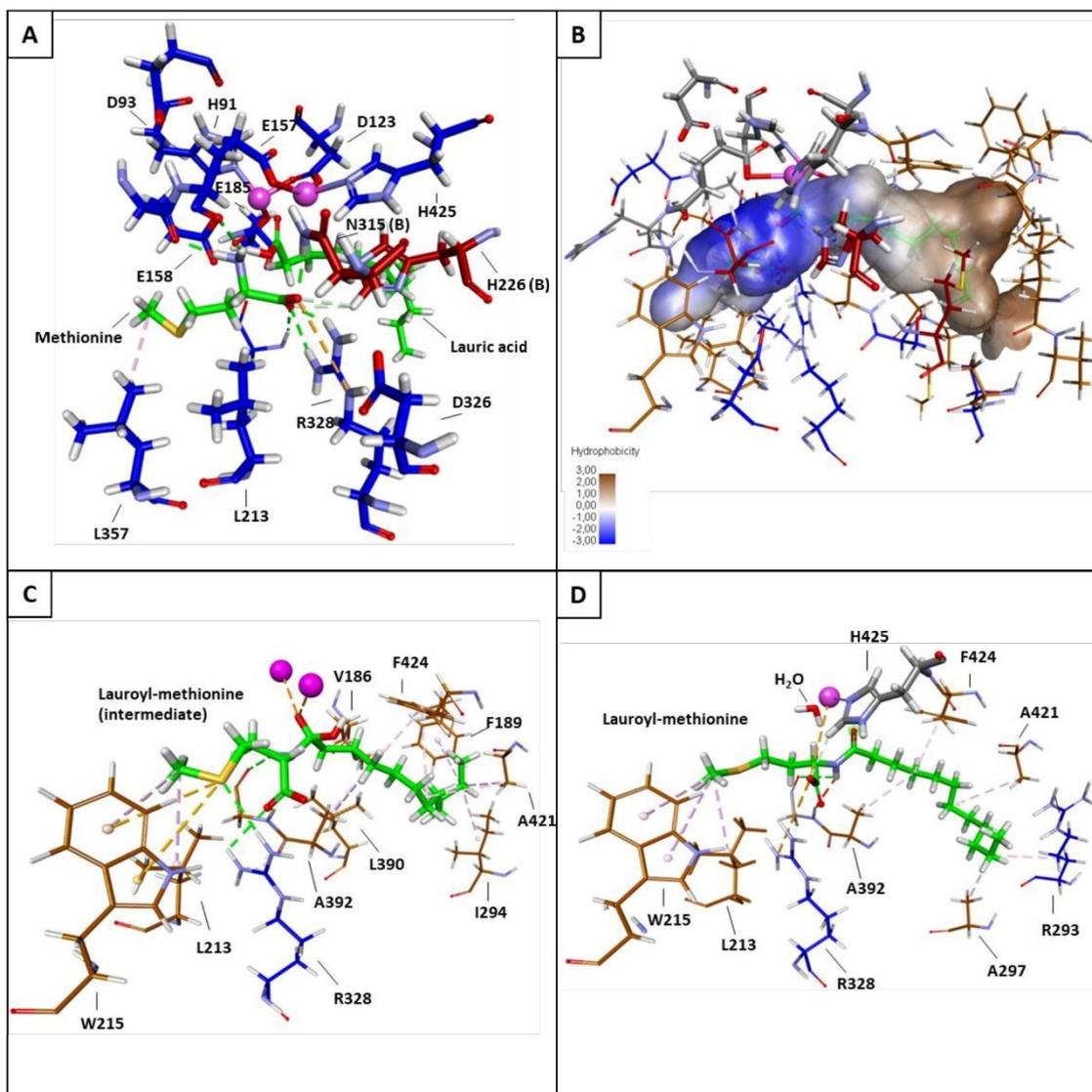


Figure 6: Results obtained from docking experiments with methionine and lauric acid, the reaction intermediate and product to MsAA.

(A) Docking of lauric acid and methionine to the active site of dimeric MsAA. The protein residues are depicted with grey carbon atoms with zinc ions shown as magenta balls. The substrates methionine and lauric acid are shown with green carbon atoms. The dimers A and B are shown with blue and red carbon atoms, respectively. The ligand interactions are shown as dashed lines and green for H-bonds, orange for electrostatic interactions, lilac for hydrophobic interactions, and yellow for sulfur- $\pi$  bonds.

## 2. Results - 2.4. Chapter IV

(B) Substrate pocket with highlighted hydrophilic and hydrophobic domains from docking with lauric acid and methionine to dimeric MsAA. Hydrophobic regions and amino acids are shown in brown color, hydrophilic regions and residues are shown in blue. Residues of the active site are shown with grey carbon atoms, and residues from the adjacent dimer are shown with red carbon atoms.

(C) Docking of tetrahedral lauroyl-methionine-intermediate (green carbon atoms) to monomeric MsAA.

(D) Docking of lauroyl-methionine (green carbon atoms) to monomeric MsAA.

Table 3: Functional amino acids of MsAA and its homologues and their proposed functions.

MsAA	SmAA	hAcy1	HiDapE	Function
H91	H107	H80	H67	Metal-binding
D123	D139	D113	D100	
E158	E174	E148	E135	
E185	E201	E175	E163	
H425	H435	H373	H349	
D93	D109	D82	D69	Catalytic, orientation of H67 <sub>HiDapE</sub> <sup>[10]</sup> or H91 <sub>MsAA</sub> (not shown for hAcy1)
E157	E173	E147	E134	Catalytic, acid-base catalyst
H226	H238	H206	H194	Catalytic, oxyanion hole-forming <sup>[12,18]</sup> ; not essential in pAcy1 <sup>[15]</sup>
L213	M225	P193	S181	Amino acid-binding in HiDapE and MsAA
L357	D367	N307	S290	
W215	W227	W195	T183	Amino acid-binding in MsAA
R328	R338	R276	R258	Binding of $\alpha$ -carboxylic acid <sup>[12,18]</sup>
K397	K407	R348	R329	Binding of $\alpha$ -carboxylic acid in hAcy1 <sup>[15]</sup> ; Possibly amino acid binding in HiDapE <sup>[10]</sup>
M229	M241	R209	Y197	Binding of succinic acid (carboxyl) in HiDapE <sup>[12]</sup> ; Acyl-binding in MsAA; hAcy1 unknown
D210	Q222	E190	R178	Binding of succinic acid (carboxyl) in HiDapE
N315	N325	N263	N245	Binding of amino acid at $\alpha$ -carboxylic group in HiDapE <sup>[12]</sup> , hAcy1 <sup>[18]</sup> , and MsAA
A314	V324	Y262	N244	Binding of amino acid at $\epsilon$ -carboxylic group in HiDapE <sup>[12]</sup> .
D326	D336	D274	N256	Orientation of arginine residue (R276 <sub>hAcy1</sub> ) <sup>[18]</sup>

In the acylation reaction, after binding of lauric acid and methionine to the active site, the amino acids nitrogen attacks the carbonyl carbon of the fatty acid, so that a tetrahedral intermediate is formed. This

intermediate breaks down to the acylated amino acid, in this case methionine, and water. In order to follow the condensation reaction *in silico*, and to get a better understanding of substrate binding during the reaction steps, a docking simulation was performed with the tetrahedral intermediate of lauroyl-methionine. As shown in Fig. 6C, the negative charge of the oxygen atom is stabilized by the zinc ions. The influence of H226(B)<sub>MsAA</sub> could not be investigated, as the docking of the intermediate was only successful with the monomeric structure. The residues binding the acyl- and amino acid parts are similar, compared to the previous docking with methionine and lauric acid. However, the position of methionyl has been slightly shifted towards the active site and lauric acid. This led to more *in silico* interactions between the substrate's thioether group and R328<sub>MsAA</sub> (hydrogen bond) and W215<sub>MsAA</sub> (sulfur- $\pi$ ).

To perform the docking with lauroyl-methionine, the second product of the condensation reaction, a water molecule was added to the protein structure. The position of the water molecule was taken from a crystal structure of HiDapE (PDB 3IC1) by first superimposing the two structures, copying the water molecule, which was located between the two zinc ions, and performing an energy minimization to position the molecule properly. Afterwards, lauroyl-methionine was docked to the enzyme with the bound water molecule. Again, as in the docking with lauric acid and methionine, or with the intermediate structure, the interacting protein side chains were the same (Fig. 6D).

In summary, the protein side chains of MsAA interacting with the ligands in various stages of the condensation reaction were similar. The substrate pocket shows a distinct hydrophobic and hydrophilic side for the lauric acid and methionine moiety, respectively. The specificity for methionine, or bias towards small, hydrophobic amino acids, might be mediated by L213<sub>MsAA</sub> and L357<sub>MsAA</sub>, and possibly W215<sub>MsAA</sub>. These residues are exchanged for S181<sub>HiDapE</sub> and S290<sub>HiDapE</sub> in HiDapE<sub>HiDapE</sub>, and W215<sub>MsAA</sub> is exchanged for T183<sub>HiDapE</sub>, which does not interact with ligand in the structure PDB 5VO3. However, R328 is conserved in HiDapE and, like in the docking with MsAA, binds the carboxyl group adjacent to the nucleophilic amino group of the ligand. The results of the docking study, compared to the co-crystallized structure of HiDapE with its bound substrates, revealed a very similar orientation of the substrates. The differences in substrate-interacting residues that were observed and pointed out could explain the differences in the substrate scope. Future work will focus on mutagenesis of MsAA to conclude the role of the residues in substrate binding.

Based on the results obtained by docking, and by comparison with proposed mechanism for hydrolysis of L,L-succinyl diaminopimelic acid by HiDapE<sup>[12]</sup>, we proposed a mechanism for synthesis of lauroyl-methionine from methionine and lauric acid by MsAA (Fig. 7). The reaction may be initiated by lauric acid replacing the water molecule that is bound by Zn1 and Zn2. The negative charge of the deprotonated lauric acid is stabilized by the zinc ions, and, as it has been shown for HiDapE<sup>[12]</sup>, the acidic oxygen of the substrate can be bound by Zn1, while the carbonyl oxygen is stabilized by Zn2. The catalytic E157<sub>MsAA</sub> can deprotonate the amino group of the amino acid, so that it can perform a nucleophilic attack on the carbonyl carbon. This results in the formation of a tetrahedral complex and addition of a negative charge of the carbonyl oxygen. Again, as it has been described for HiDapE, H226(B)<sub>MsAA</sub> could form an

oxyanion hole to stabilize the intermediate. The complex decomposes under formation of lauroyl-methionine and the zinc-bound hydroxide, which can subsequently be protonated by E157 to form water.

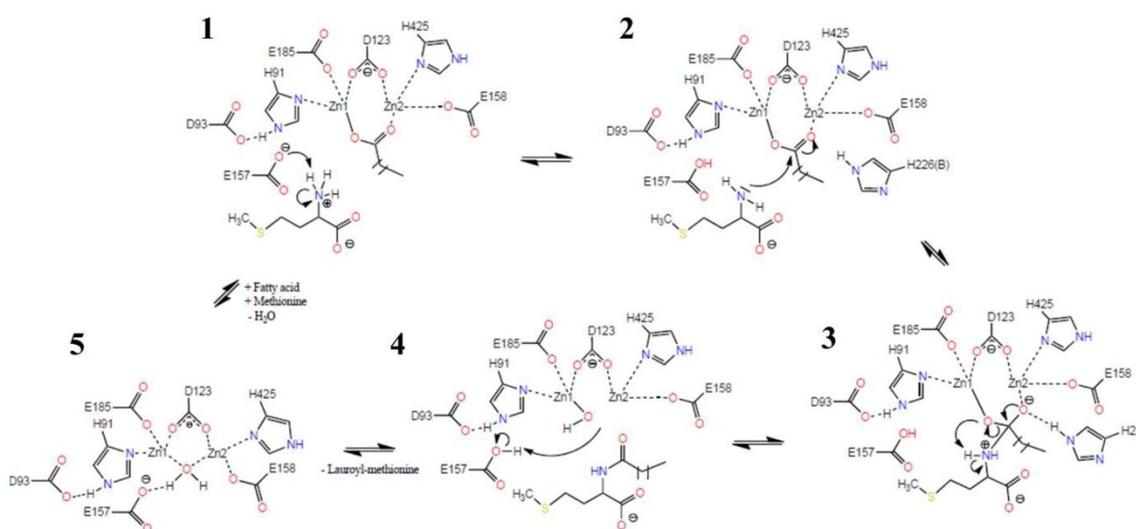


Figure 7: Proposed catalytic mechanism for the condensation of a fatty acid and methionine by aminoacylase MsAA. Step 1 shows binding of the methionine and lauric acid substrate to the active site, and the first reaction step of deprotonating the  $\alpha$ -amino group of methionine by E157. Step 2 comprises the nucleophilic attack under formation of the tetrahedral intermediate, stabilized by H226(B). In step 3, the complex decomposes under formation of a zinc-bound hydroxide ion. The hydroxide ion is protonated by E157 in step 4. Step 5 shows the water molecule between the two zinc ions. When new methionine and and lauric acid is added, the lauric acid replaces the water molecule in the active site and the reaction begins again at step 1.

## Methods

### Chemicals

The cultivation media, amino acids, metal salts, solvents, and tris(hydroxymethyl)aminomethane (Tris) were obtained from Carl Roth (Germany). *N*-acetyl-L-alanine was purchased from Sigma Aldrich (USA). The other acyl-amino acids, used as a substrate or standard for analytics, were synthesized as described previously<sup>[8]</sup>. D-desthiobiotin was obtained from Merck (Germany). The EZ Nin reagent was purchased from Biochrom (UK) and obtained from Laborservice Onken (Germany). Fatty acids and remaining chemicals were from Sigma Aldrich.

### Production of MsAA aminoacylase

The aminoacylase MsAA has been recombinantly produced and purified as described before<sup>[8]</sup>. *E. coli* ArcticExpress (DE3) (Agilent Technologies, USA) carrying the plasmid pET28a MsAA NTag was cultivated in Terrific Broth autoinduction medium (2 % tryptone from casein, 2.4 % yeast extract, 25 mM Na<sub>2</sub>HPO<sub>4</sub>, 25 mM KH<sub>2</sub>PO<sub>4</sub>, 50 mM NH<sub>4</sub>Cl, 2 mM MgSO<sub>4</sub>, 5 mM Na<sub>2</sub>SO<sub>4</sub>, 0.5 % glycerol (v/v), 0.5 % lactose and 0.05 % glucose) without antibiotics for expression. The expression temperature was 30 °C for the first 6 h and was lowered to 12 °C for the following 24 h of expression. The cells were harvested and disrupted by sonication. The buffer used for lysis and purification was 100 mM Tris-HCl pH 7.0 with 150 mM NaCl and 1 mM ZnCl<sub>2</sub>. Additives for lysis were 0.3 mg/mL lysozyme and 0.1 % Triton X-100. Recombinant MsAA was purified by Strep-TagII affinity chromatography and elution from Strep-Tactin<sup>®</sup> SuperFlow<sup>®</sup> high capacity cartridge (IBA, Germany) was done with 2.5 mM D-desthiobiotin. The eluted protein solution was rebuffered without desthiobiotin using Vivaspin centrifugal filters (10,000 MWCO; Sartorius, Germany).

### Aminoacylase activity assay

Activity of aminoacylases was assayed by quantification of released amino acids with a ninhydrin-based assay as previously described<sup>[21]</sup>. Briefly, 10 µL sample from amino acid solutions or aminoacylase reactions were mixed with 100 µL of EZ Nin:DMSO reagent, heated for 10 min at 99 °C and diluted with 100 mM Na-borate buffer pH 10.0 for measurement. In general, 200 µL reactions consisted of 190 µL substrate solution and 10 µL enzyme solution. For standard hydrolysis activity measurement, reaction with 15 mM *N*-acetyl-L-alanine in 100 mM Tris-HCl buffer pH 7.0 were performed at 30 °C for 5 min. At 1 min sampling intervals, 10 µL samples were withdrawn and assayed with the ninhydrin assay. The absorbance of diketohydrindylidene-diketohydrindamine, also called Ruhemann's purple was determined at 570 nm with the Infinite M Nano absorbance plate reader (Tecan<sup>®</sup> Group Ltd., Männedorf, CHE). One unit of MsAA was defined as the amount of enzyme that hydrolyzes one µmol of *N*-acetyl-L-alanine per minute under the given conditions, respectively.

### Biocatalytic synthesis of acyl-amino acids

Initial biocatalytic synthesis of lauroyl-methionine was investigated from 100 mM L-methionine and 100 mM lauric acid in 100 mM Tris-HCl pH 7.0 at 40 °C for 72 h without agitation. Generally, reactions were started by adding 10 µg of MsAA (1.3 U) to a reaction volume of 0.5 mL, resulting in an enzyme concentration of 0.02 mg/mL, to the reaction mixture. To stop the reaction, a 100 µL sample was withdrawn and immediately mixed with 100 µL of a mixture of 80 % acetonitrile and 20 % water containing 0.1 % trifluoroacetic acid (TFA). The optimal pH for the synthesis reaction was determined from the reaction of 100 mM methionine and 100 mM lauric acid in 100 mM Tris-HCl adjusted to various pH values at 40 °C with analysis after 72 h. All further reactions were conducted in 100 mM Tris-HCl at pH 8.0. The influence of reaction temperature on product formation was investigated in a

## 2. Results - 2.4. Chapter IV

range of 25 to 60 °C and was analyzed after 72 h. The substrate solution consisted of 100 mM methionine and 100 mM lauric acid. Note that the pH was not adjusted during the reaction.

The dependence of substrate concentrations for final product formation was investigated in a range of 25 – 400 mM for methionine and 25 – 200 mM for lauric acid. The respective other substrate was either kept constant at 100 mM or changed equimolarly, the latter only up to 200 mM concentration of lauric acid. The reactions were conducted at 45 °C and analyzed after 24 h. In the case of 400 mM methionine, concentrations of either 100 mM, 150 mM, 200 mM, or 300 mM lauric acid were used for synthesis, to determine the optimal ratio of the two substrates. For the acylation of alanine, isoleucine, leucine, phenylalanine and valine, the amino acid concentration was varied from 25 - 200 mM, while the concentration of lauric acid was kept constant at 100 mM. The reactions were conducted at 45 °C for 24 h. The acyl donor specificity was investigated by employing 400 mM methionine with 150 mM of various fatty acids. As acyl donors, caprylic acid, decanoic acid, 10-undecenoic acid, lauric acid, palmitic acid, oleic acid, cinnamic acid, and ferulic acid were investigated. Due to low solubility, myristic acid was used at 100 mM, and oleic acid and palmitic acid were used at 50 mM. The reactions were conducted at 45 °C for 24 h.

It was investigated if upscaling of the reaction had an influence on productivity. One reaction consisted of 1 mL and was conducted in 1.5 mL reaction tubes in a thermoshaker (Thermomix C; Eppendorf, Germany) at 40 °C and 500 rpm. The upscaled reaction was conducted in a 100 mL glass spinner flask (Wheaton® Celstir®; DWK Life Sciences, Germany) at 40 °C with magnetic stirring set to 500 rpm.

### Experimental design

A two-factor face-centered central composite design was used to investigate the influence of glycerol and temperature on the formation of lauroyl-methionine. The design consists of 4 corner points, 4 face-centered star points ( $\alpha = \pm 1$ ), and a central point. Glycerol content was tested at 0 %, 15 %, and 30 %. The tested temperatures were 40 °C, 50 °C and 60 °C. The 4 corner points and the 4 star points were tested in duplicates ( $n = 2$ ), while the center point was tested as quadruplicate experiment ( $n = 4$ ). Hence, 22 reactions were conducted and analyzed. The reaction consisted of 400 mM methionine and 150 mM lauric acid at pH 8.0 with 0.02 mg/mL MsAA added. After 72 h reaction, lauroyl-methionine concentration was measured. With the obtained data, a quadratic model was created to estimate the product concentration as a function of temperature and glycerol content by the following formula.

$$y = a * T + b * G + c * T * G + d * T^2 + e * G^2 + const$$

The coefficients and their statistical significance were calculated with the JMP 10 software (SAS Institute, USA) and surface response and isoresponse curves were generated with the curve fitting tool from Matlab (Mathworks, USA).

## Analytical methods

### HPLC-UV analysis

The quantitative analysis of the acyl-amino acid synthesis reactions was performed by HPLC separation and UV measurement. The HPLC system used was an Sykam S5200/S2100 (Germany) equipped with an ISAspher 100-5 C18 BDS column (C18, 5  $\mu\text{m}$ , 4.0 \* 250 mm; Isera, Germany) coupled with a Sykam UV Detector 2500 (Germany). An isocratic separation with 80 % acetonitrile, 20 % water and 0.1 % TFA was performed at 1 mL/min flow rate and a column temperature of 40 °C. The UV absorption of the eluent was measured at 210 nm. The product concentrations were calculated with an external standard of lauroyl-amino acids, which have been chemically synthesized via Schotten-Baumann acylation as described previously<sup>[8]</sup>.

### Mass spectrometry analysis

Qualitative and semi-quantitative analysis of various amino acid acylation products was performed on an HPLC-PDA-HRMS system equipped with a photodiode array (PDA) detector and an Orbitrap ID-X Tribrid™ mass spectrometer (Thermo Scientific). Separation was performed at a flow rate of 0.2 mL/min on an Alltima™ C18 column (2.1 × 100 mm, 3  $\mu\text{m}$ , Hichrom) with methanol/water (80:20, v/v) supplemented with 0.1 % trifluoroacetic acid as phase A and methanol supplemented with 0.1 % trifluoroacetic acid as solvent B, using a 10 min linear gradient elution from 0 to 98% solvent B. Mass spectrometry was performed in alternating positive/negative electrospray ionization mode (ESI<sup>+/−</sup>) with the following conditions: spray voltage was set at 3.5 kV in ESI<sup>+</sup> and -2.5 kV in ESI<sup>−</sup>; source gases were set (in arbitrary units/min) for sheath gas, auxiliary gas and sweep gas at 35, 7, and 10, respectively; vaporizer temperature and ion transfer tube temperature were both set at 300 °C. MS scans were performed from 100 to 500 m/z, at 7.5 K resolution (full width of the peak at its half maximum, fwhm, at 200 m/z) with parameters as follows: RF-lens, 35%; maximum injection time, 50 ms; data type, profile; AGC target: 100000; normalized AGC target: 25%. MS data acquisition and treatment were carried out utilizing the Xcalibur v. 3.0 software (Thermo Scientific).

### Molecular docking simulations

The Discovery Studio 2021 suite from Dassault Systemes Biovia® (France) was used to perform molecular simulations. All simulations were conducted under the CHARMM forcefield<sup>[22]</sup>. The target molecule for the docking simulations was either the dimeric or monomeric MsAA structure, as predicted by the ColabFold (AlphaFold) algorithm<sup>[13]</sup>. As previously described<sup>[8]</sup>, two zinc ions were added to the structure's active site with the Metal Ion-Binding Site Prediction and Docking Server (MIB)<sup>[23]</sup>. Lauric acid, L-methionine, *N*-lauroyl-L-methionine, and a tetrahedral intermediate of the latter molecule were non-covalently docked against the targets. The docking simulations were performed with the CDOCKER method<sup>[24]</sup>.

### Acknowledgements

The authors would like to thank Emma Boulanger from the LRGP for technical assistance in HPLC-UV measurements.

### Conflict of Interest

The authors declare no conflict of interest.

### References

- [1] K. P. Ananthapadmanabhan, *Tenside Surf. Det.* **2019**, 378 -386.
- [2] P. T. Anastas, J. C. Warner, *Green Chemistry: theory and practice*, Oxford University Press, New York, **1998**.
- [3] M. Koreishi, R. Kawasaki, H. Imanaka, K. Imamura, K. Nakanishi, *JAOCS* **2005**, 82, 631–637.
- [4] a) Y. Takakura, Y. Asano, *Biosci Biotechnol Biochem.* **2019**, 83, 1964–1973; b) J. Matsuno, S. Nagai, *J Biochem.* **1972**, 269–279; c) A. Natsch, H. Gfeller, P. Gyax, J. Schmid, G. Acuna, *J. Biol. Chem.* **2003**, 278, 5718–5727; d) I. Gentzen, H.-G. Löffler, F. Schneider, *Z. Naturforsch.* **1980**, 35, 544–550;
- [5] a) M. Koreishi, R. Kawasaki, H. Imanaka, K. Imamura, Y. Takakura, K. Nakanishi, *J Biotechnol.* **2009**, 141, 160–165; b) L. Dettori, F. Ferrari, X. Framboisier, C. Paris, Y. Guiavarc'h, L. Hôtel, A. Aymes, P. Leblond, C. Humeau, R. Kapel et al., *Eng Life Sci.* **2018**, 18, 589–599;
- [6] M. C. Bourkaib, S. Delaunay, X. Framboisier, L. Hôtel, B. Aigle, C. Humeau, Y. Guiavarc'h, I. Chevalot, *Enzyme Microb. Technol.* **2020**, 137, 109536.
- [7] M. C. Bourkaib, S. Delaunay, X. Framboisier, C. Humeau, J. Guilbot, C. Bize, E. Illous, I. Chevalot, Y. Guiavarc'h, *Process Biochem.* **2020**, 99, 307–315.
- [8] G. Haeger, J. Wirges, N. Tanzmann, S. Oyen, T. Jolmes, K. E. Jaeger, U. Schörken, J. Bongaerts, P. Siegert, *Microb Cell Fact.* **2023**, 22.
- [9] N. D. Rawlings, G. Salvesen (Eds.) *Handbook of proteolytic enzymes*, Academic Press, **2013**.
- [10] B. P. Nocek, D. M. Gillner, Y. Fan, R. C. Holz, A. Joachimiak, *J Mol Biol.* **2010**, 397, 617–626.
- [11] S. Rowsell, R. A. Pauptit, A. D. Tucker, R. G. Melton, D. M. Blow, P. Brick, *Structure (London, England : 1993)* **1997**, 5, 337–347.
- [12] B. Nocek, C. Reidl, A. Starus, T. Heath, D. Bienvenue, J. Osipiuk, R. Jedrzejczak, A. Joachimiak, D. P. Becker, R. C. and Holz, *Biochemistry* **2018**, 574–584.
- [13] M. Mirdita, K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov, M. Steinegger, *Nat Meth.* **2022**.
- [14] M. Koreishi, Y. Nakatani, M. Ooi, H. Imanaka, K. Imamura, K. Nakanishi, *Biosci Biotechnol Biochem.* **2009**, 73, 1940–1947.
- [15] Z. Liu, Z. Zhen, Z. Zuo, Y. Wu, A. Liu, Q. Yi, W. Li, *J. Biochem.* **2006**, 139, 421–430.
- [16] D. W. Christianson, J. D. Cox, *Annual review of biochemistry* **1999**, 68, 33–57.
- [17] C. D'Ambrosio, F. Talamo, R. M. Vitale, P. Amodeo, G. Tell, L. Ferrara, A. Scaloni, *Biochemistry* **2003**, 42, 4430–4443.

- [18] H. A. Lindner, A. Alary, L. I. Boju, T. Sulea, R. Ménard, *Biochemistry* **2005**, *44*, 15645–15651.
- [19] D. L. Bienvenue, D. M. Gilner, R. S. Davis, B. Bennett, R. C. Holz, *Biochemistry* **2003**, *42*, 10756–10763.
- [20] T. L. Born, R. Zheng, J. S. Blanchard, *Biochemistry* **1998**, 10478–10487.
- [21] G. Haeger, J. Bongaerts, P. Siegert, *Anal Biochem.* **2022**, *654*, 114819.
- [22] B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch et al., *J. Comp. Chem.* **2009**, *30*, 1545–1614.
- [23] Y.-F. Lin, C.-W. Cheng, C.-S. Shih, J.-K. Hwang, C.-S. Yu, C.-H. Lu, *J Chem Inf. Model.* **2016**, *56*, 2287–2291.
- [24] J. K. Gagnon, S. M. Law, C. L. Brooks, *J Comp. Chem.* **2016**, *37*, 753–762.

## 2.5. Chapter V

### **Novel aminoacylases from *Streptomyces griseus* DSM 40236 and their recombinant production in *Streptomyces lividans***

Gerrit Haeger, Johanna Probst, Karl-Erich Jaeger, Johannes Bongaerts, Petra Siegert

FEBS Open Bio 13, 12 (2023)

#### Author contributions:

GH designed the study. GH conducted cloning and bioinformatic analysis. GH and JP performed the experiments and analyzed the data for protein expression, purification, and biochemical characterization. GH and JP wrote the manuscript. GH, JB, KEJ and PS edited the manuscript. PS and JB supervised the work of GH and JP. PS and JB did funding acquisition. All authors read and approved the final manuscript.

Overall contribution GH: 85 %

## Novel aminoacylases from *Streptomyces griseus* DSM 40236 and their recombinant production in *Streptomyces lividans*

Gerrit Haeger<sup>1</sup> , Johanna Probst<sup>1</sup>, Karl-Erich Jaeger<sup>2,3</sup>, Johannes Bongaerts<sup>1</sup>  and Petra Siegert<sup>1</sup>

<sup>1</sup> Institute of Nano- and Biotechnologies, Aachen University of Applied Sciences, Jülich, Germany

<sup>2</sup> Institute of Molecular Enzyme Technology, Heinrich Heine University Düsseldorf, Jülich, Germany

<sup>3</sup> Institute of Bio- and Geosciences IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, Jülich, Germany

### Keywords

acyl amino acids; recombinant expression; *Streptomyces griseus*; *Streptomyces lividans*;  $\alpha$ -aminoacylase;  $\epsilon$ -lysine acylase

### Correspondence

P. Siegert, Institute of Nano- and Biotechnologies, Aachen University of Applied Sciences, Heinrich-Mußmann-Str. 1, Jülich 52428, Germany  
 E-mail: siegert@fh-aachen.de

(Received 17 July 2023, revised 15 September 2023, accepted 20 October 2023)

doi:10.1002/2211-5463.13723

Amino acid-based surfactants are valuable compounds for cosmetic formulations. The chemical synthesis of acyl amino acids is conventionally performed by the Schotten–Baumann reaction using fatty acyl chlorides, but aminoacylases have also been investigated for use in biocatalytic synthesis with free fatty acids. Aminoacylases and their properties are diverse; they belong to different peptidase families and show differences in substrate specificity and biocatalytic potential. Bacterial aminoacylases capable of synthesis have been isolated from *Burkholderia*, *Mycolicibacterium*, and *Streptomyces*. Although several proteases and peptidases from *S. griseus* have been described, no aminoacylases from this species have been identified yet. In this study, we investigated two novel enzymes produced by *S. griseus* DSM 40236<sup>T</sup>. We identified and cloned the respective genes and recombinantly expressed an  $\alpha$ -aminoacylase (EC3.5.1.14), designated SgAA, and an  $\epsilon$ -lysine acylase (EC3.5.1.17), designated SgELA, in *S. lividans* TK23. The purified aminoacylase SgAA was biochemically characterized, focusing on its hydrolytic activity to determine temperature- and pH optima and stabilities. The aminoacylase could hydrolyze various acetyl amino acids at the N<sub>α</sub>-position with a broad specificity regarding the sidechain. Substrates with longer acyl chains, like lauroyl amino acids, were hydrolyzed to a lesser extent. Purified aminoacylase SgELA specific for the hydrolysis of N<sub>ε</sub>-acetyl-L-lysine was unstable and lost its enzymatic activity upon storage for a longer period but could initially be characterized. The pH optimum of SgELA was pH 8.0. While synthesis of acyl amino acids was not observed with SgELA, SgAA catalyzed the synthesis of lauroyl-methionine.

N-acyl-L-amino acids are valuable compounds and are used as biosurfactants in cosmetic products. Several advantages of N-acyl-L-amino acids compared with

conventional surfactants like laureth sulfates include more skin-friendly properties, a low inflammatory potential and biodegradability [1]. Long-chain acyl

### Abbreviations

EDTA, ethylenediaminetetraacetic acid; ELSD, evaporative light scattering detector; HiDapE, N-succinyl-L,L-diaminopimelic acid desuccinylase from *Haemophilus influenzae*; LB, Lysogeny Broth; MS, mass spectrometry; MWCO, molecular weight cutoff; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; SamAA, aminoacylase from *S. ambofaciens*; SamELA,  $\epsilon$ -lysine acylase from *S. ambofaciens*; SDS, sodium dodecyl sulfate; SgAA, aminoacylase from *S. griseus*; SgELA,  $\epsilon$ -lysine acylase from *S. griseus*; SmAA, aminoacylase from *S. mobaraensis*; SmELA,  $\epsilon$ -lysine acylase from *S. mobaraensis*; TB, Terrific Broth; Tris, tris(hydroxymethyl)aminomethane; UV, ultraviolet.

2224

FEBS Open Bio 13 (2023) 2224–2238 © 2023 The Authors. FEBS Open Bio published by John Wiley & Sons Ltd on behalf of Federation of European Biochemical Societies.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

amino acids can even have physiological function due to their structural resemblance to endocannabinoids [2], and acetyl amino acids can be found in the brain [3,4]. With the motivation to establish a biocatalytic production of these biosurfactants, aminoacylases have been investigated and show high potential to replace the environmentally harmful Schotten–Baumann synthesis. Aminoacylases that were shown to be capable of synthesis of N-acyl-L-amino acids have been identified in *Streptomyces* [5–7], *Mycolicibacterium* [8], *Burkholderia* [9], and pig liver [10,11].

Several  $\alpha$ -aminoacylases with broad substrate spectrum have been identified that belong to the family of M20A peptidases. Members of this family share a similar fold and three-dimensional structure, albeit with low homology of amino acid sequences. Often, these enzymes are dimeric [12]. The protein structure can be divided into a catalytic and a dimerization domain, or lid domain for monomeric M20A peptidases. The binuclear active site is composed of metal-binding and catalytic residues. The metal ions are bound by conserved residues, namely two histidines, two glutamic acids, and aspartic acid. The catalytic residues are aspartic acid and glutamic acid, and the latter acts as a general base catalyst in hydrolysis or synthesis. Furthermore, a conserved histidine contributes to the formation of an oxyanion hole, stabilizing the tetrahedral reaction intermediate [13]. This histidine is located at the tip of the dimerization domain and reaches into the active site of the opposing dimer. In monomeric members of the M20A family, the lid domain structurally represents a doubled dimerization domain, and the histidine protrudes into the active site of the same monomer [14,15]. The aminoacylases SmAA and SamAA, both nonpeptidase homologs of the M20A family, have been isolated from *S. mobaraensis* [16] and *S. ambofaciens* [5,17], respectively. Especially SamAA showed a high potential for acylation of various amino acids. For SmAA, no synthesis has been shown yet [16]. The aminoacylase from *M. smegmatis* (MsAA) is homologous to SmAA and SamAA and has been shown to catalyze acylation of amino acids as well [8]. Further aminoacylase members of the M20A peptidase family are pAcy1 and hAcy1 from porcine and human liver, respectively [18,19]. The homologous N-succinyl-L,L-diaminopimelic acid desuccinylase from *Haemophilus influenzae* (HiDapE) has also been extensively investigated providing insights into the characteristics of this family [13,20–23].

From *S. mobaraensis* IFO 13819, *S. coelicolor* A3 (2) and *S. ambofaciens* ATCC 23877,  $\epsilon$ -lysine acylases have been identified as well, designated SmELA [6], ScELA [24], and SamELA [5], respectively. While SmAA and

SmELA have been recombinantly expressed in *S. lividans* TK24 [16,25], SamAA and SamELA were obtained from the natural producer, despite efforts to express the enzyme in *E. coli* Origami B(DE3) [26]. The recombinant expression of ScELA with *E. coli* JM109 and *C. glutamicum* YDK010 has been described in a patent [24]. The synthetic potential of SmELA is extraordinarily high, with N $\epsilon$ -lauroyl-L-lysine being synthesized with conversion rates reaching 100% [25]. Using ScELA-containing cell extract from recombinant *C. glutamicum* to synthesize N $\epsilon$ -lauroyl-L-lysine, conversions reaching 100% were also achieved [24]. In contrast to the aminoacylases belonging to the M20A family,  $\epsilon$ -lysine acylases have not been investigated as extensively, and no structure has been published. Further  $\epsilon$ -lysine acylases or  $\epsilon$ -peptidases have been identified from *Achromobacter pestifer* [27,28] and avian kidney [29] but the genes have not been cloned or sequenced. Classification of SmELA to the YtcJ-like metal-dependent amidohydrolase family has been described [25].

Several hydrolytic enzymes have been described from *S. griseus* that are highly interesting for technical applications [30–33]. A commercial preparation of proteinases and peptidases from this species is available under the name Pronase® [34]. An aminopeptidase from *S. griseus* was described, which belongs to the M28 family of metallopeptidases [35]. However, extracellular peptidases from *S. griseus* could not hydrolyze chloroacetyl amino acids [36]. Surprisingly, no aminoacylases have yet been described from *S. griseus*. The goal of this study was to identify, recombinantly express, and characterize homologous aminoacylases from *S. griseus* as a reference to the enzymes identified in *S. mobaraensis* and *S. ambofaciens* and to extend the scope of characterized aminoacylases. We searched for homologs in *S. griseus* DSM 40236<sup>T</sup> (ATCC 23345) as the type strain of this species and identified two enzymes, designated SgAA and SgELA. The respective genes were cloned for recombinant expression in *S. lividans* TK23, and the enzymes were purified and biochemically characterized. Finally, we investigated the acylation activity of the enzymes using all proteinogenic amino acids as substrates.

## Materials and methods

### Chemicals and reagents

Cultivation media, metal salts, amino acids, Tris (tris (hydroxymethyl)aminomethane), and solvents were from Carl Roth (Karlsruhe, Germany). Acetyl amino acids were purchased from Sigma-Aldrich (Taufkirchen, Germany), and other acyl amino acids were chemically synthesized as

described previously [8]. Molecular biology reagents were from Thermo Fisher Scientific (Langerwehe, Germany). Oligonucleotide synthesis and DNA sequencing were ordered from Eurofins Genomics (Ebersberg, Germany). Strep-Tactin columns were purchased from IBA (Göttingen, Germany). Reagents for native PAGE were obtained from SERVA Electrophoresis (Heidelberg, Germany). The EZ Nin reagent was from Biochrom (UK). Remaining chemicals were from Sigma-Aldrich.

### Strains and cultivation media

For cloning and plasmid maintenance, *E. coli* DH5 $\alpha$  (Thermo Fisher Scientific, USA) was used and grown in LB medium. *S. griseus* DSM 40236<sup>T</sup> was grown in liquid LB medium for isolation of genomic DNA. Heterologous expression was performed with *S. lividans* TK23 (spc-1 SLP2<sup>-</sup> SLP3<sup>-</sup>) [37] and *E. coli* BL21(DE3; Thermo Fisher Scientific, USA), *E. coli* BL21(DE3) pGro7 (transformed with GroEL/S overexpression plasmid from Takara Bio Europe, France), and *E. coli* ArcticExpress (DE3; Agilent Technologies, Santa Clara, CA, USA). Tryptic Soy Broth (TSB, MP Biomedicals, Eschwege, Germany) agar plates were used for strain maintenance and cultivation of *S. lividans* TK23. Shake flask cultures of *S. lividans* TK23 for expression and protoplast transformation were grown YEME medium [0.3% yeast extract, 0.5% peptone, 0.3% malt extract, 1% glucose, 34% sucrose, 0.2% magnesium chloride hexahydrate, all (w/v), and 25 mL of a 20% glycerol solution]. For bioreactor cultivations of *S. lividans* TK23, a fermentation medium (0.5% yeast extract, 3% TSB, 0.75% glucose, 0.3% malt extract, 10% sucrose 0.1% magnesium chloride hexahydrate) was developed. Oatmeal agar plates (20 g·L<sup>-1</sup> ground oatmeal, 10 g·L<sup>-1</sup> malt extract, 5 g·L<sup>-1</sup> yeast extract, 20 g·L<sup>-1</sup> agar) were used as sporulation plates. Preparation of spore suspensions for liquid media inoculation was performed by adding 10 mL of 0.1% Tween 80 (Carl Roth)/0.9% NaCl solution to the oatmeal agar plate. The suspension was filtered through glass wool (Carl Roth) to separate mycelium from the spores. The spore suspension was centrifuged at 4000 *g* for 10 min, and the pellet was resuspended in 0.9% NaCl solution. Expression cultures of *E. coli* BL21(DE3) were grown in Terrific Broth medium (TB; 2% tryptone from casein, 2.4% yeast extract, 25 mM NaH<sub>2</sub>PO<sub>4</sub>, 25 mM KH<sub>2</sub>PO<sub>4</sub>, 50 mM NH<sub>4</sub>Cl, 2 mM MgSO<sub>4</sub>, 5 mM Na<sub>2</sub>SO<sub>4</sub>, 0.5% glycerol (v/v), and 0.05% glucose).

### Database searches and sequence analysis

Homologs of aminoacylases SmAA, SamAA, SmELA, and SamELA were searched in *S. griseus* DSM 40236<sup>T</sup> using the BLASTp service from NCBI (<https://blast.ncbi.nlm.nih.gov/>) [38]. Pairwise protein sequence alignment was conducted with the Needleman–Wunsch algorithm using the EMBOSS Needle tool from EMBL-EBI ([https://www.ebi.ac.uk/Tools/psa/emboss\\_needle/](https://www.ebi.ac.uk/Tools/psa/emboss_needle/)) [39].

Multiple protein sequence alignment was performed with the T-Coffee algorithm (<https://www.ebi.ac.uk/Tools/msa/tcoffee/>) [40], and the results were visualized using ESPript 3.0 (<https://esprict.ibcp.fr/ESPript/ESPript/>) [41]. The classification of protein sequences in the MEROPS system was performed by MEROPS BLAST [42].

### Cloning of aminoacylase genes from *S. griseus* DSM 40236<sup>T</sup>

For isolation of genomic DNA, *S. griseus* DSM 40236<sup>T</sup> was grown and DNA was isolated with the innuPREP Bacteria DNA Kit (Analytik Jena, Jena, Germany) according to the manufacturer's instructions. Polymerase chain reaction (PCR) was performed on genomic DNA to amplify aminoacylase gene sequences using Phusion High-Fidelity PCR Master Mix with GC-buffer. The genes for SgAA and SgELA were amplified using primer pairs P1 (ATGAGCGAGAGCAGCACGGG) & P2 (TCAGGAGTGGTCGATGAACCGG) and P3 (ATGAGCCAGAGCACGCC) & P4 (TCACTCGTTCGGTTCGCACGTAG), respectively. Further PCRs were performed to attach sequences for a Strep-tag II (WSHPQFEK) with a linker (SG), either N- or C-terminally, and restriction sites. For SgAA NTag, primers P5 (CGCAGTTCGAGAAGTCCGGCATGAGCGAGAGCAGCACGGG) & P6 (ATCGAATTCAGGAGTGGTCGATGAACCGG) and P7 (GATGTAGCATGTGGTCCACCCGAGTTCGAGAAGTCCGGC) & P6 were used. For SgAA CTag, P8 (GATGCTAGCAGGAGAGCAGCACGGG) & P9 (TGCGGGTGGACCAGCCGGAGGAGTGGTCGATGAACCGGTCG) and P8 & P10 (ATCGAATTCcaCTTCTCGAACTGCGGGTGGGACCAGCC) were used. For SgAA without affinity tag, P8 & P6 were used. For SgELA NTag, P11 (CGCAGTTCGAGAAGTCCGGCATGAGCCAGACCCGCC) & P12 (ATCGAATTCCTACTCGTTCGGTTCGCACGTAG) and P7 & P12 were used. For SgELA CTag, P13 (GATGCTAGCAGGAGCAGCACGCC) & P14 (TGCGGGTGGGACCAGCCGGACTCGTTCGGTTCGCACGTAGACC) and P13 & P10 were used. For SgELA without affinity tag, P13 & P12 were used. The genes were cloned using NheI and EcoRI restriction sites and T4 ligase into pGH01. The pEM4-based [43] plasmid pGH01 is a shuttle vector for *E. coli* and *S. lividans* and carries the *ermE\** promoter from *S. erythraea* for heterologous expression [44]. The plasmid conveys resistances against ampicillin and thiostrepton for *E. coli* and *S. lividans*, respectively. The resulting plasmids were designated pGH01 SgAA (NTag/CTag/noTag) and pGH01 SgELA (NTag/CTag/noTag).

The aminoacylases from *S. griseus* DSM 40236<sup>T</sup> were also cloned for expression in *E. coli*. For this, the encoding nucleotide sequences were deduced and codon-optimized according to the *E. coli* codon usage. Sequences encoding for the

Strep-tag II and linker were attached to both termini as well. The resulting DNA sequence was commercially ordered and synthesized by GeneArt (Thermo Fisher Scientific, USA). As described previously, the genes were amplified using primers with BsaI overhangs for Golden Gate cloning into pET28-eforRED [8]. The primers used for amplification were P14 (GGTCTCCCATGTGGAGTCATCTCAATTCGAAAAATCC) and P15 (GGTCTCTCTCAGCTATGATCAATAAGCGATCCAGCACG) for SgAA Ntag, P16 (GGTCTCCCATGAGCGAAAGCAGCACCG) and P17 (GGTCTCTCTCATTTTTCGAATTGAGGATGACTCCATCC) for SgAA Ctag, P16 and P15 for SgAA without tag, P14 and P18 (GGTCTCTCTCATTCATTCGGACGAACATAAA CGGTCTG) for SgELA Ntag, P19 (GGTCTCCCATGACCCAGAGCACCGCAC) and P17 for SgELA Ctag, P19 and P18 for SgELA without tag.

### Transformation of *S. lividans* TK23

Protoplast transformation of *S. lividans* TK23 was performed according to the standard procedure presented by Kieser *et al.* [37]. Thiostrepton (Merck Millipore, Darmstadt, Germany) was used at 30  $\mu\text{g}\cdot\text{mL}^{-1}$  for selection. Isolation of plasmid DNA to verify successful transformation from *S. lividans* TK23 was performed with modifications according to the protocol of Thompson *et al.* [45]. A spore suspension was inoculated into 25 mL of YEME medium, and the culture was incubated at 180 rpm and 30 °C for 3 days. The cells were centrifuged at 8000 *g* for 2 min and washed in 10% sucrose, 100 mM glucose and 25 mM Tris-HCl pH 7.0. After another centrifugation step, the pellet was incubated in the above buffer containing 1  $\text{mg}\cdot\text{mL}^{-1}$  lysozyme (SERVA Electrophoresis, Germany) and 100  $\mu\text{g}\cdot\text{mL}^{-1}$  RNase (Carl Roth, Germany) for 1 h at 37 °C followed by plasmid purification with the GeneJET Plasmid-Miniprep-Kit (Thermo Fisher Scientific, USA).

### Recombinant expression in *E. coli*

Recombinant expression in *E. coli* was conducted with *E. coli* BL21(DE3), *E. coli* BL21(DE3) pGro7, and *E. coli* ArcticExpress(DE3) as previously described [8]. TB medium was used for growth. For induction with isopropyl  $\beta$ -D-thiogalactoside (IPTG), 1 mM IPTG was added at an  $\text{OD}_{600}$  of 0.5 and cells were harvested 4 h after induction (30 °C or 37 °C). For autoinduction, 0.2% (w/v) lactose was added to the TB medium, and the cultures were grown for 24 h at 20 °C or 30 °C. When *E. coli* BL21 (DE3) pGro7 was used, 0.5  $\text{mg}\cdot\text{mL}^{-1}$  arabinose was added to the medium. *E. coli* ArcticExpress (DE3) was first cultured for 3 h or 6 h at 30 °C and subsequently grown at 12 °C for further 24 h. No antibiotics were used for this strain. The cells were harvested at 4000 *g* and 4 °C for 30 min.

### Recombinant expression in *S. lividans* TK 23

First, recombinant expression was performed in shake flasks. Recombinant *S. lividans* TK23 was cultivated in 500 mL baffled flasks containing 100 mL YEME medium complemented with 10  $\mu\text{g}\cdot\text{mL}^{-1}$  thiostrepton and 2  $\mu\text{L}$  polyethylene glycol (PEG 2000). The medium was inoculated with a fresh spore suspension and a starting  $\text{OD}_{600}$  of 0.02. For expression, cultures were grown at 30 °C on a shaker at 180 rpm. Harvesting was performed 3 days after inoculation by centrifugation at 3000 *g* for 15 min at 4 °C.

Second, expression was performed in 1.3 L benchtop-scale bioreactors filled with 1.0 L of fermentation medium. As a starter culture, recombinant *S. lividans* TK23 was grown in shake flask in fermentation medium for 3 days with 10  $\mu\text{g}\cdot\text{mL}^{-1}$  thiostrepton at 30 °C and agitated at 180 rpm. After 3 days, the preculture was inoculated overnight from the starter culture at an  $\text{OD}_{600}$  of 0.02. Cultivation in the fermenter was performed in a DASGIP Parallel Bioreactor (Eppendorf, Hamburg, Germany). The fermentation medium was inoculated with an  $\text{OD}_{600}$  0.1 with the preculture. The cultivation temperature was 30 °C, and dissolved oxygen was set to 30%, regulated by the stirrer speed. Before 30% dissolved oxygen was reached, the stirrer was set to 350 rpm. The pH was monitored online and maintained at pH 7.0 with 2 M NaOH and 20% (v/v)  $\text{H}_2\text{SO}_4$ . The culture was fed with a 1.5  $\text{g}\cdot\text{mL}^{-1}$  glucose solution. The feed was initiated after 15 h of incubation at a rate of 1  $\text{mL}\cdot\text{h}^{-1}$  for 7 h. Subsequently, after 22 h of cultivation, the flow rate was increased to 1.5  $\text{mL}\cdot\text{h}^{-1}$  until the end of the fermentation. Polypropylene glycol (PPG 2000) was added as an antifoaming agent, regulated by the foam sensor. Fermentation lasted for 2 days for expression of SgAA and 5 days for expression of SgELA. The cultures were harvested by centrifugation for 30 min at 4 °C and 3000 *g*.

### Protein purification, concentration measurement, and electrophoresis

The harvested cells were resuspended in 100 mM Tris-HCl 150 mM NaCl pH 7.0 to a final volume of 400 mL. The cell lysis was performed by sonication on ice with a 45-s pulse followed by a 60-s pause, repeated four times. The cell debris was separated by centrifugation at 16 000 *g* for 30 min at 4 °C. The recombinant aminoacylases were purified by Strep-tag II affinity purification. The column was a 5 mL Strep-Tactin® SuperFlow® high capacity cartridge. The buffer used for column wash was 100 mM Tris-HCl 150 mM NaCl pH 7.0. For elution of the recombinant protein, 2.5 mM D-desthiobiotin was added to the wash buffer. The relevant elution fractions were pooled and rebuffered to the wash buffer without desthiobiotin in Vivaspin™ 6 concentrators (10 000 MWCO; Sartorius, Göttingen, Germany). The protein concentrations were measured with the Bradford method [46] using the Roti®-Nanoquant reagent

(Carl Roth, Germany). Gradient gels from 8% to 20% acrylamide content were poured for SDS-polyacrylamide gel electrophoresis (SDS/PAGE) [47]. The protein molecular weight marker was FastGene® BluEasy Protein Marker (Nippon Genetics Europe, Düren, Germany). Blue native PAGE was performed with SERVAGel™ N4-16% gels (SERVA, Germany), with reagents from SERVA according to the manufacturer's instructions. The protein marker was SERVA Native Marker, Liquid Mix for BN/CN, and further reference proteins were from Gel Filtration Calibration Kit (High Molecular Weight) from Cytiva (Marlborough, MA, USA). The gels were stained with Coomassie from Rotiphorese® solution (Carl Roth).

### Aminoacylase activity assay

The aminoacylase activity was measured by quantification of the released amino acids with the ninhydrin reaction as described previously [48]. The hydrolytic reactions consisted of 190  $\mu\text{L}$  substrate solution and 10  $\mu\text{L}$  enzyme solution. The reaction course was followed by withdrawing 10  $\mu\text{L}$  samples from the reaction and mixing with 100  $\mu\text{L}$  of EZ Nin:DMSO reagent, heating for 10 min at 99 °C, and diluting 25  $\mu\text{L}$  of the colored sample with 225  $\mu\text{L}$  of 100 mM Na-borate pH 10.0 for measurement. Standard hydrolytic activity assay for SgAA was performed with 15 mM N-acetyl-L-alanine or N-acetyl-L-methionine in 100 mM Tris-HCl pH 7.0 at 30 °C. The pH of all substrate solutions was measured and adjusted at the respective temperatures. The ninhydrin reaction products of all amino acids were measured at 570 nm, except for the reactions with L-cysteine and L-proline, which were measured at 410 nm. The hydrolysis of N $_{\alpha}$ -acetyl-L-lysine and N $_{\epsilon}$ -acetyl-L-lysine was quantified by reaction with an acidic variant of the ninhydrin reagent mixed with glacial acetic acid (EZ Nin:GAA) and was measured at 460 nm [48]. The hydrolysis of dipeptides was measured in the same way, except a substrate concentration of 7.5 mM was used. One unit of SgAA was defined as the amount of enzyme that hydrolyzes one  $\mu\text{mol}$  of N-acetyl-L-alanine per minute under the given conditions, and one unit of SgELA hydrolyzes one  $\mu\text{mol}$  of N $_{\epsilon}$ -acetyl-L-lysine per minute.

### Biochemical characterization of the aminoacylases

The dependency of SgAA activity on zinc or cobalt ions was investigated by incubating the purified enzyme with 1 mM  $\text{CoCl}_2$  or  $\text{ZnCl}_2$  in 100 mM Tris-HCl at pH 7.0 and 30 °C for 1 h. Likewise, the enzyme was also incubated with the chelating agent ethylenediaminetetraacetic acid (EDTA) at 1 mM concentration. After the incubation period, the hydrolytic activity was determined with 15 mM N-acetyl-L-methionine in 100 mM Tris-HCl pH 7.0 at 30 °C.

The hydrolytic substrate specificity of SgAA was determined with various 15 mM substrates in 100 mM Tris-HCl pH 7.0 with a reaction time of 5 min. The reaction temperature

was increased to 50 °C, so that all substrates were solubilized. The substrate specificity of SgELA was determined with the same acetyl-L-amino acids as for SgAA at 15 mM in 100 mM Tris-HCl pH 7.0 at 30 °C with a reaction time of 1 h.

The temperature optimum of SgAA was determined by hydrolysis of 15 mM N-acetyl-L-methionine at various temperatures from 20 °C to 70 °C at pH 7.0 in 100 mM Tris-HCl buffer. The temperature stability was assessed by incubation of the enzyme solution at temperatures from 4 °C to 80 °C in the same buffer at the respective temperatures. After 1-h incubation, residual activity was measured with 15 mM N-acetyl-L-methionine.

The pH dependency of hydrolytic activity of SgAA was determined using the following buffers at 100 mM concentration: Na-acetate for pH 4.0–5.0, Na-MES for pH 6.0–7.0, Tris-HCl for pH 6.0–9.0, and Na-borate for pH 9.0–11.0. The substrate solutions consisted of 15 mM N-acetyl-L-methionine dissolved in these buffers and adjusted at 30 °C. The optimal pH for hydrolysis with SgELA was investigated by preparing 15 mM N $_{\alpha}$ -acetyl-L-lysine in 100 mM Tris-HCl adjusted to pH 6.0–10.0 at 30 °C.

### Biocatalytic synthesis of lauroyl amino acids

For the synthesis of N-lauroyl-L-amino acids, 100 mM lauric acid and 100 mM of all proteinogenic L-amino acids in 100 mM Tris-HCl buffer at pH 7.0 were used as substrate. For the reaction, 10  $\mu\text{g}$  of enzyme per 0.5 mL reaction was used and the reaction was carried out at 40 °C. The reaction mixtures were analyzed with an HPLC system (S5200 and S2100; Sykam, Fürstenfeldbruck, Germany) equipped with a C18 column (ISAspher 100-5 C18 BDS column, 5  $\mu\text{m}$ , 4.0  $\times$  250 mm; Isera, Düren, Germany), at a column temperature of 40 °C with the flow rate set to 1 mL·min $^{-1}$ . The liquid phase consisted of 80% acetonitrile: 20% H $_2$ O, 0.1% trifluoroacetic acid, and an isocratic elution was performed. The analytes were detected with a UV Detector (2500 Sykam, Germany) at 210 nm and an ELSD detector (evaporative light scattering detector; ZAM 4000, Schambeck SFD, Bad Honnef, Germany). HPLC-MS analysis was conducted with a Shimadzu Nexera XR system equipped with a Hitachi LaChrom II column (C18, 5  $\mu\text{m}$ , 4.6  $\times$  250 mm) and a Shimadzu LCMS-2020-mass spectrometer. The column temperature was set to 40 °C. A gradient elution was applied starting at 20% acetonitrile and 80% water and 0.1% trifluoroacetic acid and going to 100% acetonitrile in 10 min, which was held for 6 min.

## Results and Discussion

### Cloning and sequence analysis of SgAA and SgELA

Homologs of the aminoacylases SmAA, SamAA, SmELA, and SamELA were found in the genome of

*S. griseus* DSM 40236<sup>T</sup>. The protein sequences were used as baits and the putative aminoacylases SgAA (Accession No. WP\_003970135.1) and SgELA (Accession No. WP\_069631407.1) were identified with NCBI BLASTp search. The genes were amplified by PCR using genomic DNA from *S. griseus* DSM 40236<sup>T</sup> as a template and were subsequently cloned into the *S. lividans* expression vector pGH01 and fused with an N-terminal Strep-tag II. Sequence identities of SgAA to SamAA (Accession No. AKZ54783.1), SmAA (Accession No. BAI44523.1), MsAA (Accession No. AWT55079.1), pAcyl (Accession No. NP 999061.1), and HiDapE (Accession No. WP 005693818.1) are 88.3%, 83.7%, 56.5%, 26.5%, and 23.7%, respectively, as determined by the Needleman–Wunsch algorithm [39]. Protein sequence alignment of SgAA with its homologs allowed to identify conserved residues in SgAA, with metal-binding (H90, D122, E157, E184, and H418) and catalytic (D92, E156, and H221) function (Fig. 1). By using the MEROPS BLAST tool, SgAA can be assigned to the M20A peptidase family in accordance with the homologous enzymes and the identification of conserved residues.

The sequence of SgELA shows identity to ScELA (GenBank HV688803.1), SamELA (Accession No. WP 053127917.1), and SmELA (Accession No. WP 053127917.1), *Burkholderia* sp. aminoacylase (Accession No. BBI47489.1), amidohydrolase Sgx9260b (PDB 3MKV), and prolidase Sgx9260c (PDB 3N2C) [49] with 79.8%, 79.6%, 75.6%, 21.0%, 21.2%, and 17.8%, respectively. The sequences of these proteins were subjected to MEROPS BLAST, which suggested their classification as M38 metallopeptidases. A protein sequence alignment revealed conserved residues of SgELA for metal-binding (H76, H78, H330, H365, and D431). The residue H177 of SgELA is conserved among SgELA, *Burkholderia* sp. aminoacylase (Accession No. BBI47489.1), amidohydrolase Sgx9260b (3MKV), and prolidase Sgx9260c (3N2C) and has been described as oxyanion hole-forming [49]. The tyrosine residue of Sgx9260b and Sgx9260c, which has been assigned the function of binding the  $\alpha$ -carboxylic group of the amino acid substrate, is not conserved in SgELA and the other  $\epsilon$ -lysine aminoacylases (Fig. 2). This might contribute to the missing or low  $\alpha$ -aminoacylase activity of these enzymes. The DNA and protein sequences of the enzymes are shown in the Appendix S1.

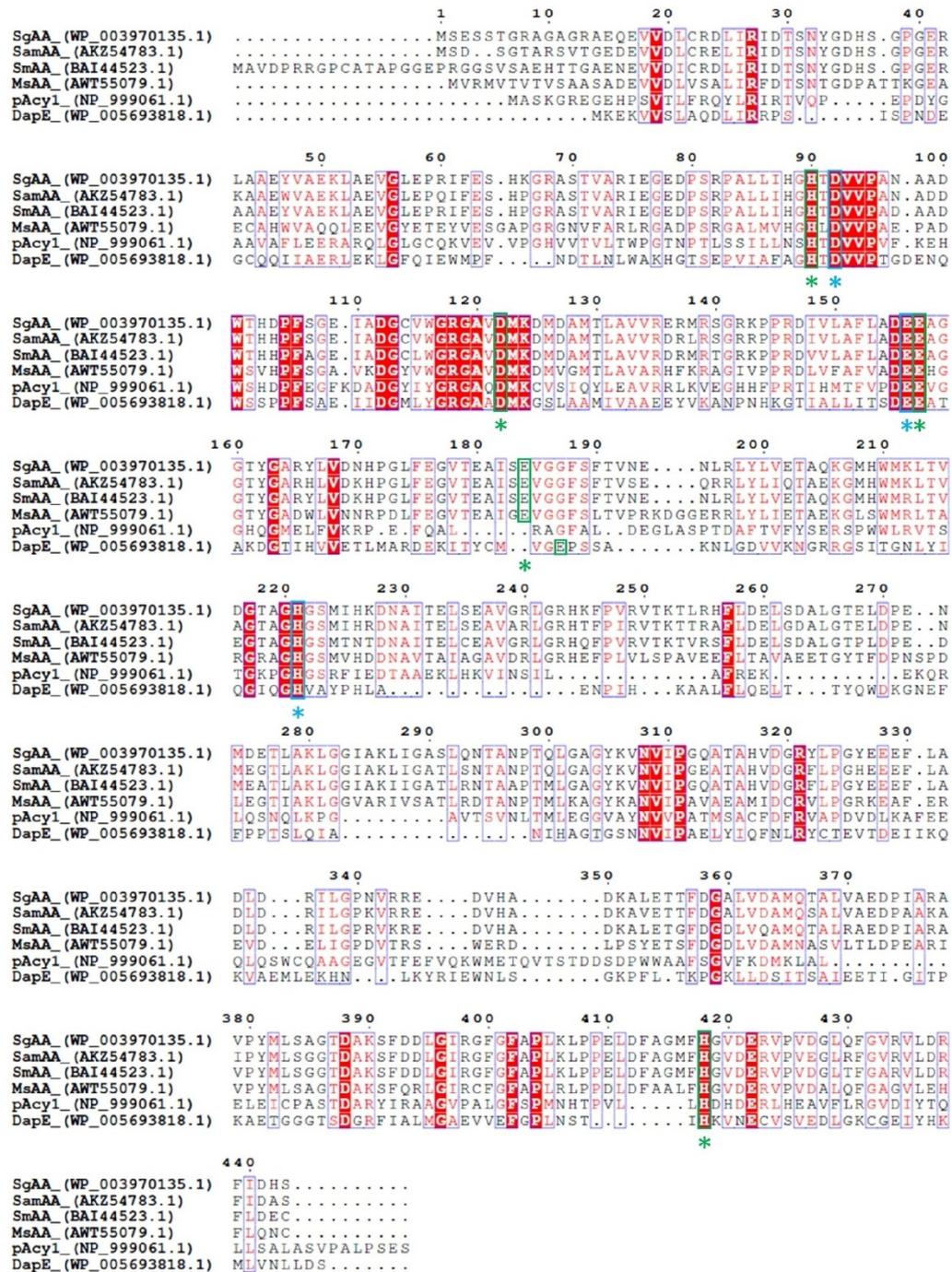
#### Expression of SgAA and SgELA in shake flasks with *E. coli* and *S. lividans*

The recombinant production of the novel aminoacylases SgAA and SgELA caused difficulties in obtaining

soluble enzymes; the enzymes were first expressed in recombinant *E. coli* BL21(DE3) after codon-optimization according to the bias of *E. coli*. Both enzymes were very prone to the formation of inclusion bodies, regardless of whether a Strep-tag was attached or not. Overexpression in *E. coli* resulted in the formation of a significant amount of protein as detected by SDS/PAGE, but the recombinant enzymes were present in the insoluble fraction, and no soluble aminoacylase could be obtained by affinity purification (data not shown). This was the case for IPTG induction or lactose autoinduction at all tested temperatures (20–37 °C). The co-expression of GroEL/S did not lead to soluble protein either. Neither did the use of *E. coli* ArcticExpress (DE3), which constitutively expresses the cold-adapted chaperonins Cpn60/10 from *Oleispira antarctica*. In a study on the expression of the aggregation-prone MsAA from *M. smegmatis*, a homolog of SgAA, the abovementioned approaches led to improvement of soluble expression [8]. This leads to the conclusion that the aminoacylases from *S. griseus* DSM 40236<sup>T</sup> are less suitable for production in *E. coli*. Hence, *S. lividans* TK23 was used as an alternative production host that is genetically closer related to the natural producer. However, the expressed aminoacylases were barely visible as bands after SDS/PAGE and no differences between the Strep-tag variants were observed. Still small amounts of SgAA NTag and SgELA NTag could be purified via affinity chromatography (not shown). Since the productivity for the aminoacylases was low, a scale-up from shake flasks to bioreactors was performed.

#### Expression and purification of SgAA and SgELA in bioreactors with *S. lividans* TK23

The aminoacylases were eventually produced with recombinant *S. lividans* TK23 transformed with pGH01 SgAA NTag or pGH01 SgELA NTag in bioreactors. After 45 h of growth, SgAA was isolated and purified. The enzyme activity in the cell-free extract was 11.8 U·mL<sup>-1</sup> (15 mM N-acetyl-L-alanine, 30 °C, pH 7.0). The cell-free extract was subjected to Strep-tag II affinity purification, and the purified enzyme had a specific activity of 65.0 U·mg<sup>-1</sup> against 15 mM acetyl-alanine (30 °C, pH 7.0). The SDS/PAGE analysis of the purification of SgAA is shown in Fig. 3, and activities and protein concentrations throughout the purification are summarized in Table S1. In relation to the total protein content, purified recombinant SgAA represents 0.56% of total cellular protein in the cell-free extract. For SmAA from *S. mobaraensis*, enzyme purified from native producer constituted 0.02% of total protein, while



**Fig. 1.** Multiple sequence alignment of SgAA and homologous proteins. SgAA from *S. griseus* (WP 003970135.1), SamAA from *S. ambofaciens* (AKZ54783.1), SmAA from *S. mobaraensis* (BAI44523.1), MsAA from *M. smegmatis* (AWT55079.1), pAcy1 from *sus scrofa* (NP 999061.1), and DapE from *H. influenzae* (WP 005693818.1) are included in the alignment. The alignment was generated with the Clustal W algorithm and displayed with ESPrict 3.0. The conserved metal-binding and catalytic residues are shown with green and blue boxes and asterisks, respectively.

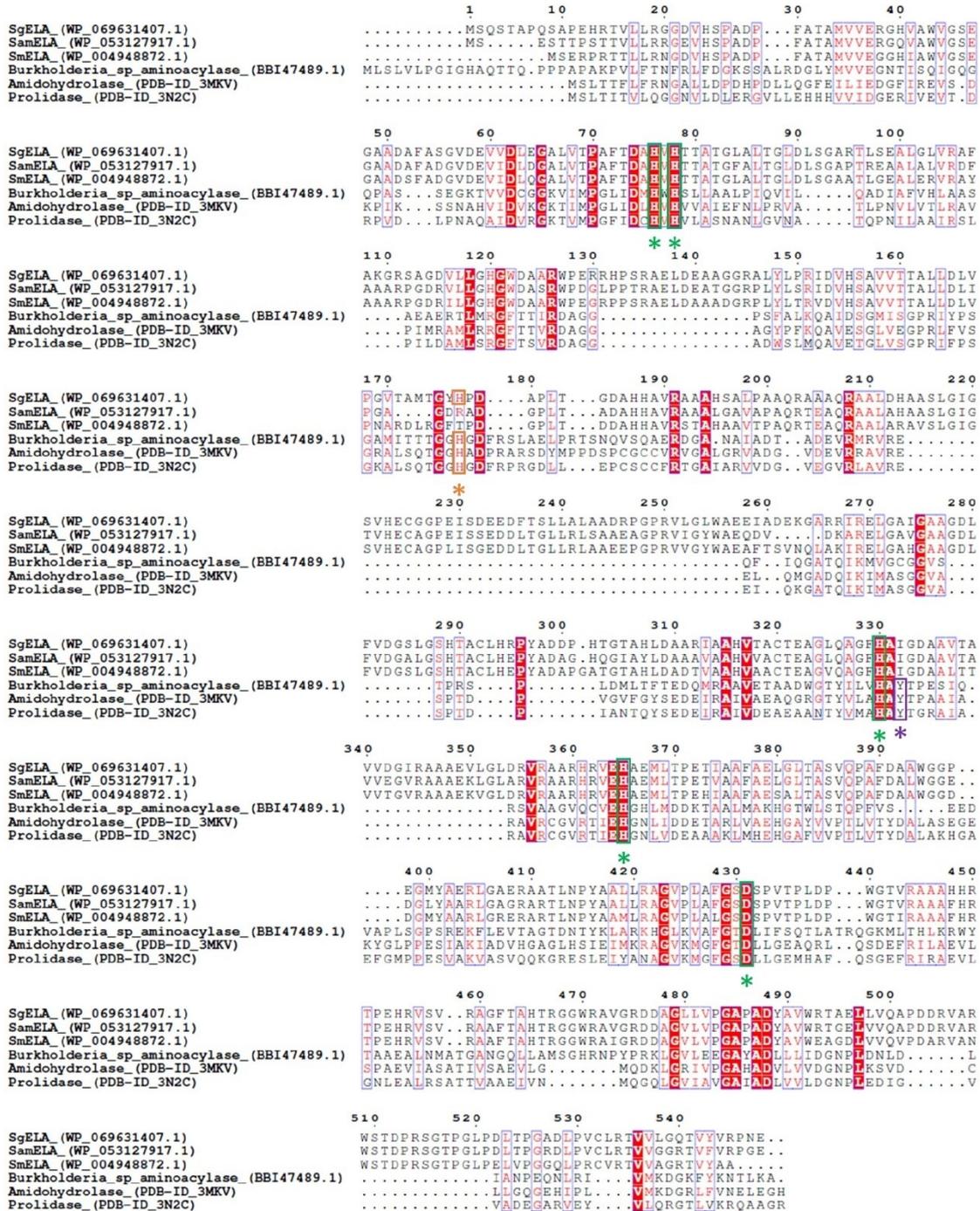
## 2. Results - 2.5. Chapter V

G. Haeger *et al.*

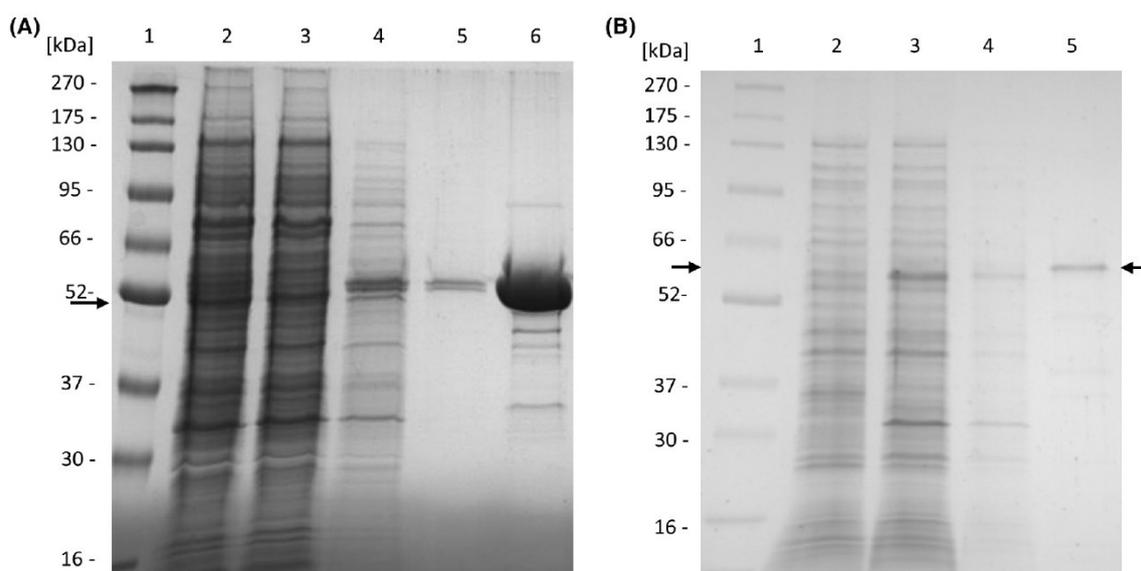
Novel aminoacylases from *Streptomyces griseus*

recombinant SmAA made up for 3.24% of total protein in *S. lividans* TK24 [16]. Purified SgAA was subjected to native PAGE (Fig. S1) and was found to be

a dimer, as it is described for several homologs of the M20A metallopeptidase family [8,13,19]. However, the aminoacylase SmAA from *S. mobaraensis* was



**Fig. 2.** Multiple sequence alignment of SgELA and homologous proteins. SgELA from *S. griseus* (WP\_069631407.1), SamELA from *S. ambofaciens* (WP\_053127917.1), SmELA from *S. mobaraensis* (WP\_004948872.1), aminoacylase from *Burkholderia* species (BBI47489.1), amidohydrolase Sgx9260b (PDB-ID\_3MKV), and prolidase Sgx9260c (PDB-ID\_3N2C) are included in the alignment. The alignment was generated with the Clustal W algorithm and displayed with ESPript 3.0. The conserved metal-binding residues are shown with green boxes and asterisks. The partly conserved oxyanion hole-forming histidine described for the amidohydrolase superfamily is highlighted in orange. The partly conserved tyrosine residue that binds the  $\alpha$ -carboxylic group of amino acid substrates in amidohydrolases is highlighted in purple [49].



**Fig. 3.** SDS/PAGE analysis of fractions obtained during purification of recombinant *S. griseus* aminoacylases. (A) Affinity chromatography of SgAA. Lane 1: Protein marker (BlueEasy Prestained protein Marker, Nippon Genetics); lane 2: cell-free extract; lane 3: flow-through; lane 4 and 5: wash fractions; lane 6: elution of SgAA. (B) Affinity chromatography of SgELA. Lane 1: Protein marker (BlueEasy Prestained protein Marker, Nippon Genetics); lane 2: cell-free extract; lane 3: flow-through; lane 4: wash fraction; lane 5: elution of SgELA.

described to be a monomeric enzyme as determined by native PAGE [16].

The production of SgELA was performed in bioreactors with an extended cultivation time of 144 h, cells were harvested, and SgELA could be purified from the cell-free extract in a small amount of 123  $\mu\text{g}$  at a concentration of 16  $\mu\text{g}\cdot\text{mL}^{-1}$ . The specific activity of purified SgELA was 11.1  $\text{U}\cdot\text{mg}^{-1}$  against 15 mM N $\epsilon$ -acetyl-lysine (30  $^{\circ}\text{C}$ , pH 7.0). The SDS/PAGE analysis of the purification fractions of SgELA is shown in Fig. 3.

#### Biochemical characterization of SgAA and SgELA

The purified aminoacylases were biochemically characterized, investigating the effect of metal ions, screening for hydrolytic activity with various acyl amino acids and studying the effect on pH and temperature on activity and stability. The effect of

divalent zinc or cobalt ions and the chelating agent EDTA on the enzymatic activity of SgAA was investigated with the purified enzyme. The addition of 1 mM  $\text{ZnCl}_2$  and 1 mM  $\text{CoCl}_2$  reduced the specific activity of SgAA to 50.7% and 90.5%, respectively. Despite the addition of 1 mM of the chelator EDTA, enzyme activity was retained with a specific activity to 107.3%. The homologous enzyme SmAA also showed decreased activity from an excess of  $\text{Zn}^{2+}$  ions, and EDTA slightly increased the enzyme activity as well [16]. Given the conserved metal-binding residues of SgAA and the presence of cocatalytic metals in M20A peptidases, it is likely that metal ions are needed for SgAA activity, but at 1 mM ion concentration tested here, an inhibition of aminoacylase activity was observed. The aminoacylase SamAA showed the highest synthetic activity upon the addition of 0.1 mM  $\text{CoCl}_2$ , but decreased activity with 0.1 mM  $\text{ZnSO}_4$  or EDTA [5].

**Table 1.** Substrate specificity of SgAA in the hydrolysis of N-acyl-L-amino acids and dipeptides. N-acyl-L-amino acids and dipeptides were used at 15 mM and 7.5 mM concentrations, respectively, in 100 mM Tris-HCl pH 7.0. The reaction temperature was set at 50 °C to solubilize all substrates.

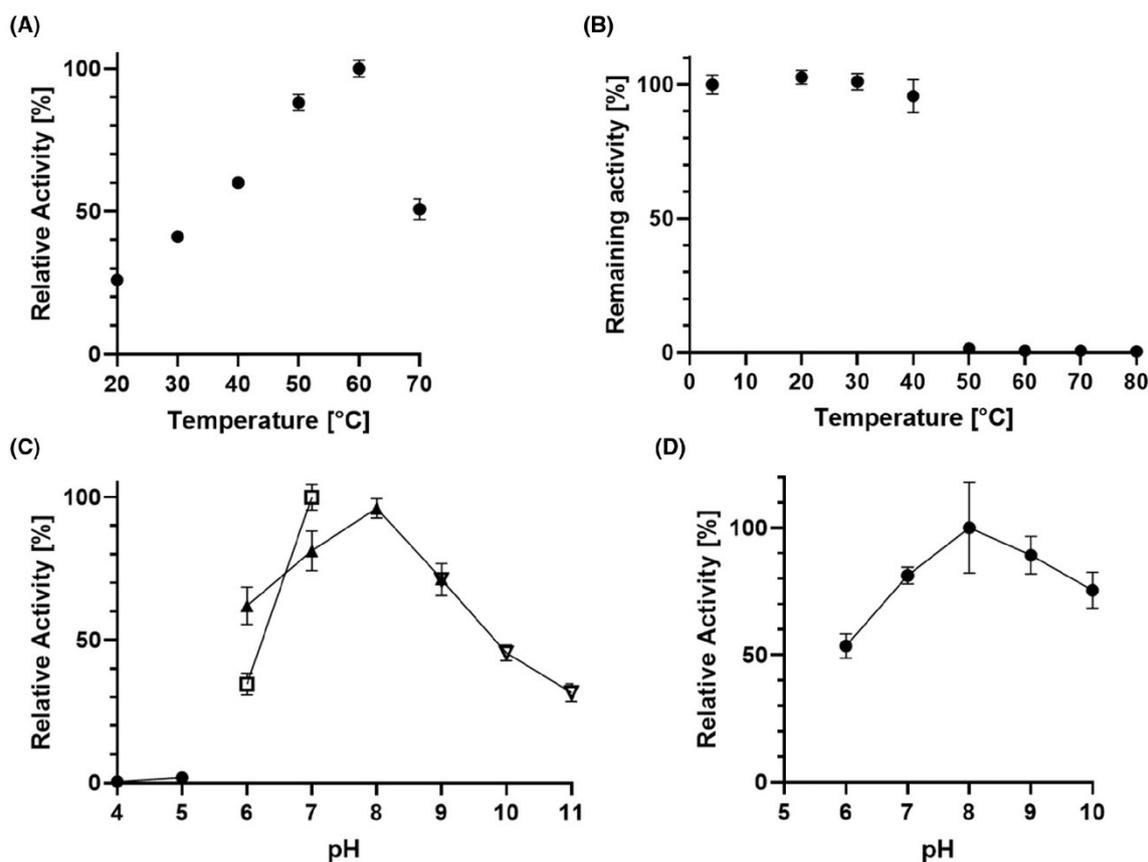
Substrate	Specific activity [U·mg <sup>-1</sup> ]
N-Acetyl-L-methionine	301.1 ± 10.1
N-Acetyl-L-alanine	155.0 ± 4.2
N <sub>α</sub> -Acetyl-L-arginine	99.4 ± 0.6
N-Acetyl-L-leucine	43.4 ± 4.2
N-Acetyl-L-valine	25.7 ± 7.0
N-Acetyl-L-phenylalanine	23.5 ± 3.2
N-Acetyl-L-glutamine	16.7 ± 1.0
N-Acetyl-L-asparagine	14.1 ± 0.5
N-Acetyl-L-glycine	13.3 ± 1.7
N-Acetyl-L-proline	12.0 ± 2.2
N-Acetyl-L-threonine	7.7 ± 0.3
N-Acetyl-L-tyrosine	7.5 ± 0.7
N-Acetyl-L-tryptophan	7.0 ± 0.6
N-Acetyl-L-aspartic acid	5.9 ± 0.3
N-Acetyl-L-isoleucine	5.6 ± 0.5
N-Benzoyl-L-alanine	3.3 ± 0.9
N-Acetyl-L-glutamic acid	2.4 ± 0.2
N-Palmitoyl-L-alanine	2.8 ± 0.9
N-Lauroyl-L-methionine	1.4 ± 0.0
N-Lauroyl-L-alanine	1.1 ± 0.0
N-Caproyl-L-glutamine	0.4 ± 0.0
N-Lauroyl-L-glutamine	0.3 ± 0.2
N-Palmitoyl-L-glutamine	0.2 ± 0.2
N-Lauroyl-L-serine	0
N-Lauroyl-L-glycine	0
N-Lauroyl-L-arginine	0
N <sub>ε</sub> -acetyl-lysine	0
N <sub>α</sub> -acetyl-lysine	0
Phenylalanyl-alanine	0
Alanyl-phenylalanine	0

The substrate specificity of purified SgAA was determined for the hydrolysis of various acyl amino acids with a concentration of 15 mM at 50 °C and pH 7.0 (Table 1). The enzyme could hydrolyze various acetyl amino acids with a broad specificity regarding the amino acid moiety. Highest specific activities were measured for N-acetyl-methionine with 301.1 (± 10.1) U·mg<sup>-1</sup>, N-acetyl-alanine with 155.0 (± 4.2) U·mg<sup>-1</sup>, and N<sub>α</sub>-acetyl-arginine with 94.4 (± 0.6) U·mg<sup>-1</sup>. In general, SgAA shows higher activities for amino acids with hydrophobic residues, but some amino acids with polar or charged side chains were deacylated as well. Interestingly, in contrast to SmAA and MsAA, which do not hydrolyze N-acetyl-proline, SgAA could hydrolyze N-acetyl-proline with an activity of 12.0 (± 2.2) U·mg<sup>-1</sup> under the chosen conditions. Furthermore, SmAA did not hydrolyze acetyl-glutamic acid, acetyl-tryptophane, and acetyl-tyrosine. Comparing to

MsAA, the aminoacylase SgAA showed higher activity with N<sub>α</sub>-acetyl-arginine, N-acetyl-phenylalanine, and N-acetyl-tyrosine, which were hydrolyzed with low activity by MsAA under the chosen conditions. Acetyl-tryptophan was not hydrolyzed by MsAA but was accepted as a substrate by SgAA. The aminoacylase SgAA strongly favors short-chain acyl residues, showing very low hydrolytic activity with the tested caproyl-, lauroyl-, and palmitoyl amino acids. The homologs SmAA and MsAA also favor a short acyl chain length, but activity against various lauroyl amino acids was higher compared with SgAA. No dipeptidase activity of SgAA was detected with alanyl-phenylalanine or phenylalanyl-alanine. The purified aminoacylase SgELA showed activity only against 15 mM N<sub>ε</sub>-acetyl-lysine with 11.1 U·mg<sup>-1</sup> (30 °C, pH 7.0). No activity was detected with 15 mM N<sub>α</sub>-acetyl-lysine or other 15 mM N<sub>α</sub>-acyl amino acids which were tested for substrate scope of SgAA (not shown). Hence, the putative classification as an ε-lysine aminoacylase was confirmed.

The temperature optimum of SgAA was 60 °C with a specific activity of 313 U·mg<sup>-1</sup> tested with 15 mM acetyl-methionine in 100 mM Tris-HCl at pH 7.0. The temperature stability was measured after 24-h incubation at different temperatures and subsequent activity assay with 15 mM acetyl-methionine in 100 mM Tris-HCl at pH 7.0 and 30 °C. The enzyme was stable up to 40 °C, and at 50 °C and higher temperatures, no remaining hydrolytic activity could be detected. The pH optimum of SgAA was pH 7.0–8.0 at 30 °C (Fig. 4). These characteristics of SgAA are similar to its homologs SmAA and MsAA. The optimal temperature and pH for hydrolysis with SmAA were 50 °C and pH 7.0–8.0, respectively [16]. Highest hydrolytic activity with MsAA was measured at 70 °C and pH 7.0 [8]. Both enzymes showed good stability at 40 °C but decreasing activity when incubated at 50 °C or higher.

The pH optimum of SgELA in hydrolysis of 15 mM N<sub>ε</sub>-acetyl-lysine was tested in a range of pH 6.0–10.0 at 30 °C, and the highest activity was measured at pH 8.0. No hydrolytic activity was measured at 50 °C. The enzyme was not stable during storage (in 100 mM Tris-HCl pH 7.0 with 150 mM NaCl) and quickly lost its activity after several hours at 4 °C or -20 °C, so no detailed investigations on stability could be conducted. The homologous enzyme SmELA also had its optimal pH at 8.0 but was significantly more stable, and highest activity was measured at 55 °C [6]. The reasons for the observed instability of SgELA remain unclear, as both SgELA and SmELA were produced with recombinant *S. lividans* cells, and the enzymes are homologs with a sequence identity of 75.6%.



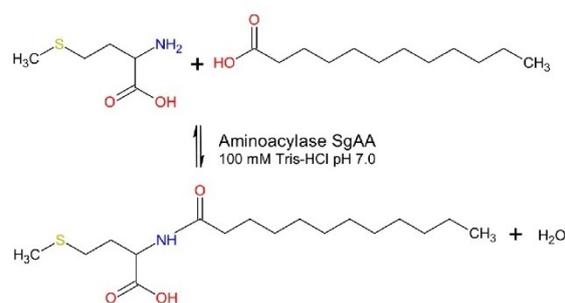
**Fig. 4.** Biochemical characterization of SgAA and SgELA. (A) Temperature optimum of SgAA. Activity was measured at 20–70 °C with 15 mM acetyl-methionine in 100 mM Tris-HCl pH 7.0. (B) Temperature stability of SgAA. Residual activity after 24 h incubation at temperatures from 4 to 80 °C was measured with 15 mM acetyl-methionine in 100 mM Tris-HCl pH 7.0 at 30 °C. (C) pH optimum of hydrolysis with SgAA. The activity was measured with 15 mM acetyl-methionine at 30 °C in various 100 mM buffers. Na-acetate (●) was used for pH 4.0–5.0, Na-MES (□) was used for pH 6.0–7.0, Tris-HCl (▲) was used for pH 6.0–9.0, and Na-borate (▼) was used for pH 9.0–11.0. (D) pH optimum of SgELA in hydrolysis of 15 mM  $N_\epsilon$ -acetyl-lysine. All measurements were conducted in triplicate ( $n = 3$ ). The error bars indicate the standard deviations.

### Biocatalytic synthesis of lauroyl amino acids

Some aminoacylases are capable of amino acid acylation, which makes these enzymes interesting for the synthesis of amino acid-based surfactants. Hence, we investigated the biocatalytic potential of the aminoacylases from *S. griseus*. In a screen for acylation of all proteinogenic amino acids, the reactions were performed in aqueous buffers and a substrate excess of 100 mM L-amino acid and 100 mM lauric acid in 100 mM Tris-HCl pH 7.0. After 24 h at 40 °C, and the reaction mixtures were analyzed to detect acylation products. From the reaction conducted with SgELA, no acylation products could be detected under these conditions, whereas the homolog SmELA could

produce  $N_\epsilon$ -lauroyl-lysine to high yields [25]. SgELA presumably suffered from its low stability. However, the reactions performed with SgAA yielded an acylation product from lauric acid and methionine. A final concentration of 4 mM lauroyl-methionine was observed (MS spectrum of the product shown in supplements, Fig. S2). The reaction scheme is shown in Fig. 5. The reactions with the remaining proteinogenic amino acids did not result in product formation. Lauroyl-methionine is a biosurfactant with antioxidative properties and might thus be an interesting additive in cosmetic formulations [50]. The acceptance of further fatty acids is likely and can be investigated in future experiments.

**Fig. 5.** Reaction scheme for production of lauroyl-methionine catalyzed by SgAA. The condensation reaction was performed with 100 mM L-methionine and 100 mM lauric acid in 100 mM Tris-HCl at pH 7.0 at 40 °C and 20  $\mu\text{g}\cdot\text{mL}^{-1}$  of the aminoacylase SgAA.



Despite several aminoacylases from *S. mobaraensis* being capable of acylation, no acylation has been reported with SmAA (83.7% sequence identity to SgAA) [16]. In contrast, with SamAA from *S. ambofaciens* (88.3% sequence identity to SgAA), high yields for acylation of various amino acids have been reported. The SamAA enzyme was responsible for the main acylating activity of *S. ambofaciens* [5]. We showed that the newly characterized aminoacylase SgAA is also capable of synthesis of acyl amino acids, specifically lauroyl-methionine. Interestingly, the homologous MsAA from *M. smegmatis* (56.5% sequence identity) also acylated methionine with highest conversions [8]. The findings are significant for exploring the synthetic potential of streptomycetal M20A aminoacylases, confirming that it is well worth further investigating this sequence space. Future work will focus on the optimization of the synthesis of acyl amino acids with SgAA.

## Conclusions

We present the identification and cloning of aminoacylases SgAA and SgELA from *S. griseus* DSM 40236<sup>T</sup> that are homologous to aminoacylases from *S. mobaraensis*, *S. ambofaciens*, and *M. smegmatis*. The enzymes were successfully expressed in *S. lividans* TK23 and a protocol for high cell density fermentation in bioreactors was established, which ensures dispersed growth of the bacterial mycelium. Recombinant production and purification of Strep-tag II-fused protein in *S. lividans* was performed, which has previously only been reported once for secreted antigens from *M. tuberculosis* [51]. The putative aminoacylase activity could be verified for SgAA and SgELA, and they show characteristics in hydrolysis similar to their homologs. The dimeric enzyme SgAA is a short-chain acyl aminoacylase with a broad substrate spectrum regarding the amino acid moiety. On the contrary, SgELA is specific of N<sub>ε</sub>-acetyl-lysine. Due to its low stability,

SgELA did not prove to be suitable for biocatalysis. Future work will deal with the optimization of the conditions for heterologous production. With SgAA, synthesis of lauroyl-methionine in aqueous buffer was shown, which renders this enzyme interesting for biocatalytic applications and future work to optimize acylation conditions.

## Acknowledgements

Open Access funding enabled and organized by Projekt DEAL.

## Conflict of interest

The authors declare no conflict of interest.

## Data accessibility

The datasets supporting the conclusions of this article are included within the article and its additional files.

## Author contributions

GH designed the study. GH conducted cloning and bioinformatic analysis. GH and JP performed the experiments and analyzed the data for protein expression, purification, and biochemical characterization. GH and JP wrote the manuscript. GH, JB, K-EJ, and PS edited the manuscript. PS and JB supervised the work of GH and JP. PS and JB did fund acquisition. All authors read and approved the final manuscript.

## References

- 1 Farias CBB, Almeida FC, Silva IA, Souza TC, Meira HM, Da Soares Silva RCF, Luna JM, Santos VA, Converti A, Banat IM *et al.* (2021) Production of green surfactants: market prospects. *Electron J Biotechnol* **51**, 28–39.

- 2 Battista N, Bari M and Bisogno T (2019) N-acyl amino acids: metabolism, molecular targets, and role in biological processes. *Biomolecules* **9**, 822.
- 3 Yan H-D, Ishihara K, Serikawa T and Sasa M (2003) Activation by N-acetyl-L-aspartate of acutely dissociated hippocampal neurons in rats via metabotropic glutamate receptors. *Epilepsia* **44**, 1153–1159.
- 4 Morland C and Nordengen K (2022) N-acetyl-aspartyl-glutamate in brain health and disease. *Int J Mol Sci* **23**, 1268.
- 5 Bourkaib MC, Delaunay S, Framboisier X, Hôtel L, Aigle B, Humeau C, Guiavarc'h Y and Chevalot I (2020) N-acylation of L-amino acids in aqueous media: evaluation of the catalytic performances of *Streptomyces ambofaciens* aminoacylases. *Enzyme Microb Technol* **137**, 109536.
- 6 Koreishi M, Kawasaki R, Imanaka H, Imamura K and Nakanishi K (2005) A novel  $\epsilon$ -lysine acylase from *Streptomyces mobaraensis* for synthesis of N $\epsilon$ -acyl-L-lysines. *J Am Oil Chem Soc* **82**, 631–637.
- 7 Dettori L, Ferrari F, Framboisier X, Paris C, Guiavarc'h Y, Hôtel L, Aymes A, Leblond P, Humeau C, Kapel R *et al.* (2018) An aminoacylase activity from *Streptomyces ambofaciens* catalyzes the acylation of lysine on  $\alpha$ -position and peptides on N-terminal position. *Eng Life Sci* **18**, 589–599.
- 8 Haeger G, Wirges J, Tanzmann N, Oyen S, Jolmes T, Jaeger KE, Schörken U, Bongaerts J and Siegert P (2023) Chaperone assisted recombinant expression of a mycobacterial aminoacylase in *vibrio natriegens* and *Escherichia coli* capable of N-lauroyl-L-amino acid synthesis. *Microb Cell Fact* **22**, 77.
- 9 Takakura Y and Asano Y (2019) Purification, characterization, and gene cloning of a novel aminoacylase from *Burkholderia* sp. strain LP5\_18B that efficiently catalyzes the synthesis of N-lauroyl-L-amino acids. *Biosci Biotechnol Biochem* **83**, 1964–1973.
- 10 Wardenga R, Lindner HA, Hollmann F, Thum O and Bornscheuer U (2010) Increasing the synthesis/hydrolysis ratio of aminoacylase 1 by site-directed mutagenesis. *Biochimie* **92**, 102–109.
- 11 Wada E, Handa M, Imamura K, Sakiyama T, Adachib S, Matsunob R and Nakanishia K (2002) Enzymatic synthesis of N-acyl-L-amino acids in a glycerol-water system using acylase I from pig kidney. *J Am Oil Chem Soc* **79**, 41–46.
- 12 Rawlings ND and Barrett AJ (2013) Chapter 77 – Introduction: metallopeptidases and their clans. In *Handbook of Proteolytic Enzymes* (Rawlings ND and Salvesen G, eds), pp. 325–370. Academic Press, doi: 10.1016/B978-0-12-382219-2.00077-6
- 13 Nocek B, Reidl C, Starus A, Heath T, Bienvenue D, Osipiuk J, Jedrzejczak R, Joachimiak A, Becker DP and Holz RC (2018) Structural evidence of a major conformational change triggered by substrate binding in DapE enzymes: impact on the catalytic mechanism. *Biochemistry* **57**, 574–584.
- 14 Jozic D, Bourenkow G, Bartunik H, Scholze H, Dive V, Henrich B, Huber R, Bode W and Maskos K (2002) Crystal structure of the dinuclear zinc aminopeptidase PepV from *Lactobacillus delbrueckii* unravels its preference for dipeptides. *Structure* **10**, 1097–1106.
- 15 Girish TS and Gopal B (2010) Crystal structure of *Staphylococcus aureus* metallopeptidase (Sapep) reveals large domain motions between the manganese-bound and apo-states. *J Biol Chem* **285**, 29406–29415.
- 16 Koreishi M, Nakatani Y, Ooi M, Imanaka H, Imamura K and Nakanishi K (2009) Purification, characterization, molecular cloning, and expression of a new aminoacylase from *Streptomyces mobaraensis* that can hydrolyze N-(middle/long)-chain-fatty-acyl-L-amino acids as well as N-short-chain-acyl-L-amino acids. *Biosci Biotechnol Biochem* **73**, 1940–1947.
- 17 Bourkaib MC, Delaunay S, Framboisier X, Humeau C, Guilbot J, Bize C, Illous E, Chevalot I and Guiavarc'h Y (2020) Enzymatic synthesis of N-10-undecenoyl-phenylalanine catalysed by aminoacylases from *Streptomyces ambofaciens*. *Process Biochem* **99**, 307–315.
- 18 Liu Z, Zhen Z, Zuo Z, Wu Y, Liu A, Yi Q and Li W (2006) Probing the catalytic center of porcine aminoacylase 1 by site-directed mutagenesis, homology modeling and substrate docking. *J Biochem* **139**, 421–430.
- 19 Lindner HA, Lunin VV, Alary A, Hecker R, Cygler M and Ménard R (2003) Essential roles of zinc ligation and enzyme dimerization for catalysis in the aminoacylase-1/M20 family. *J Biol Chem* **278**, 44496–44504.
- 20 Nocek BP, Gillner DM, Fan Y, Holz RC and Joachimiak A (2010) Structural basis for catalysis by the mono- and dimetalated forms of the dapE-encoded N-succinyl-L,L-diaminopimelic acid desuccinylase. *J Mol Biol* **397**, 617–626.
- 21 Nocek B, Starus A, Makowska-Grzyska M, Gutierrez B, Sanchez S, Jedrzejczak R, Mack JC, Olsen KW, Joachimiak A and Holz RC (2014) The dimerization domain in DapE enzymes is required for catalysis. *PLoS One* **9**, e93593.
- 22 Gillner DM, Bienvenue DL, Nocek BP, Joachimiak A, Zachary V, Bennett B and Holz RC (2009) The dapE-encoded N-succinyl-L,L-diaminopimelic acid desuccinylase from *Haemophilus influenzae* contains two active-site histidine residues. *J Biol Inorg Chem* **14**, 1–10.
- 23 Davis R, Bienvenue D, Swierczek SI, Gilner DM, Rajagopal L, Bennett B and Holz RC (2006) Kinetic

- and spectroscopic characterization of the E134A- and E134D-altered dapE-encoded N-succinyl-L,L-diaminopimelic acid desuccinylase from *Haemophilus influenzae*. *J Biol Inorg Chem* **11**, 206–216.
- 24 Takakura Y, Nakanishi K, Suzuki S, Nio N, Koreishi M, Imamura T and Imanaka H N<sub>ε</sub>-acyl-L-lysine specific aminoacylase (N<sub>ε</sub>-アシル-L-リジン特異的アミノアシラーゼ, in Japanese) – JP2012039878A (2009): Japanese Patent Application.
- 25 Koreishi M, Kawasaki R, Imanaka H, Imamura K, Takakura Y and Nakanishi K (2009) Efficient N<sub>ε</sub>-lauroyl-L-lysine production by recombinant ε-lysine acylase from *Streptomyces mobaraensis*. *J Biotechnol* **141**, 160–165.
- 26 Bourkaib MC (2020) Doctoral thesis: Procédé enzymatique intensifié et durable de N-acylation d'acides aminés en milieux éocompatibles: caractérisation, cinétiques et immobilisation de nouveaux biocatalyseurs (in French), Nancy, France.
- 27 Padayatty JD and van Kley H (1967) Specificity of ε-peptidase of *Achromobacter pestifer* EA. *Arch Biochem Biophys* **120**, 296–302.
- 28 Chibata I, Tosa T and Ishikawa T (1964) On the substrate specificity of the bacterial ε-acylase. *Arch Biochem Biophys* **104**, 231–237.
- 29 Leclerc J and Benoiton L (1968) Further studies on ε-lysine acylase. The ω-N-acyl-diamino acid hydrolase activity of avian kidney. *Can J Biochem* **46**, 471–475.
- 30 Uhoraningoga A, Kinsella GK, Henehan GT and Ryan BJ (2021) β-Glucosidase from *Streptomyces griseus*: Ester hydrolysis and alkyl glucoside synthesis in the presence of Deep Eutectic Solvents. *Curr Res Green Sust Chem* **4**, 100129.
- 31 Usuki H, Uesugi Y, Iwabuchi M and Hatanaka T (2009) Putative “acylaminoacyl” peptidases from *Streptomyces griseus* and *S. coelicolor* display “aminopeptidase” activities with distinct substrate specificities and sensitivities to reducing reagent. *Biochim Biophys Acta* **1794**, 468–475.
- 32 Wu L-L, Li L-P, Xiang Y, Guan Z and He Y-H (2017) Enzyme-promoted direct asymmetric Michael reaction by using protease from *Streptomyces griseus*. *Catal Lett* **147**, 2209–2214.
- 33 Zhang Y, Liang Q, Zhang C, Zhang J, Du G and Kang Z (2020) Improving production of *Streptomyces griseus* trypsin for enzymatic processing of insulin precursor. *Microb Cell Fact* **19**, 88.
- 34 Trop M and Birk Y (1970) The specificity of proteinases from *Streptomyces griseus* (pronase). *Biochem J* **116**, 19–25.
- 35 Greenblatt HM, Almog O, Maras B, Spungin-Bialik A, Barra D, Blumberg S and Shoham G (1997) *Streptomyces griseus* aminopeptidase: X-ray crystallographic structure at 1.75 Å resolution. *J Mol Biol* **265**, 620–636.
- 36 Ouchi T (1962) Studies on proteolytic enzymes of *Streptomyces griseus*. *Agric Biol Chem* **26**, 728–733.
- 37 Kieser T, Bibb MJ, Buttner MJ, Chater KF and Hopwood DA (2000) *Practical Streptomyces Genetics*. Innes, Norwich.
- 38 Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* **215**, 403–410.
- 39 Needleman SB and Wunsch CD (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol* **48**, 443–453.
- 40 Notredame C, Higgins DG and Heringa J (2000) T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J Mol Biol* **302**, 205–217.
- 41 Robert X and Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**, W320–W324.
- 42 Rawlings ND and Morton FR (2008) The MEROPS batch BLAST: a tool to detect peptidases and their non-peptidase homologues in a genome. *Biochimie* **90**, 243–259.
- 43 Quirós LM, Aguirrezabalaga I, Olano C, Méndez C and Salas JA (1998) Two glycosyltransferases and a glycosidase are involved in oleandomycin modification during its biosynthesis by *Streptomyces antibioticus*. *Mol Microbiol* **28**, 1177–1185.
- 44 Bibb MJ, Janssen GR and Ward JM (1958) Cloning and analysis of the promoter region of the erythromycin resistance gene (ermE) of *Streptomyces erythraeus*. *Gene* **38**, 215–226.
- 45 Thompson CJ, Ward JM and Hopwood DA (1982) Cloning of antibiotic resistance and nutritional genes in *streptomycetes*. *J Bacteriol* **668**–677.
- 46 Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248–254.
- 47 Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
- 48 Haeger G, Bongaerts J and Siegert P (2022) A convenient ninhydrin assay in 96-well format for amino acid-releasing enzymes using an air-stable reagent. *Anal Biochem* **654**, 114819.
- 49 Xiang DF, Patskovsky Y, Xu C, Fedorov AA, Fedorov EV, Sisco AA, Sauder JM, Burley SK, Almo SC and Raushel FM (2010) Functional identification and structure determination of two novel prolidases from cog1228 in the amidohydrolase superfamily. *Biochemistry* **49**, 6791–6803.

- 50 Ohta A, Hossain F, Asakawa H and Asakawa T (2020) Study of the antioxidative properties of several amino acid-type surfactants and their synergistic effect in mixed micelle. *J Surfactant Deterg* **23**, 99–108.
- 51 Ayala JC, Pimienta E, Rodríguez C, Anné J, Vallín C, Milanés MT, King-Batsios E, Huygen K and van Mellaert L (2013) Use of Strep-tag II for rapid detection and purification of *Mycobacterium tuberculosis* recombinant antigens secreted by *Streptomyces lividans*. *J Microbiol Methods* **94**, 192–198.

### Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Fig. S1.** Native PAGE of SgAA and reference proteins.

**Fig. S2.** Mass spectrum of N-lauroyl-L-methionine produced by SgAA.

**Table S1.** Purification of SgAA from recombinant *S. lividans* culture by Strep-tag affinity chromatography.

## 3. Discussion

Biobased surfactants are in increasing demand. Acyl-amino acids not only have desirable chemical properties, being mild and skin-friendly, low inflammatory, and showing good foaming ability. They are also environmentally benign and are readily degraded, thus being non-persistent, which is important for regulatory aspects. Detergents have the largest market for chemical products, especially in the consumer market. Hence, it is important to find sustainable synthetic routes. Biocatalysis with aminoacylases is promising for this endeavor yet represents an underexplored field. The sustainability of acyl-amino acid production is not only defined by biocatalytic process, but also by the efficient production of the active biocatalyst. Therefore, the recombinant protein expression was a major part of this thesis. The work of this thesis comprises the identification, cloning, recombinant expression, purification, characterization, and evaluation of novel L-aminoacylases for the synthesis of acyl-amino acids.

### 3.1. Overview of results

**Chapter 1:** Chapter 1 presents the development of the ninhydrin-based assay for hydrolytic aminoacylase activity that was used throughout this thesis. Ninhydrin reacts with primary amino groups under the formation of Ruhemann's purple. Conventional ninhydrin reagents are susceptible to oxidation, which makes frequent reagent and standard preparation necessary. Furthermore, protocols described in literature for ninhydrin-based aminoacylase activity measurements are laborious and inconvenient for high throughput. In our study, a commercial, air-stable ninhydrin reagent, modified by mixture with DMSO, was found to be suitable for batch analysis. All proteinogenic amino acids could be measured. With conventional ninhydrin reagents, lysine and  $N_{\alpha/\epsilon}$ -acetyl-lysine lead to the same color yield, despite the two free amino groups in lysine. An acidic reagent variant was developed for the specific quantification of lysine from hydrolysis of  $N_{\alpha/\epsilon}$ -acetyl-lysine by mixture of the reagent with glacial acetic acid. Under these conditions, reaction with lysine yields an orange color, while  $N_{\epsilon}$ -acetyl-lysine develops a weak purple color and reaction with  $N_{\alpha}$ -acetyl-lysine remains colorless. Furthermore, the experimental layout was adjusted to a 96-well format, so that multichannel

pipettes and microtiter plates could be used. This allowed for a higher throughput compared to other ninhydrin-based aminoacylase assays presented in literature.

**Chapter 2:** The identification, cloning, expression, and characterization of the aminoacylase PmAcy from *Paraburkholderia monticola* DSM 100849 is described in chapter 2. The recombinant expression of soluble, active enzyme was challenging due to the formation of inclusion bodies. Soluble PmAcy could be obtained through co-expression of the chaperonin GroEL/S and lactose-autoinduction of the T7 expression system, combined with a low expression temperature of 20 °C. The purified enzyme was biochemically characterized. Extraordinary activity and stability at high temperatures and alkaline pH values was observed. Regarding substrate specificity, PmAcy can be classified as a long-chain acyl aminoacylase with a bias for hydrophobic amino acid residues. In biocatalysis, high conversion rates for the synthesis of acyl-amino acids could be achieved. A variety of amino acids could be acylated with conversions of 20 – 60 %, including positively charged substrate arginine and lysine, the imidazole-containing histidine, and hydrophobic amino acids like leucine or phenylalanine.

**Chapter 3:** The recombinant expression and initial characterization of the aminoacylase MsAA from *M. smegmatis* MKD 8 is presented in chapter 3. An expression study was conducted with various *E. coli* strains and *V. natriegens* Vmax™ to enhance soluble expression of the enzyme. The approaches were switched from IPTG induction to lactose autoinduction, lowering the expression temperature, and the co-expression of molecular chaperones, which led to significant improvements in soluble expression. The study presents the first report about chaperone co-expression in *V. natriegens*, and the first example of an aminoacylase being heterologously expressed by the species. The purified enzyme was characterized and biochemical parameters, such as temperature and pH optima, stabilities, and substrate specificity, were determined. Finally, a screen for synthetic acylation activity towards all proteinogenic amino acids showed that several hydrophobic lauroyl-amino acids can be produced.

**Chapter 4:** The study presented in chapter 4 builds upon chapter 3. The biocatalytic reaction with MsAA was optimized, focusing on the synthesis of lauroyl-methionine. A product concentration of 100 mM was reached with a conversion of 67 %, which are 13.5-fold and 9.0-fold respective increases from the initial reaction conditions. To get insight into substrate binding and to comprehend the observed bias towards methionine, docking simulations on the AlphaFold2-generated protein structure was performed. From the obtained results and comparison to homologs from literature, a reaction mechanism was proposed.

**Chapter 5:** In chapter 5, cloning and expression of the aminoacylases SgAA and SgELA from *S. griseus* DSM 40236<sup>T</sup> is presented. The enzymes were initially expressed in various *E. coli* strains and through chaperone co-expression, which, however, only resulted in overexpression as inclusion bodies. Hence, a fermentation protocol was developed that led to successful expression of soluble enzymes with recombinant *S. lividans* TK 23. The aminoacylases were initially characterized upon purification and tested in acylation of proteinogenic amino acids. The aminoacylase SgELA showed hydrolytic activity with N<sub>ε</sub>-acetyl-lysine, but not N<sub>α</sub>-acetyl-lysine, thus verifying the putative function as an ε-lysine acylase. The α-aminoacylase SgAA showed several similarities to MsAA, was found to be capable to synthesize lauroyl-methionine.

### 3.2. Novel aminoacylases were identified

In the work of this thesis, it was anticipated to find and produce novel aminoacylases for the synthesis of acyl-amino acids, preferentially using longer acyl chains like in lauric acid. While aminoacylases from *Aspergillus* sp. are used for racemic cleavage of acetyl-amino acids [150], the use of aminoacylases in the synthesis of acyl-amino acids is still an emerging field. In general, few aminoacylases have been characterized and their sequence published. However, recent research presented several aminoacylases with promising potential in biocatalytic synthesis [30, 31, 33]. Hence, enzymes from literature were evaluated and used as baits for BLASTp search for homologous enzymes.

The aminoacylases described in literature were evaluated with a focus on synthesis of acyl-amino acids. Only L-aminoacylases were considered, because proteinogenic amino acids, eventually obtained from renewable biological materials, are intended to be used as substrates in biocatalysis. Ideally, a broad range of amino acid substrates or peptides should be accepted as substrate to maximize the scope of possible compounds. On the other hand, no dipeptidase activity should be exerted by the enzymes to keep the possibility of peptide acylation. Due to evolutionary linkage between aminoacylases and peptidases, promiscuous activity might occur [151]. Regarding the acyl moiety, longer acyl chains, like lauric or palmitic acid, are preferred substrates in biocatalysis. High stability of the enzymes is also important for industrial application. After identifying proprietary sequences, the aminoacylases candidates were intended to be heterologously produced, preferably in *E. coli*. However, the recombinant expression of aminoacylases is often described as difficult or unsuccessful in *E. coli*, with

inclusion body formation being a common drawback [33, 66, 112]. Still, not least due to possible difficulties in expression, bacterial aminoacylases were preferred over eukaryotic enzymes.

By the start of the thesis, the aminoacylase isolated from *Burkholderia* sp. LP5\_18B showed the most favorable characteristics, opening a promising sequence space. High conversions for the synthesis of lauroyl-amino acids, and good stability at a broad pH range and high temperatures render this enzyme potent for industrial application. However, attempts to produce the enzyme in *E. coli* were not successful [33]. A homologous enzyme that shows 85.5 % sequence identity was found in *Paraburkholderia monticola* DSM 100849. The putative aminoacylase was designated PmAcy and was chosen a candidate enzyme for this thesis (GenBank: KXU85199.1). By MEROPS BLAST [39], the identified enzyme can be classified as a member of the M38 family. Sequence identities to further M38 peptidases were 33.9 % for Sgx9260b, 31.3 % for Sgx9260c, 26.2 % for Cc2672, 17.6 % for isoaspartyl dipeptidase from *E. coli*, and 18.3 % for the non-peptidase homolog SmELA.

Several aminoacylases have been characterized that belong to the M20A metallopeptidase subfamily. Among these are the porcine pAcy1 [51] and human hAcy1 [62], the streptomycetal SmAA from *S. mobaraensis* [69] and SamAA from *S. ambofaciens* [31], and acetylornithine deacetylase (EC 3.5.1.16) from *E. coli* [151, 152]. For other related M20A metallopeptidases, in particular N-succinyl-L,L-diaminopimelate desuccinylase (EC 3.5.1.18) from *H. influenzae* (HiDapE) [46] and carboxypeptidase G2 (EC 3.4.17.11) from *Pseudomonas* sp. RS-16 (CPG2) [153], crystal protein structures have been published, and a mechanism for HiDapE has been proposed [47]. Due to conserved residues essential for catalysis among these enzymes, a common mode of action is likely. Good acylation activity has been described for SamAA towards a broad range of amino acids [31]. The conversion yields, however, were not as high as for the *Burkholderia* aminoacylase [33], and the enzyme was not as stable, but studies on SamAA were conducted with partially purified crude cell extract from the natural producer.

Furthermore, an  $\epsilon$ -lysine acylase has been cloned and characterized from *S. mobaraensis* [30], and  $\epsilon$ -lysine acylase activity has been observed from *S. ambofaciens* extract [113], which been attributed to the enzyme SamELA [31]. Acylation with an  $\epsilon$ -lysine acylase from *S. coelicolor* (ScELA) has been described in a patent [77]. The production of N $\epsilon$ -lauroyl-lysine could be achieved with up to 100 % conversion using SmELA and ScELA [30, 74, 77]. Heterologous expression of SamELA was not successful in *E. coli* Origami B(DE3) [112], and SmELA was expressed in *S. lividans* TK24 in literature, but the patent also described expression of both

### 3. Discussion - 3.3. Conserved residues of metallopeptidases in the novel aminoacylases

SmELA and ScELA in *E. coli* JM109 and ScELA with *C. glutamicum* [77]. The streptomycetal aminoacylases were interesting because of high level synthesis with SamAA against amino acids, and extremely high yields with SmELA and ScELA. We identified SgAA, which showed sequence similarity to SamAA, SmAA, pAcy1, HiDapE of 88.3 %, 83.7 %, 26.5 %, and 23.7 %, respectively, and SgELA, which showed sequence identity to SamELA, SmELA, and the *Burkholderia* aminoacylase of 79.6 %, 75.6 %, and 21.0 %, respectively.

Aminoacylases have also been described from *M. smegmatis*, but no gene or protein was sequenced. One of these two aminoacylases preferred long-chain acyl residues, while the other was a short-chain acyl aminoacylase [9, 105]. As Actinomycetes with high G+C content, *Mycobacterium* and *Streptomyces* are distantly related, and similarities can be found on molecular and cellular level [154]. Mycobacteria produce mycolic acid and have a waxy layer in their cell wall [155]. DapE enzymes are involved in synthesis of cell wall components [156]. Even though mycolic acids are not directly linked to diamino pimelic acid, the occurrence of these long acyl-chain molecules hints at homologous mycobacterial aminoacylases possibly accepting long-chain fatty acids, which would be of benefit for biocatalysis. Hence, with the same baits, namely the protein sequences of SmAA, SamAA, SmELA, SamELA, BLASTp-search for homologs was performed on protein sequences of *M. smegmatis* MKD 8. We found a putative aminoacylase, which we designated MsAA, with a sequence identity to SgAA, SamAA, SmAA, pAcy1, hAcy1, ArgE, and HiDapE of 56.5 %, 56.3 %, 54.9 %, 25.5 %, 25.2 %, 22.1 %, and 21.4 %, respectively. Furthermore, a sequence with homology to  $\epsilon$ -lysine acylases, designated MsELA, was found with sequence identity to SamELA, SgELA, SmELA, and the *Burkholderia* aminoacylase of 50.8 %, 50.0 %, 49.7 %, and 19.5 %, respectively.

### 3.3. Conserved residues of metallopeptidases in the novel aminoacylases

The aminoacylase MsAA from *M. smegmatis* MKD 8 has a length of 450 amino acids and a molecular weight of 48.6 kDa, with a theoretical isoelectric point (pI) of 4.82. By MEROPS BLAST, the enzyme can be classified as a M20A metallopeptidase and shows the conserved sites of this protein family. These represent metal-binding residues (H91, D123, E158, E185, H425) and catalytic residues (D93, E157, H226). From sequence comparison with dimeric members of M20A family (HiDapE, pAcy1, hAcy1) and monomeric M20A peptidases (PepV), it can already be assumed that MsAA is likely a dimeric enzyme. Among dimeric M20A peptidases, the putatively catalytically active H226<sub>MsAA</sub> is conserved, but not in monomeric

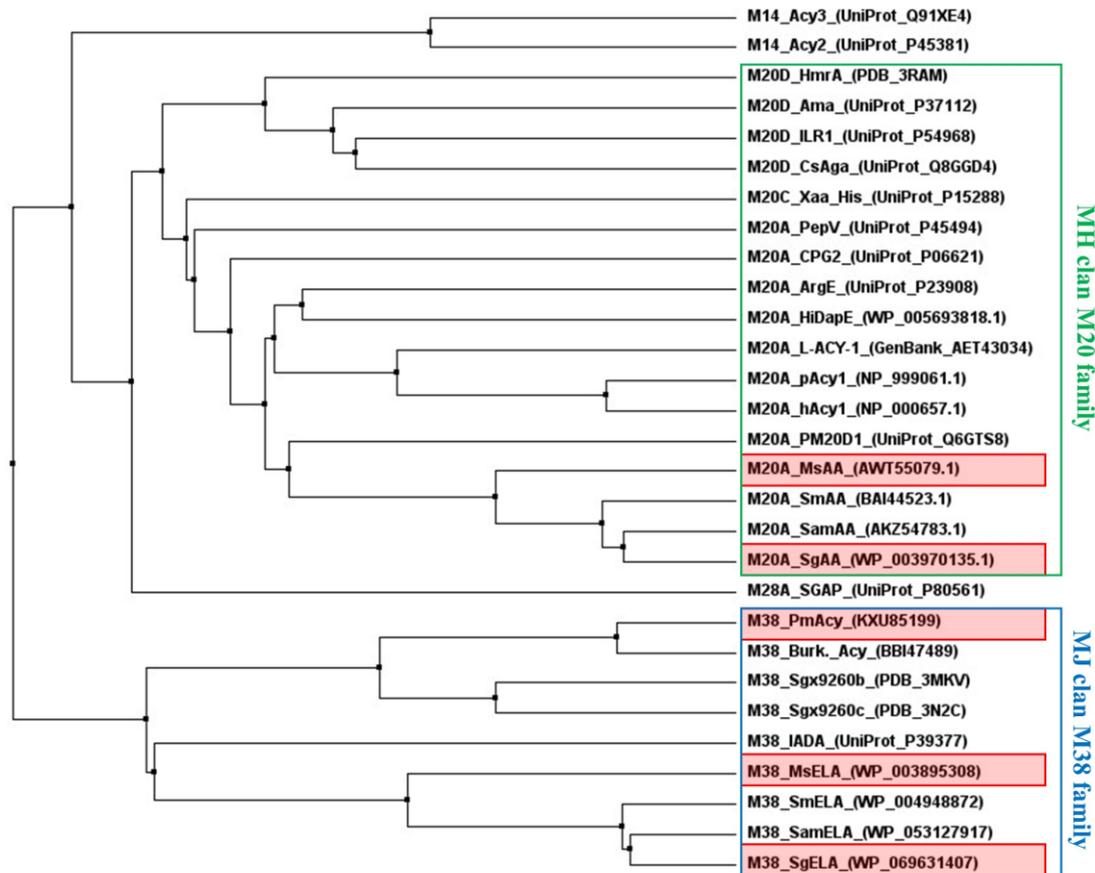
enzymes. For the dimeric HiDapE, it has been shown that the conserved H194<sub>HiDapE</sub> is located at the tip of the dimerization domain and protrudes into the active site of the opposing dimer. The histidine residue contributes to the formation of an oxyanion hole and stabilizes the reaction intermediate [47]. In turn, H269<sub>PepV</sub>, which is oxyanion hole-forming in PepV [54], is not conserved in the dimeric M20A peptidases. The homologous aminoacylase SgAA from *S. griseus* DSM 40236<sup>T</sup> also shows the metal-binding (H90, D122, E157, E184, H418) and catalytic (D92, E156, H221) residues conserved for M20A peptidases. Hence, the same conclusions as for MsAA can be drawn for this enzyme. The enzyme has a length of 443 amino acids, a molecular weight of 48.0 kDa, and a theoretical pI of 4.95.

The aminoacylase PmAcy from *P. monticola* DSM 100849 does not share these features of the M20A peptidases but can be assigned to M38 metallopeptidases. The enzyme consists of 440 amino acids and its monomer has a molecular weight of 47.4 kDa and a theoretical pI of 6.52. The conserved sites for metal-binding typical for M38 peptidases can be found in PmAcy (H85, H87, H253, H273, D340). Furthermore, K212 is conserved, which has been shown to be carboxylated in homologous enzymes as the bridging residue between the metal ions in the binuclear active site [60]. For homologous enzymes, H162 has been attributed an oxyanion-hole forming function, and Y255 is involved in binding the  $\alpha$ -carboxylic group of the substrate [60]. The homologous aminoacylase from *Burkholderia* sp. LP5\_18B is an octameric enzyme [33], as shown for other homologous M38 peptidases as well [58, 60]. However, no estimation for the oligomerization state of PmAcy can be done based on its amino acid sequence.

The putative  $\epsilon$ -lysine acylases SgELA from *S. griseus* DSM 40236<sup>T</sup> and MsELA from *M. smegmatis* MKD 8 can also be assigned to the M38 family of metallopeptidases. The aminoacylase SgELA has a length of 548 amino acids, a molecular weight of 56.9 kDa, and a theoretical pI of 5.56. Compared to PmAcy or the *Burkholderia* sp. LP5\_18B aminoacylase, the metal-binding residues are conserved in SgELA (H76, H78, H330, H365, D431). On the other hand, the (carboxylated) lysine residue described to bridge the cocatalytic metal ions is not conserved. Interestingly, the oxyanion-hole forming histidine is conserved in SgELA as H177, but not in SamELA or SmELA. In neither of the putative  $\epsilon$ -lysine acylases, the tyrosine residue that binds the substrates  $\alpha$ -carboxylic group is conserved. This might explain the missing or low  $\alpha$ -aminoacylase activity of these enzymes [74]. In MsELA, the metal-binding residues are conserved as well (H62, H64, H314, H349, D415). The oxyanion hole-forming histidine, however, is not conserved in MsELA. The sequence alignments of the aminoacylases are presented in the result chapters on cloning and characterization of the respective enzymes.

### 3. Discussion - 3.3. Conserved residues of metallopeptidases in the novel aminoacylases

A phylogenetic tree with the aminoacylases and their homologs mentioned in this thesis is shown in figure 5.



**Figure 5:** Phylogenetic tree of the relevant metallopeptidases and non-peptidase homologs described in this thesis. The alignment was generated with the Clustal Omega algorithm [157] and displayed by average distance. Sequences of the aminoacylases investigated in this work are highlighted with red boxes. The enzymes are listed with their acronyms and peptidase family, and their respective NCBI accession number, or PDB- or UniProt identifiers. The following sequences were included: **Acy3**: murine aminoacylase-3 (M14; UniProt: Q91XE4); **Acy2**: human aminoacylase-2 or aspartoacylase (M14; UniProt: P45381); **HmrA**: peptidase from *Staphylococcus aureus* (M20D; PDB: 3RAM); **Ama**: Aminoacylase from *Geoacillus stearothermophilus* (M20D; Uniprot P37112); **ILR1**: IAA-amino acid hydrolase from *Arabidopsis thaliana* (M20D; UniProt\_P54968); **CsAga**: alpha-glutamine aminoacylase from *Corynebacterium striatum* Ax20 (M20D; UniProt\_Q8GGD4) **Xaa His**: Dipeptidase from *Escherichia coli* (M20C; UniProt\_P15288); **PepV**: beta-Ala-Xaa dipeptidase from *Lactobacillus delbrueckii* (M20A; UniProt\_P45494); **CPG2**: Carboxypeptidase G2 from *Pseudomonas* sp. RS-16 (M20A; UniProt\_P06621); **ArgE**: acetylmithine deacetylase from *E. coli* (M20A; UniProt: P23908); **HiDapE**: succinyl-diaminopimelate desuccinylase from *Haemophilus influenzae* (M20A; Accession Number: WP\_005693818); **L-ACY-1**: aminoacylase 1 from *Heliothis virescens* (M20A; GenBank: AET43034); **pAcy1**: porcine aminoacylase-1 (M20A; Accession Number: NP\_999061); **hAcy1**: human aminoacylase-1 (M20A; Accession Number: NP\_000657); **PM20D1**: human N-fatty-acyl-amino acid hydrolase (M20A; UniProt: Q6GTS8); **MsAA**: aminoacylase from *Mycobacterium smegmatis* MKD8 (M20A; GenBank: AWT55079); **SmAA**: aminoacylase from *Streptomyces mobaraensis* (M20A; GenBank: BAI44523); **SamAA**: aminoacylase from *Streptomyces ambofaciens* ATCC 23877 (M20A; GenBank: AKZ54783); **SgAA**: aminoacylase from *Streptomyces griseus* (M20A; Accession Number: WP\_003970135); **PmAcy**: aminoacylase from *Paraburkholderia monticola* (M38; GenBank: KXU85199); **Burk. Acy**: aminoacylase from *Burkholderia* sp. LP5\_18B (M38; GenBank: BBI47489); **Sgx9260b**: amidohydrolase (M38; PDB: 3MKV); **Sgx9260c**: prolidase (M38; PDB: 3N2C); **IADA**: isoaspartyl dipeptidase from *E. coli* (M38; UniProt P39377); **MsELA**: epsilon-lysine acylase from *M. smegmatis* (M38; Accession Number: WP\_003895308); **SmELA**: epsilon-lysine acylase from *S. mobaraensis* (M38; Accession Number: WP\_004948872); **SamELA**: epsilon-lysine acylase from *S. ambofaciens* (M38; Accession Number: WP\_053127917); **SgELA**: epsilon-lysine acylase from *S. ambofaciens* (M38; Accession Number: WP\_069631407).

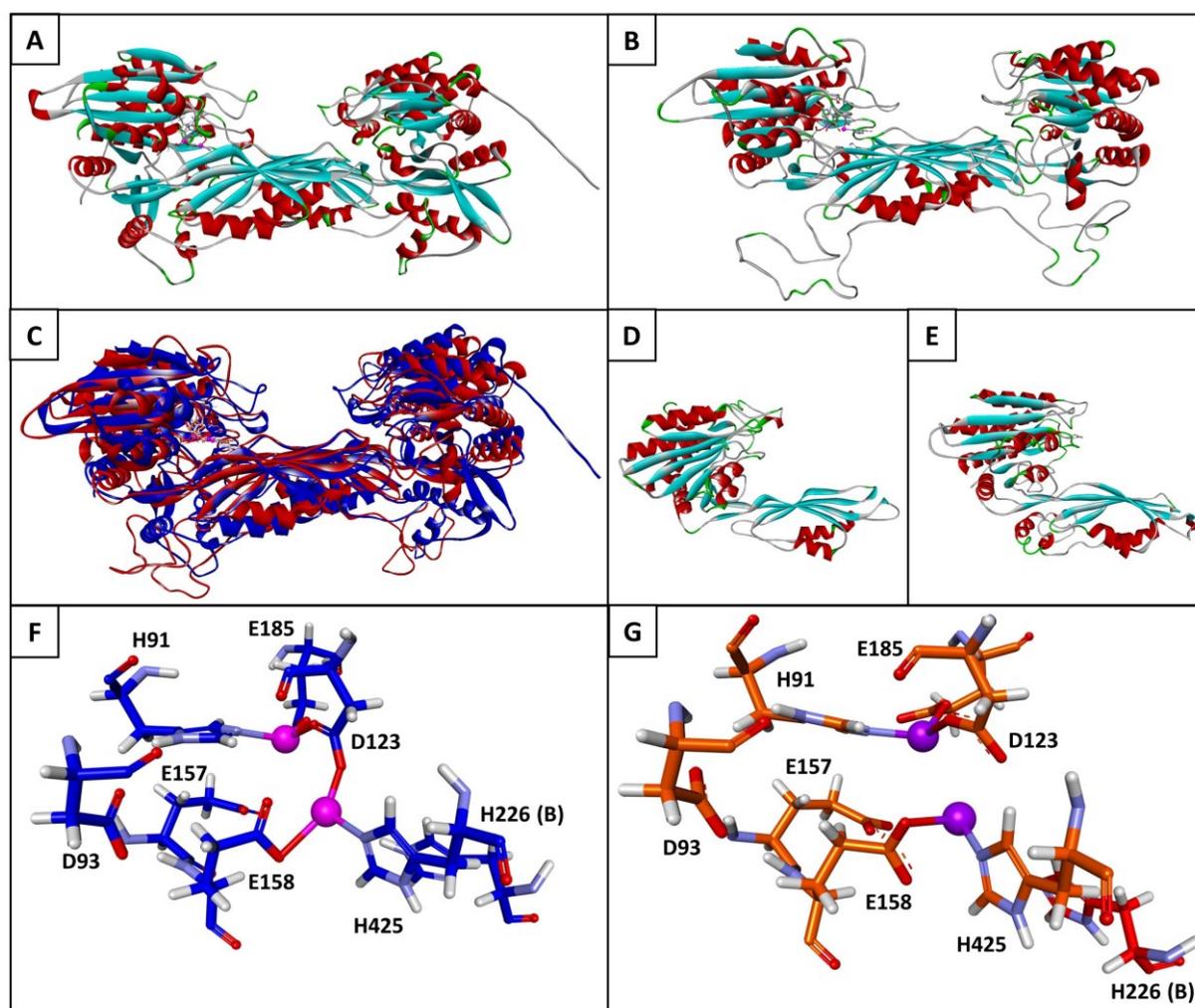
### 3.4. Predicted structures of the novel aminoacylases

In addition to sequence analysis, protein structure prediction can provide further insight into the properties of the novel aminoacylases. For instance, the enzymes can be studied in terms of similarities and differences in protein folding or domain architecture within their protein family. Furthermore, insight structure-function relationships can be gained. With high quality models, even studies on molecular docking or rational protein engineering can be conducted. However, only few structures of L-aminoacylases have been solved and published (see table 1, section 1.2.4). No structures of closely related aminoacylases (>40 % identity) are available for the novel aminoacylases in this thesis. Still, the structures of some homologous peptidases (<34 %) have been solved [47, 60]. In general, protein structure is more conserved than sequence [158]. Especially for M20 peptidases, a high level of fold- and domain conservation can be observed [62], despite low sequence homology, which should be beneficial for structure predictions.

Protein structure prediction can be template-based or *ab initio* free modeling. Hybrid approaches can be realized in that homology-based modeling uses energy-guided model refinement, or that free modeling uses fragment-based sampling approaches [159]. The template-based SWISS-MODEL was the first protein modeling service on the internet and has been constantly improved [160]. From the amino acid input sequence, evolutionary related structures are searched, so that suitable templates can be found. For the selected template, a model is generated and refined, and insertions or deletions in the alignment are predicted by loop modeling [161]. Due to the template-based approach, the quality of the model greatly depends on the alignments and will be poorer for low sequence homology [162]. At least ~30 % sequence identity between the target and template is needed for acceptable prediction, and below that threshold, the models become unreliable [163, 164]. The Phyre2 protocol uses threading, which essentially extracts structure information from the alignment of the template and target, and is more suitable for remote homology [165]. However, the service cannot predict multimer structures, and tends to delete regions predicted with poor confidence from the final structure. The I-TASSER protocol is a hybrid method using homology templates and *ab initio* prediction. First, homologous structure templates are identified using multiple threading, and the topology of the full-length model is constructed by reassembling the aligned fragment structures. Unaligned regions are created through *ab initio* folding approaches. The assembly of structures and following energy minimizations are iterative [166]. The novel AlphaFold software uses artificial intelligence and contains a neural network that is trained by PDB structures to predict distances between the residues of the model [159, 167]. It has been shown

### 3. Discussion - 3.4. Predicted structures of the novel aminoacylases

to outperform other prediction algorithms and is considered a break-through in protein structure prediction [159]. Its current version is AlphaFold2. A publicly accessible Jupyter Notebook has been made available through Google Colaboratory, called ColabFold [168]. In the following, the predicted structures from AlphaFold and SWISS-MODEL are shown and discussed briefly. For MsAA, structures predicted by Phyre2 and I-TASSER are shown as well. The SWISS-MODEL templates for MsAA, PmAcy, and SgELA were HiDapE (5VO3), Sgx9260c (3N2C), and an uncharacterized amidohydrolase from *Pyrococcus furiosus* (3ICJ), with sequence identities of 21.4 %, 31.3 %, and 25.6 %, respectively. To the predicted structures, Zn<sup>2+</sup> ions were added with the Metal Ion-Binding Site Prediction and Docking Server (MIB) [169].

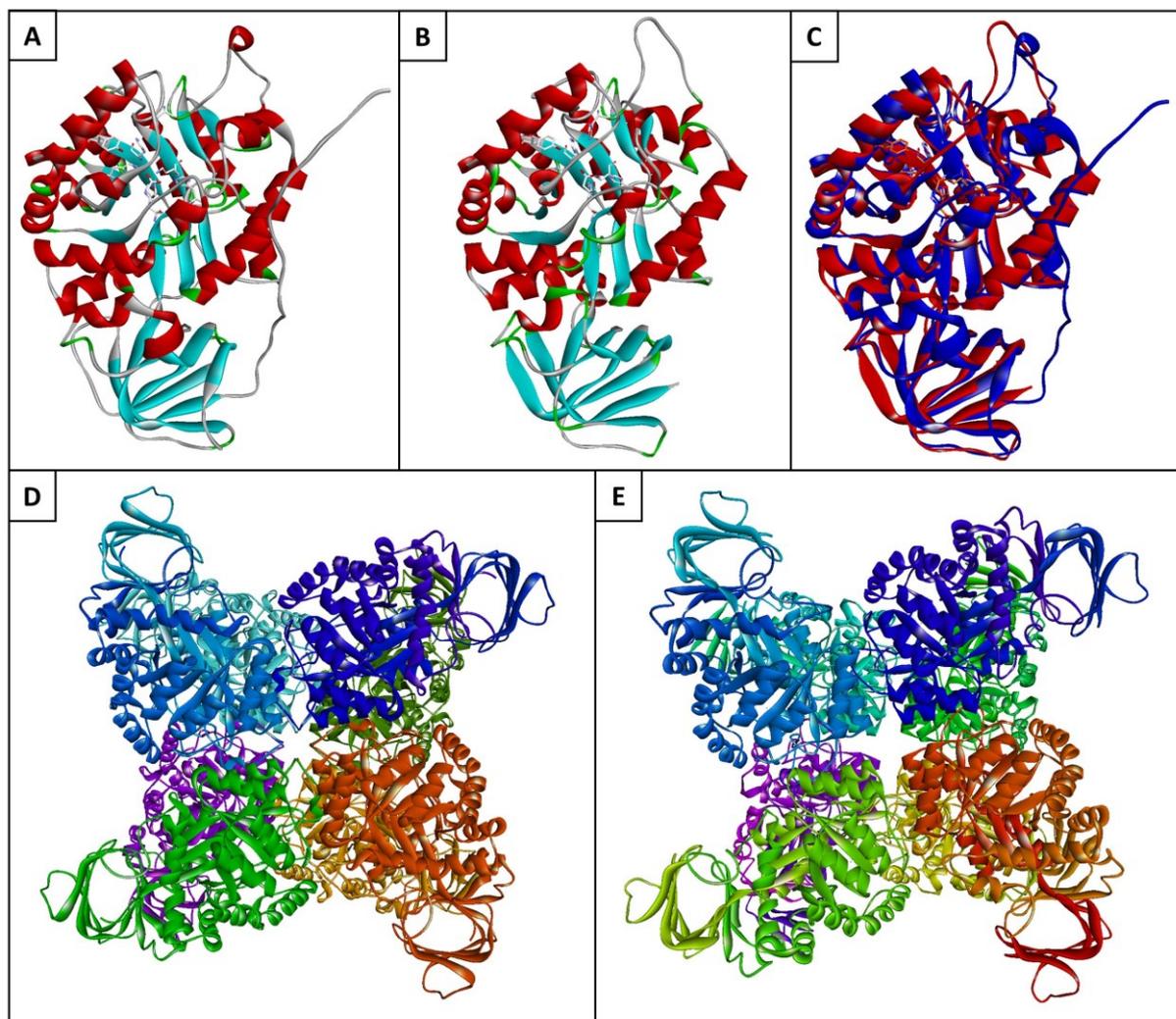


**Figure 6:** Predicted protein structures of MsAA. (A) AlphaFold2-generated structure. (B) SWISS-MODEL-generated structure using HiDapE (PDB 5VO3) as homology template. The SWISS-MODEL structure is truncated at the N-terminus and starts with S13. (C) Overlay of AlphaFold2 structure (blue) and SWISS-MODEL structure (red). (D) Phyre2-generated structure. (E) I-TASSER-generated structure. (F) Active site of the AlphaFold2-generated structure. The carbon atoms are shown in blue color. The zinc ions are shown as magenta balls. (G) Active site of the SWISS-MODEL-generated structure. The carbon atoms are shown in orange color. The zinc ions are shown as purple balls.

The AlphaFold2-predicted structure of MsAA (Fig. 6A) has already been shown in chapter 2.3, and was used to perform docking studies, presented in chapter 2.4. The predicted structure shows the characteristic catalytic and dimerization domains and the architecture of  $\alpha$ -helices and  $\beta$ -folds typical for the M20A metallopeptidase family [38, 62]. The active site comprises the metal-binding and catalytic residues as previously described (Fig. 6FG). Furthermore, the residue H226(B) protrudes into the active site. The SWISS-MODEL structure of MsAA was based on HiDapE (PDB 5VO3) as a homology model (Fig. 6B). For HiDapE, it has been shown that the hinge between the catalytic and dimerization domains closes and opens with substrate binding and release, respectively [47]. In PDB 5VO3, the structure was solved with an inhibitor. Because both the AlphaFold2- and SWISS-MODEL-generated structures are similar to PDB 5VO3 in that regard, the AlphaFold2 structure should represent closed, substrate-bound conformation. The predictions from Phyre2 (Fig. 6D) and I-TASSER (Fig. 6E) resulted in a monomeric structure (Fig. 6D). Several differences between the predicted structures can be observed. In the active site of the SWISS-MODEL-generated structure, the bridging aspartate D123 is not well-oriented, so that it does not bridge the added zinc ions (Fig. 6G). In the AlphaFold2 structure, the N-terminus has no fold and low confidence of prediction. The N-terminus also has a loose fold with a short  $\alpha$ -helix in the I-TASSER model. This region was not predicted by either SWISS-MODEL or Phyre2, but is truncated, so that the structures start with the residue S13 and A11, respectively. Despite this loosely folded N-terminus being too short for regular secretion signals, we subjected the sequence to the signal peptide prediction with the SignalP algorithm. With SignalP-5.0 [170], the first 13 amino acids had a probability of 0.3189 to be a Sec signal peptide. However, with the newer SignalP-6.0 algorithm [171], zero probability for the occurrence of a signal peptide was calculated. The homologous human PM20D1 enzyme has a secretion signal, with a SignalP-6.0-calculated probability of 0.9997, and was also described as an extracellular enzyme [17]. Its AlphaFold-generated structure, accessible under UniProt identifier Q6GTS8, is more similar to MsAA than the human hAcy1, its AlphaFold structure being accessible under UniProt identifier Q03154. In PM20D1 and MsAA, 3-4 additional  $\alpha$ -helices were predicted by AlphaFold2 at the outer hinge region, which are missing in HiDapE or predicted hAcy1 structures. A similar, but less organized region can be found in the I-TASSER model as well. In the SWISS-MODEL structure of MsAA, these residues are predicted as a large loop from H249 to D301. The Phyre2-generated structure has a gap between P252 and R293, and the remaining adjacent residues form a shorter loop. Despite observations that AlphaFold2 overpredicts  $\alpha$ -helices, and secondary structures in general [172], more confidence can be put into the AlphaFold2 structure, especially with the region in question

### 3. Discussion - 3.4. Predicted structures of the novel aminoacylases

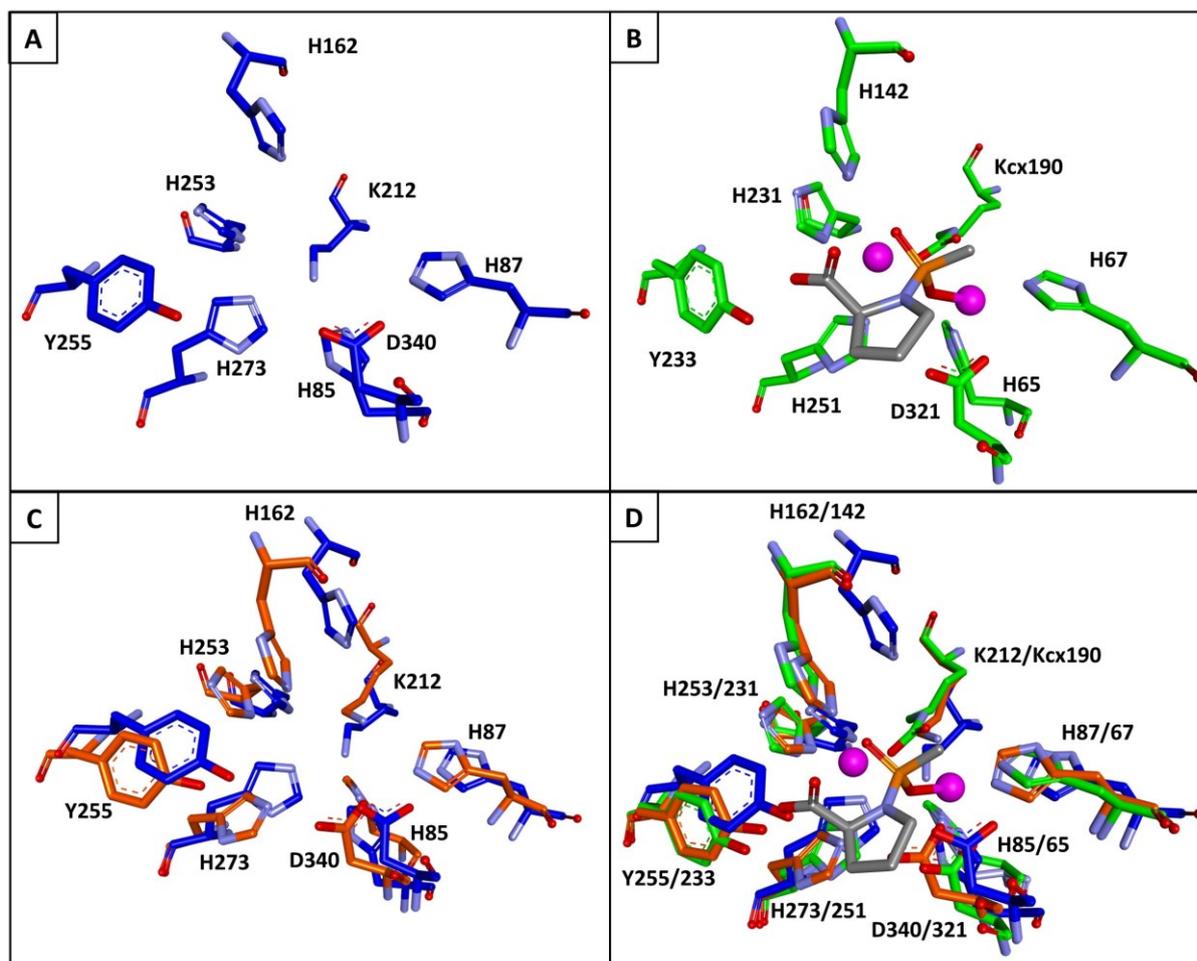
being deleted in the Phyre2 model. The function of these helices can only be speculated, possibly being beneficial for extracellular or vesicular folding for the secreted PM20D1. The elongated, unfolded N-terminus of MsAA could be a remnant or artefact from an evolutionarily former secretion signal, or might be responsible for membrane-association, due to its hydrophobic character. The AlphaFold2 and SWISS-MODEL-predicted models of SgAA (Fig. S1, supplementary materials) are very similar to the respective predicted structures of MsAA. The N-terminus of AlphaFold2-structure of SgAA has no fold, and the SWISS-MODEL-predicted model is truncated, starting with A14. The SignalP algorithms did not find a probability for signal peptides in SgAA either. Still, for future work, it would certainly be interesting to fuse both MsAA and SgAA to heterologous signal peptides and to attempt to secrete the enzymes, preferably in recombinant *Bacillus* or *Streptomyces* species.



**Figure 7:** Predicted protein structures of PmAcy. **(A)** AlphaFold-generated monomeric structure. **(B)** SWISS-MODEL-generated monomeric structure based on Sgx9260c (PDB 3N2C). The SWISS-MODEL structure is truncated at both termini and starts at K23 and ends with T437 (instead of A440). **(C)** Overlay of AlphaFold structure (blue) and SWISS-MODEL structure (red). **(D)** Homotetrameric structure of the AlphaFold structure. **(E)** Homooctameric structure of the SWISS-MODEL structure.

The predicted structures of PmAcy are shown in figure 7. For the SWISS-MODEL prediction, the prolidase Sgx9260c served as the homology template (PDB 3N2C). Overall, the AlphaFold2- (Fig. 7A) and SWISS-MODEL- (Fig. 7B) predicted structures show similar folds. The two domains that are characteristic for M38 peptidases, namely the N-terminal domain of eight  $\beta$ -strands and the catalytic C-terminal domain with  $(\beta/\alpha)_8$ -barrel fold, are present in the models (compare Fig. 4, chapter 1.2.3.). The N-terminus of the AlphaFold2 structure has a loose fold, somewhat similar to the AlphaFold2 structure of MsAA (Fig. 6A). Again, the SWISS-MODEL structure is truncated at the N-terminus, starting with K23, but the model is also missing three amino acids at the C-terminus. As the homology template, the SWISS-MODEL structure of PmAcy is homooctameric (Fig. 7E). Hence, we also generated an octameric structure in AlphaFold2 (Fig. 7D). The quaternary orientation of the multimer is similar, and the elongated N-terminus does not seem to disturb multimerization in the model. The orientation of the metal-binding and catalytic residues in the active site are similar among the AlphaFold2- and SWISS-MODEL structures, and Sgx9260c (Fig. 8). The AlphaFold2- and SWISS-MODEL structures were generated with non-carboxylated lysine. Hence, no metal ions were added to these models. The positioning of the oxyanion hole-forming histidine (H142<sub>Sgx9260c</sub> or H162<sub>PmAcy</sub>) and the tyrosine residue responsible for binding of the  $\alpha$ -carboxyl group of the amino acyl substrate (Y233<sub>Sgx9260c</sub> or Y255<sub>PmAcy</sub>) are conserved in the models, suggesting a similar role of these residues in PmAcy as in Sgx9260c [60].

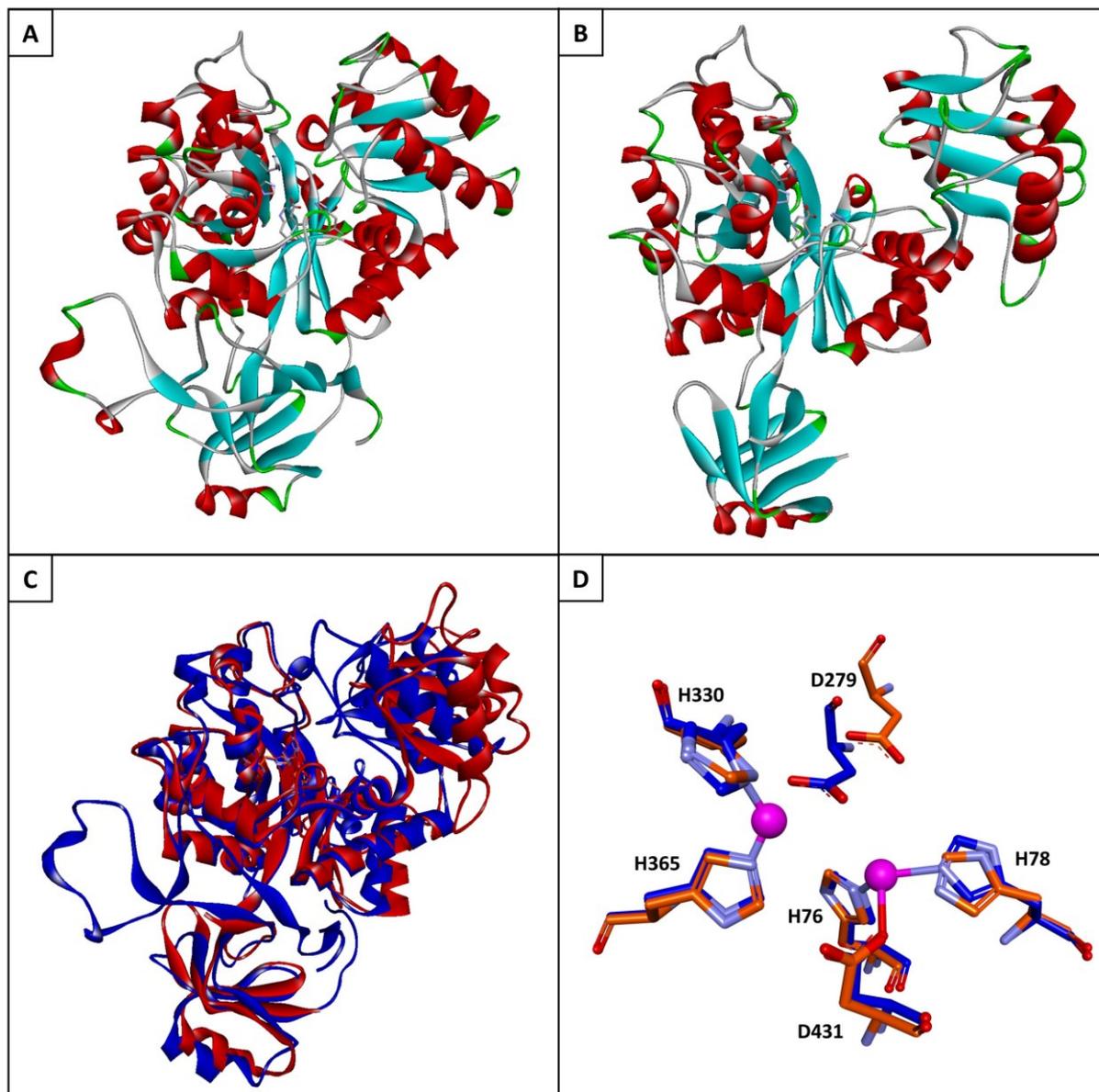
### 3. Discussion - 3.4. Predicted structures of the novel aminoacylases



**Figure 8:** Active sites of predicted PmAcy-structures and the homolog Sgx9260c (PDB 3N2C). **(A)** Active site of the AlphaFold2-generated PmAcy structure. The carbon atoms are shown in blue color. **(B)** Active site of Sgx9260c (PDB 3N2C). The carbon atoms are shown in green color. The zinc ions are shown as magenta balls. The inhibitor N-methylphosphonate-L-proline is shown with grey carbon atoms. **(C)** Overlay of active site of the AlphaFold2-generated (blue carbon atoms) and SWISS-MODEL (orange carbon atoms) active sites. **(D)** Overlay of active site of the AlphaFold2-generated (blue carbon atoms) and SWISS-MODEL (orange carbon atoms) and Sgx9260c (PDB 3N2C, green carbon atoms) active sites. The zinc ions of Sgx9260c are shown as magenta balls and the inhibitor N-methylphosphonate-L-proline is shown with grey carbon atoms.

The predicted protein structures SgELA are shown in figure 9. The SWISS-MODEL prediction was based on an uncharacterized metal-dependent hydrolase from *Pyrococcus furiosus* (PDB 3ICJ), and is truncated at both termini, starting with R14 and ending with T494, missing 54 residues. The homologous SmELA was described as monomeric [74], like the used template. The loop-domain containing two short  $\alpha$ -helices (V499-G519) close to the C-terminus is only present in the AlphaFold2 structure and predicted with poor confidence. Furthermore, when comparing the two predicted structures, the catalytic, active-site containing domain has an open conformation in the SWISS-MODEL structure, but a closed conformation in the AlphaFold2-structure. Since no structure-function information on other  $\epsilon$ -lysine acylases is available, no certain conclusions can be drawn. The predictions only hint at a conformational change due to substrate binding. A major difference of SgELA to other characterized members of the M38

metallopeptidase family is the absence of Kcx as bridging residue. In the AlphaFold2-generated structure, when adding zinc ions through the MIB service, the aspartate D279 was found as a possible bridging residue (Fig. 9D). Interestingly, the homology template (PDB 3ICJ) has a Kcx residue in the active site as the bridging residue. In the homologous SmELA, D505<sub>SmELA</sub> is described as a metal-binding residue, which is conserved in SgELA as D512<sub>SgELA</sub>. However, this residue is located in the described loop that contains two  $\alpha$ -helices, far from the active site. Therefore, the assigned metal-binding function is questionable. The oxyanion-hole forming histidine H142<sub>Sgx9260c</sub> is conserved in the alignment SgELA as H177<sub>SgELA</sub>, but not in SamELA or SmELA (chapter 2.5). However, this residue is not located near the active site of the predicted SgELA structures. The predicted structures of MsELA are similar and shown in Fig. S2 (supplementary material).



**Figure 9:** Predicted protein structures of SgELA. **(A)** AlphaFold2-generated structure. **(B)** SWISS-MODEL-generated structure based on PDB 3ICJ. The SWISS-MODEL structure is truncated at both termini and starts at R14 and ends with T494 (instead of E548). **(C)** Overlay of AlphaFold2 structure (blue) and SWISS-MODEL structure (red). **(D)** Overlay of active site of the AlphaFold2-generated (blue carbon atoms) and SWISS-MODEL (orange carbon atoms) active sites. The zinc ions added to the AlphaFold2 structure are shown as magenta balls.

### 3.5. The aminoacylases could be overexpressed in various hosts

The approach of this thesis was to recombinantly produce the selected aminoacylase candidates. As *E. coli* is the most common host for heterologous production, we chose this organism for the initial expression of all selected enzymes. From the protein sequences of selected aminoacylases, the DNA sequence was deduced by *in vitro* reverse translation, with the codon usage adjusted to the bias of *E. coli*. Furthermore, to enable one-step affinity purification of the

### 3. Discussion - 3.5. The aminoacylases could be overexpressed in various hosts

recombinant proteins, sequences coding for Strep-tag II affinity tags (WSHPQFEK) with a linker (SG) were attached to either the 5'- or 3'-end of the genes. Additionally, the aminoacylases were cloned without any tag, to control whether the affinity tag has influence on expression and activity. The strain *E. coli* BL21(DE3) was used in combination with the T7 expression system through pET-plasmid based expression. Hence, the aminoacylase genes were cloned into the plasmid pET28a-eforRED [173], which carries a chromoprotein for screening of successful cloning, and BsaI restriction sites that enable Golden Gate cloning.

In the initial expression studies, all aminoacylase genes were expressed in *E. coli* BL21(DE3) in Terrific Broth (TB) medium and induced with 1 mM IPTG. Strong overexpression was observed for all enzymes, but only inclusion bodies were formed so that no soluble aminoacylases could be detected. The induction with IPTG is known to be a possible cause of problems in recombinant expression. Despite its advantage of having the potential of high-level expression, the high promoter strength can lead to an overload of the expression machinery [174]. IPTG is an efficient inducer and is not metabolically degraded, and even has a toxic effect on the cells [175]. High levels of mRNA transcript leads to high levels of translation, which causes metabolic burden to occur, accompanied by overloading of protein folding apparatus [176]. Lowering IPTG concentration from 1 mM to 0.1 mM did not lead to soluble protein, only a lesser total amount of inclusion bodies.

Several means exist to improve soluble expression in *E. coli*. The cultivation or expression conditions can be adjusted regarding temperature, inducer, or medium composition. Different *E. coli* strains with distinct characteristics can be used, which carry the DE3 prophage to be compatible with the T7 expression system. Examples for derivatives of the *E. coli* B and K-12 strain series are *E. coli* Tuner(DE3) or *E. coli* JM109(DE3), respectively [177]. The Tuner strain has a *lacY* deletion, which encodes the lactose permease, so that intracellular IPTG levels, mediated by passive diffusion through the cell membrane, can be adjusted by inducer concentration [124]. The co-expression of chaperones is also a crucial approach to enhance expression of aggregation-prone proteins [127]. Codon-optimization has been deliberately conducted, but it has been reported that introduction of rare codons can improve soluble production due to slowed down translation [178]. The recombinant protein can also be fused to a solubility-enhancing protein, like the maltose binding protein [179]. There are further ways to improve heterologous expression for special cases, like toxic proteins, genes with rare codons, especially eukaryotic genes, disulfide-containing proteins, or proteins susceptible to degradation by proteases.

### 3. Discussion - 3.5. The aminoacylases could be overexpressed in various hosts

In this thesis, the approaches chosen were changing expression temperature, inducer, the use of various strains and alternative expression hosts, co-expression of chaperones, and combinatorial approaches. Because the induction with IPTG only led to the formation of inclusion bodies, both at 1 mM and 0.1 mM concentrations, the inducer was changed to autoinduction with lactose. This is realized by adding 0.2 % lactose to the medium, which already contains glucose. As long as glucose is still present, the *lac* operon is repressed, so that no lactose is taken up or converted to allolactose. Once glucose is consumed, the *lac* operon is activated and eventually the heterologous expression is induced [125]. Glycerol is catabolized independently. Furthermore, *V. natriegens* Vmax™ was used as an alternative host to *E. coli*. Since this organism does not have a *lac* operon, autoinduction was not possible with *V. natriegens*.

For the soluble expression of PmAcy, a combinatorial approach was necessary: With IPTG and lactose autoinduction, no soluble protein was produced, both at 37 °C and at 20 °C. Only when the enzyme was produced at 20 °C by lactose-autoinduction, and the chaperonine GroEL/S was co-expressed, soluble protein could be obtained. Other chaperones were not tested in this case. The expression of PmAcy with *V. natriegens* Vmax™ did not lead to an overexpression, neither as soluble enzyme, nor as inclusion bodies (data not shown), the underlying reasons being unclear.

A systematic expression study was conducted for the enzyme MsAA (Chapter 3). The study comprises the use of various *E. coli* strains, comparison of IPTG- and lactose induction, varying temperatures, *Vibrio natriegens* as an alternative expression host, and chaperone co-expression in both organisms. In *E. coli*, IPTG induction led to only inclusion bodies, without soluble protein detectable via SDS-PAGE and no aminoacylase activity in the soluble fraction of the cell-free extract. In contrast, *V. natriegens* readily produces soluble enzyme upon induction with IPTG. The co-expression of GroEL/S in *V. natriegens* led to a 1.8-fold increase in activity and specific activity in cell-free extract at 30 °C cultivation temperature. Conversely, co-expression of DnaK/J/GrpE and DnaK/J/GrpE in combination with GroEL/S were detrimental to aminoacylase activity, with no soluble enzyme or activity detected. Similar effect of the chaperone co-expression on aminoacylase expression was observed in *E. coli*. Aminoacylase activity and specific activity were increased 1.8-fold and 1.6-fold with GroEL/S co-expression, respectively, at 30 °C. At 20 °C, overall protein expression was higher, but effect of GroEL/S was not as pronounced. GroEL/S has optimum at 30 °C and might not work as efficiently at 20 °C [180]. Furthermore, protein formation at 20 °C is slower, possibly improving folding of

the recombinant aminoacylase. In an effort to combine the findings and further improve soluble expression, *E. coli* ArcticExpress(DE3) was used. This organism constitutively expresses the cold-adapted chaperonins Cpn60/10 from *Oleispira antarctica*, which are active at cold temperatures [180, 181]. This even led two times the aminoacylase activity than that observed with co-expression of GroEL/S in *E. coli* BL21(DE3) at 20 °C. The specific activity of MsAA purified from the described organisms was identical. In conclusion, *E. coli* ArcticExpress(DE3) is best used for overnight expression of MsAA at 12 °C. For short expression times of 4 h, *V. natriegens* Vmax™ with GroEL/S co-expression at 30 °C is the most suitable host and protocol. Because lactose autoinduction was beneficial for expressions with *E. coli*, it would be interesting to perform autoinduction in *V. natriegens* as well. However, this organism does not carry a *lac*-operon and cannot utilize lactose [136]. Hence, future work could aim at integrating a *lac*-operon, possibly sourced from *E. coli*, into the genome of *V. natriegens*. This would allow autoinduction and could further improve the heterologous expression of soluble enzyme.

Attempts to optimize expression, however, did not always lead to soluble recombinant protein. The aminoacylase MsELA was expressed in *E. coli* BL21(DE3) with 1 mM or 0.1 mM IPTG, lactose autoinduction, at 20 °C and 37 °C, with no success in soluble expression. The co-expression of GroEL/S, or expression with *E. coli* ArcticExpress(DE3), did not lead to soluble protein either. The co-expression of other chaperones or expression in *V. natriegens* were not tested. Therefore, this aminoacylase could not be purified and biochemically characterized.

The streptomycetal aminoacylases SgAA and SgELA were also prone to form inclusion bodies during expression in *E. coli*. To attempt soluble expression of these enzymes, the same approaches as previously described were followed. The expression was induced with either IPTG or lactose autoinduction in *E. coli* BL21(DE3). Expression temperatures of 37 °C, 30 °C, and 20 °C were tested, as well as co-expression of GroEL/S and the use of *E. coli* ArcticExpress(DE3). Under all conditions, only insoluble inclusion bodies were obtained. In the publications from Koreishi *et al.*, *S. lividans* TK24 was used as the expression host for the aminoacylases SmAA [69] and SmELA [30] from *S. mobaraensis*. In a patent, SmELA and the homologous ScELA were also expressed in *E. coli* JM109 [77]. The productivity was highest in recombinant *S. lividans* with 210 U per ml of culture, compared to 6 U/ml for *E. coli* and 1.4 U/ml with the natural producer *S. mobaraensis* IFO 13819. Heterologous expression can be more successful when the host and donor are closely related. Hence, we chose *S. lividans* TK23 for expression of the aminoacylases from *S. griseus* DSM 40236<sup>T</sup>. The genes were amplified from genomic DNA of *S. griseus* DSM 40236<sup>T</sup> and cloned into the expression vector pGH01,

### 3. Discussion - 3.6. Development of an optimized ninhydrin-based activity assay

which is based on pEM4 [182]. This was done despite the corresponding genes being already synthesized for *E. coli* expression in order to preserve the streptomycetal, high G+C codon bias. The aminoacylase genes were under control of the constitutive *ermE\** promoter from *Saccharopolyspora erythraea* [183]. During cultivation of *S. lividans*, due to the filamentous morphology of the bacteria, the mycelium can adapt either a dispersed or pelleted growth. In general, evenly dispersed mycelium is considered best for protein production [148]. For cultivation and expression in shake flasks, a high-sucrose medium was used. The high sucrose concentration of 34 % leads to dispersed growth, and the sugar is not catabolized by *S. lividans* [145]. In contrast, when using medium that does not contain sucrose, pelleted growth was observed. From the expression experiments performed with shake flasks, only low amount of recombinant SgAA and no SgELA could be obtained. To optimize the culture conditions and ensure longer periods of controlled growth, fermentations in bioreactors were performed. For this, a new growth medium was developed by combining compositions of YEME and TSB media [145]. Sucrose composition was lowered from 34 % to 10 %, as overall osmolarity was higher so that no growth was observed with 34 % sucrose. The stirrer speed was controlled by dissolved oxygen concentration to maintain 30 % saturation. A low glucose feed was set to ensure supply of nutrient without increasing the volume too significantly. After 3 days fermentation, soluble SgAA could be purified, yielding 14 mg from 400 ml lysate, which corresponds to 83 g wet cell weight. From the production of SgELA, only a small amount of protein could be purified after 3 days, which was insufficient for characterization. Thus, fermentation time was extended to 5 days, so that soluble protein could be purified. However, only 0.13 mg pure SgELA was obtained.

### **3.6. Development of an optimized ninhydrin-based activity assay**

Activity assays are crucial when working with enzymes. They are used to determine activity with various substrates, the presence and amount of enzyme in a sample, or to characterize the enzymes. This can be done by directly detecting reaction products, decrease of substrate concentration, following the consumption of a cofactor, or by using chromogenic substrates. For aminoacylases, enzyme activity can refer to either hydrolytic or synthetic reaction. Despite synthesis being the intended use of the aminoacylases and this thesis, routine activities were based on hydrolysis, as well as the initial characterization prior to evaluation of synthesis. This is due to the synthetic reaction usually being performed over a longer period of time, at least

several hours to days. Usually, the reaction mixture is then analyzed by high-performance liquid chromatography (HPLC) for reliable, direct quantification. This is, however, too time consuming for routine activity measurements. Furthermore, not all substrates might be accepted for synthesis, and optima and characteristics can be different for synthesis and hydrolysis. Synthesis often needs high substrate excess, while at low substrate concentration, hydrolysis prevails. Not least, in literature, aminoacylases are often characterized based on their hydrolytic activities, which enables comparing enzymes from this thesis.

For the detection of reaction products from the hydrolysis of N-acetyl-L-amino acids, we chose the ninhydrin assay to detect the released primary amino groups. Ninhydrin (2,2-dihydroxyindane-1,3-dione) reacts with (primary) amino groups and forms Ruhemann's purple (RP; 2-(1,3-dioxindan-2-yl)iminoindane-1,3-dione). First, indane-1,2,3-trione is formed under acidic conditions, then the amino acid's nitrogen attacks the medial carbonyl function, which leads to the formation of an aldehyde derivate of the amino acid and 2-aminoindan-1,3-dione. The latter molecule reacts with ninhydrin or 2-hydroxyindane-1,3-dione to form RP [184]. Some of the proteinogenic amino acids, like cysteine and proline, form different chromophores that have an orange-brown color. Several other ninhydrin derivatives were observed after reaction with proteinogenic amino acids, which can also produce color other than RP. This leads to slightly different visual color variations and different sensitivity for the amino acids, but in general, the products can be measured at 570 nm. Special attention had to be paid to the ninhydrin reaction with lysine in the context of aminoacylase activity. Lysine gives the same color yield as most other amino acids from ninhydrin reaction, despite the additional primary amino group in its side chain. However, either N $_{\alpha}$ -acetyl-lysine or N $_{\epsilon}$ -acetyl-lysine leads to the same color yield as free lysine. Hence, aminoacylase-catalyzed hydrolysis of N $_{\alpha/\epsilon}$ -acetyl-lysine cannot be followed by conventional ninhydrin reagents. It has been described that under strongly acidic conditions, lysine reacts selectively with ninhydrin under formation of an orange chromophore that can be measured at 470 nm [184, 185].

The ninhydrin reaction is often used to determine aminoacylase activity [51, 64, 69, 72, 74, 91, 97, 105, 107, 114, 186–190]. The aminoacylase reaction itself is independent of the ninhydrin reaction, because samples are withdrawn for batch analysis. A disadvantage of the ninhydrin assay is its susceptibility to oxidation. During the ninhydrin reaction, a reduced form of ninhydrin is formed and crucial for the color-forming reaction. For quantitative analysis of amino acid with a constant sensitivity and linear response, a reducing agent is necessary. This can for example be stannous chloride [191], borohydride [192], or cadmium chloride [193].

### 3. Discussion - 3.6. Development of an optimized ninhydrin-based activity assay

However, these reducing agents will oxidize unless the reagent is stored under nitrogen or other inert gas. We found a commercial ninhydrin reagent with a patent-protected formula that contains a temperature-sensitive reducing agent that is stable at air (EZ Nin™, invented by JPP chromatography [194], distributed by Biochrom UK). This eliminates the need for frequent preparation of reagent and amino acid standards. Hence, we used the reagent and developed a protocol for straight-forward aminoacylase activity assays. However, the reagent was intended for use in HPLC-coupled automated amino acid analyzers and needed modification to be suitable for batch analysis. With the pure, unmodified EZ Nin™ reagent, which mainly contains ethylene glycol as the solvent, the formed RP precipitated when the amino acid concentration in the samples was above 16 mM. Mixing the reagent with DMSO to equal parts increased the solubility of RP, thus solving the issue. Furthermore, DMSO decreased viscosity of the reagent, and more importantly can increase enzyme-inactivating properties of the reagent. The variant for aminoacylase activity against acetyl-lysine was implemented by mixing the EZ NIN™ reagent mixed with glacial acetic acid instead of DMSO. The reaction with lysine yields an orange dye, while reaction with N<sub>α</sub>-acetyl-lysine did not lead to any color formation, and reaction N<sub>ε</sub>-acetyl-lysine yielded slight purple color.

In summary, an aminoacylase activity assay was developed using an air-stable reagent that eliminates the need for reagent preparation or storing under inert gas and gassing out oxygen. Furthermore, a 96-well format was adapted that enables parallel processing with multichannel pipettes and microtiter plates. In principle, this protocol can be adapted to be used in automated liquid handling robotics, which would not be feasible with other protocols. Not least, the stability of RP could be enhanced by dilution in alkaline buffers.

A viable alternative to ninhydrin-based assay is the quantification of amino groups with 2,4,6-trinitrobenzenesulfonic acid (TNBS). The reaction of TNBS with a primary amine produces an orange-colored derivative, which can be measured at 335 nm. However, the TNBS assay has several disadvantages. The reagent is susceptible to hydrolysis, producing picric acid, which lowers the sensitivity. The quantification of amines is described to be hampered by sodium dodecyl anions [195] In general, use of TNBS was not often described for aminoacylase activity assays [63, 118, 196, 197], but rather for derivatization in HPLC analysis.

### **3.7. The purified aminoacylases were characterized and compared**

The aminoacylases PmAcy, MsAA, SgAA, and SgELA were successfully expressed and purified. Upon affinity purification, the aminoacylases were biochemically characterized. The monomeric molecular weight of PmAcy is 47.4 kDa, or 48.7 kDa with the Strep-tag, which could be confirmed by MALDI-TOF analysis (48.6 kDa). The pI was determined by isoelectric focusing and is approximately 6.6. The aminoacylase is multimeric and appears dodecameric in native PAGE with a weight of approximately 550 kDa. The homologous enzymes, namely aminoacylase from *Burkholderia* sp. LP18\_5B [33], L-pipecolic acid acylase from *Pseudomonas* sp. AK2 (LpipACY) [108], Sgx9260c and Sgx9260b [60], or isoaspartyl dipeptidase [58] were found to be octameric.

The mycobacterial aminoacylase MsAA has a monomeric molecular weight of 49.9 kDa with an attached N-terminal Strep-tag, which was confirmed by MALDI-TOF analysis. The experimental pI of MsAA was 4.3. MsAA was shown to be a dimeric enzyme, as it is often described for members of the M20A family. However, SmAA is described to be monomeric, as determined by native PAGE [69]. The aminoacylase SmAA produced in *S. mobaraensis* or *S. lividans* showed a truncated N-terminus [69]. It could be that MsAA is also truncated by the natural producer through degradation by peptidases. This was not observed with recombinant MsAA NTag, where the N-terminus is essential for affinity purification, to preserve the N-terminal Strep-tag. The aminoacylase SgAA was also determined to be dimeric by native PAGE. It can also be excluded that the N-terminus is truncated, because N-terminal affinity tag was successfully used for purification.

#### **3.7.1. The aminoacylases have different substrate specificities**

The substrate specificity of aminoacylases can be described for both the amino acid and fatty acid moiety. A possible bias towards certain amino acids can be grouped or interpreted regarding the properties of the side chain. The enzyme can hence prefer polar, charged, hydrophobic, small, or large side chains. Referring to the fatty acid specificity, aminoacylases can be short-chain or long-chain aminoacylases and have different acceptance for branched acyl chains or aromatic acyl residues.

### 3. Discussion - 3.7. The purified aminoacylases were characterized and compared

The enzyme PmAcy can be classified as a long chain aminoacylase. The substrate that was best accepted in hydrolysis was lauroyl-alanine (184.4 U/mg with 1 mM Zn<sup>2+</sup>, 238.9 U/mg with 1 mM Co<sup>2+</sup>). In contrast, acetyl-alanine was barely hydrolyzed. Benzoyl-alanine was also not accepted as a substrate. However, palmitoyl-alanine (C16) was hydrolyzed with 32.8 % activity compared to lauroyl-alanine with 60.4 U/mg (52.5 % or 125.3 U/mg with Co<sup>2+</sup>). The hydrolysis of acyl-glutamines of various acyl chain lengths could also be investigated. The aminoacylase preferred lauroyl-glutamine (69.1 U/mg) over the shorter capryloyl-glutamine (10.5 U/mg) and the longer palmitoyl-glutamine (38.8 U/mg). Regarding the amino acid moiety, a bias towards hydrophobic amino acids could be detected. As lauroyl-amino acids, favored substrates were alanine (184.4 U/mg), methionine (177.8 U/mg), isoleucine (148.9 U/mg), valine (130.6 U/mg), and glycine (123.9 U/mg). The lauroyl-amino acids with aromatic side chain were also hydrolyzed, namely phenylalanine (78.0 U/mg), tyrosine (63.5 U/mg), and tryptophan (4.9 U/mg). Among the polar amino acids, only lauroyl-serine was hydrolyzed (20.4 U/mg). No activity could be detected with lauroyl-aspartic acid, -glutamic acid, and -cysteine. Hence, no amino acids with carboxyl group in their side chain was accepted. Furthermore, no dipeptidase activity was detected with L-alanyl-L-phenylalanine and L-phenylalanyl-L-alanine. Compared to the homologous aminoacylase from *Burkholderia* sp. LP5\_18B, the substrate scope shows comparable relative activities [33]. The aminoacylase from *Burkholderia* sp. LP5\_18B had higher relative activity with lauroyl-phenylalanine (100 %, at 10 mM substrate concentration, pH 10.0), compared to PmAcy (42 %, at 15 mM substrate concentration, pH 9.0).

A few other long-chain acyl-specific aminoacylases have been described in literature. Two enzymes were characterized from *Pseudomonas diminuta*, now renamed to *Brevundimonas diminuta*. The aminoacylase I has a native weight of 300 kDa and prefers long-chain acyl-glutamates. No activity was measured with acetyl-glutamic acid and highest activity was observed against pentadecenoyl-glutamic acid [107]. Another long-chain aminoacylase II with a native weight of 220 kDa was isolated from the same organism [198]. This enzyme had a broader specificity and could hydrolyze lauroyl-glycine, -valine, -aspartic acid and -phenylalanine besides lauroyl-glutamic acid. No activity with acetyl-glutamic acid was observed either. No protein or gene sequences were published for these aminoacylases. Furthermore, a long-chain aminoacylase from *M. smegmatis* ATCC 607 was characterized [9, 105] that hydrolyzes palmitoyl-aspartate, -valine, -phenylalanine, but shows no activity with acetyl-valine, -methionine, -leucine, -aspartate, -glutamate, -tryptophan. The enzyme can even hydrolyze long acyl chain (C20) arachidoyl-aspartate. The enzyme has a native molecular

weight of 40-48 kDa, suggesting the presence as a monomeric enzyme. Again, no sequence has been published for the aminoacylase.

The aminoacylase MsAA can be considered short-middle/long chain aminoacylase. It prefers acetyl-amino acids in hydrolysis, and activity with corresponding lauroyl-amino acids is lower or non-detectable. Some exceptions of long chain acyl-amino acids that were hydrolyzed with high activity were lauroyl-methionine (54.9 U/mg), lauroyl-alanine (20.6 U/mg), palmitoyl-alanine (20.7 U/mg), and lauroyl-valine (7.1 U/mg). Among acetyl-amino acids, the best substrates were acetyl-valine (326.1 U/mg), acetyl-isoleucine (284.2 U/mg), acetyl-alanine (265.9 U/mg), acetyl-leucine (147.1 U/mg), and acetyl-methionine (140.9 U/mg). The aminoacylase has a bias for amino acid substrates with a hydrophobic side chain. Substrates with bulkier hydrophobic or aromatic side chains, however, are not hydrolyzed well, such as acetyl-phenylalanine (3.4 U/mg) or acetyl-tyrosine (1.2 U/mg), and acetyl-tryptophan was not accepted. Some polar or charged acetyl-amino acids were hydrolyzed, like acetyl-aspartate (41.0 U/mg), acetyl-glutamate (27.5 U/mg), acetyl-arginine (17.5 U/mg), acetyl-glutamine (44.6 U/mg), or acetyl-threonine (63.5 U/mg). Acetyl-cysteine was also accepted (39.4 U/mg). Furthermore, the aminoacylase MsAA was shown to be an  $\alpha$ -aminoacylase, hydrolyzing  $N_{\alpha}$ -acetyl-lysine with 18.1 U/mg but not accepting  $N_{\epsilon}$ -acetyl-lysine. In summary, despite having a bias for small, hydrophobic amino acids, MsAA has a broad substrate specificity regarding the amino acid moiety and prefers short acyl-chains in hydrolysis. In comparison to the homologous aminoacylase SmAA from *S. mobaraensis*, several differences in substrate scope can be observed under comparable assay conditions [69]. The acetyl-amino acids that were hydrolyzed best were acetyl-methionine, acetyl-cysteine, and acetyl-alanine. Among the tested lauroyl-amino acids, SmAA showed highest activity with lauroyl-alanine and lauroyl-methionine, similar to MsAA. However, in contrast to MsAA, SmAA did not hydrolyze acetyl-glutamic acid and had only low activity with acetyl-aspartic acid. The activity of SmAA with acyl-methionines of various acyl chain length has also been tested. The specific activity increased from acetyl-methionine to octanoyl-methionine, decreased with decanoyl- and lauroyl-methionine, and was slightly higher with myristoyl- and palmitoyl-methionine [69]. Hence, SmAA has a broad acyl-chain specificity and can hydrolyze short-chain acyl-amino acids, as well as middle/long chain acyl-amino acids.

The aminoacylase SgAA from *S. griseus* DSM 40236<sup>T</sup> is also homologous to MsAA, but more closely related to SmAA. Overall, both streptomycetal aminoacylases have a similar preference for the various acetyl-amino acids. The substrates of SgAA with highest activity were acetyl-

### 3. Discussion - 3.7. The purified aminoacylases were characterized and compared

methionine (301.1 U/mg), acetyl-alanine (266.0 U/mg), and acetyl-arginine (99.4 U/mg). The aminoacylase could hydrolyze acetyl-glutamic acid and acetyl-aspartic acid with activities of 2.4 U/mg and 5.9 U/mg, respectively. Thus, like MsAA, the determination of the substrate scope of SgAA revealed that the enzyme can hydrolyze broad spectrum of acetyl-amino acids. However, no activity against N $\alpha$ -acetyl-lysine was determined with SgAA, and the enzyme also had no  $\epsilon$ -acylase activity. Compared to MsAA, the acceptance for longer acyl chains was much lower under the same conditions, so that barely any lauroyl-amino acids were accepted. The activity with lauroyl-methionine and lauroyl-alanine were 1.4 U/mg and 1.1 U/mg, respectively. Hence, SgAA can be classified as a short-chain-acyl aminoacylase.

For SgELA from *S. griseus* DSM 40236<sup>T</sup>, the putative classification as an  $\epsilon$ -lysine acylase could be confirmed. The enzyme only showed activity with N $\epsilon$ -acetyl-lysine (11.1 U/mg), but not with N $\alpha$ -acetyl-lysine or any other tested N $\alpha$ -acetyl-amino acid. The specificity towards various acyl chain lengths was not tested due to unavailability of further substrates.

#### 3.7.2. Comparative biochemical characterization

The aminoacylases were biochemically characterized regarding the hydrolytic activities to determine optima for pH and temperature, stabilities, and dependency on metal ions. Some characteristics, like optimal reaction conditions, were investigated anew for the synthetic reaction.

The pH optimum of PmAcy for hydrolysis was remarkably high at pH 12.0 and PmAcy can thus be considered an alkaline enzyme. The homologous aminoacylase from *Burkholderia* sp. LP5\_18B also had a pH optimum at pH 12.0. Compared to other aminoacylases, this high optimum is unusual, and most have higher activity at neutral pH values. Interestingly, the optimal pH for synthesis of acyl-amino acids differs vastly and lies at the slightly basic pH values of 8.0. The optimal pH of MsAA was at pH 7.0 in hydrolysis, and the optimum for synthesis pH 8.0. The optimal pH of SgAA in hydrolysis was pH 7.0 – 8.0. The optimal pH of SgELA was pH 8.0, with a broader optimum. PmAcy stands out from these neutral to slightly basic optima, but reasons for this are not obvious. The aminoacylases MsAA and SgAA belong to the M20A family, whereas PmAcy belongs to the M38 family, like SgELA. Other homologous M38 peptidases, except the *Burkholderia* aminoacylase, such as isoaspartyl dipeptidase from *E. coli* [199] or LpipACY [108], have an pH optimum of 7.4 - 8.0, and 8.0, respectively. In general, the hydrolysis reaction is started by the formation of a hydroxide ion

### 3. Discussion - 3.7. The purified aminoacylases were characterized and compared

from the metal-bound water molecule, which should be formed better at alkaline conditions. Maybe a high pH has a negative influence on dimerization of M20A peptidases and could lead to falsely charged surface at the dimerization domain.

The aminoacylase PmAcy also showed extraordinary stability at basic pH values. The enzyme was stable even at pH 13.0 for 1 h, and only showed loss of activity at pH 13.0 after 24 h with 30 % residual activity and was stable at pH 12.0 for 24 h. The other aminoacylases were stable only at neutral pH values, with MsAA being most stable at pH 6.0 - 7.0, relatively stable at pH 5.0 - 10.0 for 24 h, and losing approximately 50 % activity at after 5 days incubation at pH 8.0. The aminoacylase SgAA and SgELA have not yet been investigated regarding pH stability.

The temperature optima were similar among the aminoacylases and were 70 °C, 70 °C, and 60 °C for PmAcy, MsAA, and SgAA, respectively. In contrast, the aminoacylase PmAcy was significantly more stable at higher temperatures. The enzyme showed no loss of activity after 1 h incubation at 80 °C and retains 23 % activity after 24 h incubation. After 4 days at 70 °C, 38 % residual activity was measured. This thermostability renders the aminoacylase interesting for industrial application, because acylation reactions are usually performed over a longer period, often 24 h to several days. Moreover, the stability offers the possibility enzyme recycling. An interesting observation is that the purified enzyme could be permanently activated by application of heat, with 3-fold activity after 1 h incubation at 80 °C. The reason for this is not directly evident. The aminoacylase was prone to aggregation during recombinant expression and formation of multimers, which could have been dissociated, possibly accompanied by partial unfolding and refolding, so that activity was enhanced. For the investigation of temperature stability, the enzyme was previously activated and then subsequently subjected to analysis of temperature stability. The  $\alpha$ -aminoacylases MsAA and SgAA show similar, mesophilic stability. The enzymes are stable at 40 °C for several days, but activity declines when incubating at 50 °C. This can lead to loss of activity during biocatalysis at higher temperatures. Compared to the aminoacylases described in literature, the enzyme from *Burkholderia* also exhibits good stability at high temperatures, but was only investigated for 1 h incubation time [33]. The stability of SmAA was similar to MsAA and SgAA with decreasing activity upon incubation over 50 °C [69]. The crude preparation of SamAA lost over 80 % synthetic activity after 16 h incubation at 37 °C, and was almost completely inactivated at 55 °C [31].

Both the members of M20A and M38 metallopeptidase have binuclear active sites with divalent metal ions. The dependency of activity from addition of various divalent metal ions was

### 3. Discussion - 3.7. The purified aminoacylases were characterized and compared

investigated with CaCl<sub>2</sub>, CoCl<sub>2</sub>, CuCl<sub>2</sub>, FeSO<sub>4</sub>, MgCl<sub>2</sub>, MnCl<sub>2</sub>, NiCl<sub>2</sub>, and ZnCl<sub>2</sub>. When PmAcy was purified without divalent metals, no hydrolytic activity was observed. Upon addition of ZnCl<sub>2</sub>, CoCl<sub>2</sub>, and MnCl<sub>2</sub>, activity was completely restored. High activity was measured with ZnCl<sub>2</sub> and CoCl<sub>2</sub>, and lower activity with MnCl<sub>2</sub>. The substrate scope did not change in that no additional substrates were accepted, but relative activity to the various substrates were different. With MnCl<sub>2</sub>, acyl chain acceptance was broader, meaning that e.g., lauroyl-alanine is hydrolyzed with the same activity as palmitoyl-alanine, whereas with ZnCl<sub>2</sub>, activity with palmitoyl-alanine was 32.7 % compared to lauroyl-alanine. With N-palmitoyl-glutamine, activity was even 69.9 % higher than with lauroyl-glutamine for PmAcy incubated with MnCl<sub>2</sub>, which was again not the case with ZnCl<sub>2</sub> and CoCl<sub>2</sub>. This might possibly be due to conformational change in the active site. Still, absolute specific activity not higher than with ZnCl<sub>2</sub> or CoCl<sub>2</sub>.

MsAA only accepts ZnCl<sub>2</sub>, and no other divalent metal ions increased its activity. In contrast, SgAA was inhibited by 1 mM ZnCl<sub>2</sub> and CoCl<sub>2</sub>. This effect might be caused by an excess of ions. Conserved metal-binding sites and similarity with other metallopeptidases suggest that SgAA also is a metal-dependent enzyme. In general, members of the M38 and M20 peptidase family can accept Zn<sup>2+</sup>, Co<sup>2+</sup>, or Mn<sup>2+</sup> [151]. The homologous aminoacylase SamAA shows highest synthetic activity with cobalt [68]. The porcine pAcy1 has Zn<sup>2+</sup> in its metal center, which can be substituted by Mn<sup>2+</sup> [200]. HiDapE accepts Co<sup>2+</sup>, Zn<sup>2+</sup>, Cd<sup>2+</sup>, and Mn<sup>2+</sup>, with highest activity with two Co<sup>2+</sup> ions [48].

Table 2: Optimal temperature and pH values for hydrolysis and synthesis for the enzymes characterized within this thesis. For hydrolysis and synthesis, respectively, lauroyl-alanine and lauroyl-phenylalanine were investigated for PmAcy, acetyl-alanine and lauroyl-methionine were investigated for MsAA, acetyl-alanine and acetyl-methionine were investigated for SgAA, and N<sub>ε</sub>-acetyl-lysine was used for SgELA in hydrolysis.

<b>Enzyme</b>	<b>pH optimum hydrolysis</b>	<b>pH optimum synthesis</b>	<b>Temperature optimum hydrolysis</b>
<b>PmAcy</b>	12.0	8.0	70
<b>MsAA</b>	7.0	8.0	70
<b>SgAA</b>	7.0 – 8.0	n.d.	60
<b>SgELA</b>	pH 8.0	n.d.	n.d.

### 3.8. Biocatalytic synthesis

The best results for acylation of amino acids within this thesis were achieved with PmAcy, with high levels of product formation for a diversity of acylated amino acids. Amino acids with polar or positively charged residues were acylated as well. However, glutamic acid was not accepted in synthesis. The optimal pH for the production of lauroyl-phenylalanine was at 8.0, with 48 % conversion. Interestingly, despite the hydrolytic optimum being at pH 12.0, and PmAcy being stable at alkaline conditions for a longer period, acylation yield significantly dropped at pH values higher than 10.0, with barely any acylation product detected at pH 11.0. A possible explanation might lie in the acid-base properties of the substrates. For synthesis, a deprotonated amino group is necessary, so that a free electron pair is available for nucleophilic attack. However, at high pH values, the fatty acid is also predominantly deprotonated, making the carbonyl group less reactive. Possibly, a slightly basic pH is favored to facilitate protonation of lauric acid during acylation. The enzyme accepted various acyl donors, like caprylic, lauric, palmitic, stearic, and oleic acid, with tendency of higher conversion towards longer acyl chains with conversions of 49 % to 75 %. In a screening, amino acids with highly diverse side chain properties could be efficiently acylated. Among the synthesized lauroyl-amino acids are lauroyl-arginine (61 %), lauroyl-histidine (62 %), lauroyl-leucine (55 %), and lauroyl-phenylalanine (48 %). Other products are lauroyl-lysine (30 %), lauroyl-isoleucine (14 %), lauroyl-methionine (12 %), and lauroyl-valine (8 %). The conditions for these acylations were 200 mM amino acid, 100 mM lauric acid, in 50 mM Tris-HCl pH 8.0 at 50 °C. Lauroyl-arginine precipitated during reaction because of its low solubility. This likely shifted the equilibrium towards the product side.

In biocatalytic synthesis with MsAA, a strong preference for lauroyl-methionine was observed. In a first screening at unoptimized conditions, lauroyl-methionine was produced best. Other lauroyl-amino acids produced were lauroyl-isoleucine, lauroyl-leucine, lauroyl-valine, lauroyl-alanine, and lauroyl-phenylalanine. An optimization of reaction was performed, focusing on the synthesis of lauroyl-methionine. The optimal pH was determined to be 8.0, and the reaction is very sensitive to change in pH. The optimal temperature for reaction was 40 - 45 °C. At 45 °C the speed of the reaction was higher, but the synthesis was counteracted by stability of the enzyme, so that final yield was slightly lower. The optimal substrate concentrations were 400 mM methionine and 150 mM lauric acid. A scale-up of the reaction from 0.5 ml to 20 ml was performed with agitation, conversion reaching 67 % with 100 mM product concentration. No difference between the two scales was observed. Hence, a 13.5-fold increase of product

### 3. Discussion - 3.8. Biocatalytic synthesis

concentration and a 9.0-fold increase in conversion rate was achieved. For the remaining amino acids, the conversion could also be optimized by increasing amino acids concentration, but to a lesser extent than methionine. With MsAA, a specificity for hydrophobic amino acids was observed in acylation. A bias for hydrophobic amino acids was also present in hydrolytic substrate scope, but some hydrophilic or charged acetyl-amino acids were hydrolyzed. For synthesis, acylation of methionine with various acyl donors was tested. The aliphatic fatty acids caprylic acid, decanoic acid, 10-undecenoic acid, lauric acid, palmitic acid, and oleic acid were accepted. However, for the long-chain palmitic acid and oleic acid, products were only detected with MS, not with the UV detector, and was thus not quantifiable. Acylation with the aromatic cinnamic acid or ferulic acid did not yield detectable products.

A further approach to optimization of the acylation reaction with MsAA was a two-factor variation of glycerol content and temperature. This experimental design addressed two strategies: First, glycerol can stabilize enzymes at higher temperatures [201, 202]. Second, the addition of glycerol leads to a lower water content in the reaction mixture. This can shift the equilibrium towards the synthesis, as the by-product of the condensation reaction is water. At a moderate temperature of 40 °C, synthesis could be improved by addition of glycerol, but only to a small extent. At higher temperatures, however, the effect was more pronounced. Without glycerol, only 2.4 mM lauroyl-methionine were produced at 60 °C, but with 30 % glycerol, 37.1 mM product concentration was measured. Hence, the conversion rate was improved over 15-fold and an improvement of enzyme stability can be implied. Still, acylation at 40 °C surpassed synthesis at higher temperatures.

Further insight into the acylation of methionine with lauric acid was gained by performing molecular docking experiments with MsAA. For this, the AlphaFold2-generated structure with added zinc ions was used. It was shown that M20A peptidases change conformation upon substrate binding [47]. For HiDapE, crystal structures for both conformations have been published (open: PDB 3CI1, closed with inhibitor: PDB 5VO3). By superimposing the two structures with predicted MsAA structure, we found that the MsAA structure represents the closed conformation, which is beneficial for performing the docking experiments. The metal-binding and catalytic residues of MsAA show a similar orientation as in the crystal structure of HiDapE [47]. The catalytic E157 forms interactions with both the amino group of methionine and the carboxylic group of lauric acid, confirming the putative function as an acid-base catalyst. In our proposed acylation mechanism, E157 first deprotonates the positively charged amino group of methionine, so that it can perform a nucleophilic attack on the carbonyl atom

of lauric acid. After decomposition of the tetrahedral intermediate, lauroyl-methionine is formed as the acylation product. A hydroxide remains bound in the binuclear zinc site, which is protonated by E157 to form a water molecule, the second product of the condensation reaction. This proposed mechanism represents the reverse hydrolytic mechanism proposed for HiDapE [47]. The substrate binding pocket of MsAA is bipartite in that it shows a hydrophilic and a hydrophobic cavity. These participated in binding of methionine and lauric acid, respectively. In MsAA, R328 and N315 bind  $\alpha$ -carboxylic group of methionine, and the corresponding conserved residues were also shown to have the same function in HiDapE [47]. The bias for methionine or specificity for hydrophobic amino acids could be explained by L213, L357, and W215, which bind the side chain of methionine. They are exchanged for two serines and a threonine in HiDapE, which accepts the hydrophilic diaminopimelic acid as the amino acyl moiety [47]. On the other hand, N244 of HiDapE forms hydrogen bonds with the  $\epsilon$ -carboxyl group of diaminopimelic acid [47], which is not conserved in MsAA, but exchanged to A314. These observations lead to possible targets for protein engineering in order to change the substrate scope of MsAA. This way, more target compounds could be synthesized, preferably from acylation of amino acids with hydrophilic side chains.

In an initial screening for acylation of proteinogenic amino acids with SgAA, only lauroyl-methionine was produced. However, the reaction conditions were not optimized yet. Studies to optimize the reaction conditions, as conducted with MsAA, should lead to an improvement of acylation yield, and could enable the formation other acyl-amino acids. Synthesis with SgELA, on the other hand, was not successful. This is presumably because of the low stability of the enzyme, which lost most of its activity even in cold storage. Either the enzyme itself is not stable, the preparation is not stable due to very low protein concentration, non-optimized metal ion concentration, or the affinity tag decreases stability. The homologous SmELA [69] and ScELA [77] showed extraordinary synthetic activity towards  $N_\epsilon$ -lauroyl-lysine, with conversions reaching 100 %. However, acylation was performed with cell-free extract, not with purified enzyme. The other homolog from SamELA was not active for  $N_\epsilon$ -acyl-lysine production [113].

### 3. Discussion - 3.9. Comparison to chemical synthesis

Table 3: Overview about the compounds that could be synthesized with aminoacylases within this thesis. The products in written bold were synthesized with conversions over 20 %. The synthesis of SgAA was performed with initial, unoptimized conditions.

Enzyme	Products	pH <sub>Opt</sub>	T <sub>Opt</sub> [°C]
<b>PmAcy</b>	<b>Lauroyl-arginine, -histidine, -leucine, -phenylalanine, -lysine, -isoleucine, -methionine, -valine</b> <b>Acylation of phenylalanine with caprylic-, lauric-, palmitic-, stearic-, and oleic acid</b>	8.0	-
<b>MsAA</b>	<b>Lauroyl-methionine, -valine, -alanine, -isoleucine, -leucine, -phenylalanine</b> <b>Acylation of methionine with caprylic-, decanoic-, undecenoic-, lauric-, myristic-, palmitic-, and oleic acid</b>	8.0	40
<b>SgAA</b>	Lauroyl-methionine	-	-

### 3.9. Comparison to chemical synthesis

In general, chemical synthesis with the Schotten-Baumann method is easy to perform and results in high yields. This method was conducted by the project partners at Technische Hochschule Köln to obtain reference substances and substrates, which are usually difficult to source commercially. As stated in the introduction, chemical synthesis is hazardous due to toxic chemicals and environmentally harmful due to waste products and the need for chlorinated fatty acids. Not all acyl-amino acid products can be synthesized. Furthermore, lysine has two amino groups, at the  $\alpha$ - and  $\epsilon$ -position, which can both be nucleophiles in Schotten-Baumann synthesis. Hence, for selective chemical synthesis of either N $\alpha$ -lauroyl-lysine or N $\epsilon$ -lauroyl-lysine, the use of protective groups is necessary. In biocatalysis, enzymes with opposing specificity can be used to synthesize these molecules, without protective groups. The chemical synthesis is performed with equimolar substrate concentrations. Biocatalytic synthesis can be performed equimolarly, but an excess of amino acid is beneficial, while the fatty acid inhibitory can be inhibitory.

By expanding the enzyme toolbox, it should be possible for more compound molecules to be synthesized by aminoacylases. In contrast to the other aminoacylases described in this thesis, the pAcy1 can synthesize lauroyl-glutamic acid [64]. Furthermore, protein engineering with aminoacylases should open new possibilities in synthesis. Few examples can be found for

protein engineering of aminoacylases. Mainly, variants to understand functions of certain residues were generated, investigating metal-binding residues, catalytic residues, residues, and residues for substrate binding [51, 52, 62, 203]. In one study, pAcy1 variants were generated to increase the ratio of synthesis to hydrolysis [32]. Furthermore, variants of *Thermococcus litoralis* aminoacylase TliACY have been generated that exhibit improved thermostability [120].

### 3.10. Outlook

#### 3.10.1. Aminoacylase PmAcy

Biocatalytic potential of PmAcy was extraordinary and surpasses other processes for  $\alpha$ -acylation of amino acids presented thus far, especially with the enzyme being heterologously expressed in *E. coli*. Still, the total process needs to be optimized for industrial application, both regarding recombinant expression and biocatalysis. During expression in *E. coli*, PmAcy was prone to aggregation and formation to inclusion bodies. The production of soluble enzyme could be further improved by co-expression of other chaperones apart from GroEL/S or by using *E. coli* ArcticExpress(DE3). Due to the high thermal stability of PmAcy, heat-precipitation of contaminant proteins from cell-extract can be performed, which could abolish the need for purification. It has been tested that PmAcy is still active after incubating the recombinant cell-free extract for up to 2 h at 70 °C (data not shown). This way, precipitated proteins can be separated by centrifugation. Interfering enzymes, acting on amino acids or fatty acids, could be heat-inactivated, so that the extract can be used in biocatalysis. Lyophilization of the supernatant from heat-precipitation could be used as an inexpensive enzyme preparation for industrial use.

We furthermore generated a predicted protein structure model of PmAcy with AlphaFold (Fig. 7). Since the N-terminus of the predicted structure has no fold and mainly consists of hydrophobic amino acids, we used the SignalP algorithm to predict the occurrence of a signal peptide. With SignalP-5.0 [170], a probability of 0.4855 for the occurrence of a Sec signal peptide was calculated, and with SignalP-6.0 [171], the probability is even 0.7514. Still, with 12 amino acids, this N-terminus is shorter than average length signal peptides, and no positively charged N-terminal part can be observed. Similar results are obtained with the aminoacylase from *Burkholderia* sp. LP5\_18B, which was obtained intracellularly [33]. Hence, the function

### 3. Discussion - 3.10. Outlook

as a signal peptide is unlikely, and the involvement in multimerization cannot be excluded. The truncation of this N-terminus in PmAcy might be beneficial for soluble expression and should be investigated in future work.

Regarding biocatalysis with PmAcy, reaction conditions could be optimized for each product compound, especially focusing on pH-value and substrate concentrations. The pH optimum was only tested for lauroyl-phenylalanine production. The side chains of amino acid substrates can be charged, which also depends on the pH, and may influence substrate acceptance. In the case of lauroyl-arginine, product precipitation occurs, which is beneficial for the production due to the law of mass action. The product can easily be separated by centrifugation, and new substrates can be added to the reaction mixture, thus recycling the enzyme. Further work is in progress to synthesize N $\alpha$ -lauroyl-arginine by biocatalysis with further reaction with ethanol to N $\alpha$ -lauroyl-arginine ethyl ester (LAE), catalyzed by ion-exchange resin or lipase. LAE has antimicrobial properties [204] with simultaneously low animal toxicity [205] and is thus interesting for the use as a food additive. This process would represent a green and sustainable alternative route to the chemical pathway, which esterifies the  $\alpha$ -carboxylic group with ethanol and thionyl chloride, with subsequent N $\alpha$ -acylation with acyl chlorides [206]. Furthermore, determination of the protein structure through crystallization would give more insight to the catalytic mechanism and could enable docking and protein engineering approaches. If crystallization is performed with an inhibitor or directly with substrates and products, detailed conclusion to substrate-binding residues can be drawn.

#### 3.10.2. Aminoacylase MsAA

The recombinant expression of soluble MsAA could be further optimized, as the enzyme still is found predominantly in inclusion bodies. N-terminal sequencing of the homologous aminoacylase SmAA revealed that the enzyme is truncated, both from production in *S. mobaraensis* or *S. lividans* [69]. In the AlphaFold2 structure of MsAA (Fig. 6), the N-terminus has no distinct predicted fold and implies a solvent-exposed hydrophobic stretch, which may lead to aggregation. The Strep-tag added to the N-terminus further increases its length. Since no fold or catalytic function can be assigned to these residues, this N-terminal sequences could be deleted without affecting the activity, in anticipation of improved solubility of the aminoacylase. Conversely, a signal sequence could be added to the N-terminus to enable secretion of the recombinant enzyme.

The results obtained from *in silico* docking of methionine and lauric acid reveal possible targets for rational protein engineering. Variants of MsAA could be generated by changing putative substrate-binding residues. The residues that bind the side chain of methionine were predicted to be L213 and L357. In HiDapE, the residues that bind the  $\epsilon$ -carboxylic and  $\epsilon$ -amino groups of diaminopimelic acid were from two serines in the respective position [47]. Hence, L213S and L357S are interesting variants for MsAA. As W215 also formed  $\pi$ -sulfur interactions with methionine, another variant could be W215T to again match HiDapE sequence. Furthermore, N244 was shown to bind the  $\epsilon$ -carboxyl group of the diaminopimelic acid, hence, a A314N variant of MsAA would be interesting. These engineering approaches could lead to the acceptance of glutamic acid, lysine, diaminopimelic acid, or other hydrophilic amino acids for acylation. Furthermore, a K397R variant of MsAA would be interesting to characterize, since the arginine residue in this position was described to bind the  $\alpha$ -carboxylic acid in hAcy1 [51] and HiDapE [46]. In general, mechanistic studies could also be conducted by variation of putative catalytic- and metal-binding residues of MsAA to verify their function.

Despite low sequence homology, M20A enzymes show high structural homology. Their structure can be divided into catalytic and dimerization domain [62]. In molecular docking, MsAA was shown to bind lauric acid mainly with catalytic domain, and the hydrophilic functional groups of diamino pimelic acid were bound by HiDapE with residues located in the dimerization domain [47]. This opens the possibility for chimeric enzymes, for example in form of fusions of the dimerization domain of HiDapE with the catalytic domain of MsAA to achieve the acyl acceptance of MsAA and the amino acyl acceptance of HiDapE. In a similar approach, randomized gene shuffling could be performed among various M20A aminoacylases or non-peptidase homologs.

### 3.10.3. Streptomycetal aminoacylases SgAA and SgELA

Through recombinant expression in *S. lividans* TK23, SgAA and SgELA could be successfully produced and purified. The substrate scope of SgAA was comparable to that of MsAA. A 5.6-fold higher specific activity with  $N_\alpha$ -acetyl-arginine was observed for SgAA under the same assay conditions, and the enzyme showed activity with acetyl-proline, which MsAA did not hydrolyze. The conditions for the synthetic reaction with SgAA could be improved in the future, similar to synthesis with MsAA. Furthermore, the biochemical characterization should be completed in more detail. It would be interesting to succeed in heterologous expression of

SamAA from *S. ambofaciens*, and to check if the good synthetic activity and broad synthetic substrate scope [31] is also exhibited by the recombinant enzyme. Currently, recombinant expression in *E. coli* and *S. lividans* are worked on (personal communication). As described previously, the homologous aminoacylase SmAA is N-terminally truncated [69]. Since SgAA was cloned and produced with a functional N-terminal Strep-tag, truncation can be excluded, despite also being produced by *S. lividans*. Possibly, the recombinant enzyme is not recognized by streptomycetal peptidases due to the affinity tag. The expression of a truncated version of SgAA could also lead to an improved expression. Despite the broad amino acid specificity, SamAA could not acylate the negatively charged glutamic acid or aspartic acid [31], and neither did the aminoacylases PmAcy, MsAA, SgAA. It would be interesting to investigate underlying reasons. With pAcy1, glutamic acid could be acylated, but only with low conversion rates of 1.5 % and 3.8 mM product from 1 M glutamic acid and 250 mM lauric acid methyl ester in 50 mM Tris-HCl pH 8.5 at 37 °C [207]. With SgELA, the predominant problems were insufficient stability and low concentration of purified protein. Aminoacylase production yields in fermentation were also much lower than with SgAA, even though longer fermentation time was chosen. This implies the need for further process development, and to optimize production on the molecular level by variation of promoter and plasmids, or even secretion of the aminoacylase.

### 3.11. Conclusion

The main objective of this work was to find novel aminoacylases and establish biocatalytic synthesis of acyl-amino acids. The motivation of the collaboration of academic and industrial partners laid within the sustainable synthesis of green surfactants. In this thesis, four novel aminoacylases were found and successfully produced and characterized. Among the achievements of this work was the establishment of a platform for recombinant expression of aminoacylases. This comprises various strains of *E. coli*, and furthermore *V. natriegens* and *S. lividans* as alternative expression hosts, and the use of chaperone toolboxes. Both the use of chaperone co-expression and aminoacylase expression in *V. natriegens* is a scientific novelty. For the characterization of the enzymes, I developed an improved ninhydrin-based aminoacylase assay, which allowed for high throughput. Biocatalytic acylation reaction were established with our aminoacylases, along with systematic optimization of the reaction parameters. The achieved conversions were very competitive to values described in literature,

often outperforming other syntheses. With the PmAcy aminoacylase, I could establish a process for  $\alpha$ -acylation of amino acids that is unparalleled in literature, regarding enzyme production and synthetic yields. Furthermore, I self-initiated an international academic collaboration. During my research stay, we not only improved acylation with MsAA, but also performed *in silico* experiments for molecular docking with MsAA. We described docking experiments with regard to synthesis and proposed the first acylation mechanism for aminoacylases.

The work and the finding of this thesis represent significant advances in the field of aminoacylase biocatalysis, both methodological and in application. These pave the way for future work on improving the catalysts and reactions, and on further exploration of the aminoacylase sequence space.

## 4. References

1. Campos JM, Stamford TLM, Sarubbo LA, Luna JM de, Rufino RD, Banat IM. Microbial biosurfactants as additives for food industries. *Biotechnol Prog.* 2013;29:1097–108. doi:10.1002/btpr.1796.
2. Ananthapadmanabhan KP. Amino-Acid Surfactants in Personal Cleansing (Review). *Tenside Surf. Det.* 2019;378 -386. doi:10.3139/113.110641.
3. Farias CBB, Almeida FC, Silva IA, Souza TC, Meira HM, Da Soares Silva RdCF, et al. Production of green surfactants: Market prospects. *Electronic Journal of Biotechnology.* 2021;51:28–39. doi:10.1016/j.ejbt.2021.02.002.
4. Nagtode VS, Cardoza C, Yasin HKA, Mali SN, Tambe SM, Roy P, et al. Green surfactants (biosurfactants): A petroleum-free substitute for sustainability-Comparison, applications, market, and future prospects. *ACS Omega.* 2023;8:11674–99. doi:10.1021/acsomega.3c00591.
5. König WA, Aydin M, Lucht N, Winkelmann G, Lupp R, Jung G, editors. Structure elucidation of the antimycetic glycolipodepsipeptide herbicolin A. Odessa: De Gruyter; 1985.
6. Pinheiro L, Faustino C. Amino acid-based surfactants for biomedical applications. In: Najjar R, editor. *Application and characterization of surfactants.* London: IntechOpen; 2017. p. 207–232. doi:10.5772/67977.
7. Czakaj A, Jarek E, Krzan M, Warszyński P. Ethyl lauroyl arginate, an inherently multicomponent surfactant system. *Molecules* 2021. doi:10.3390/molecules26195894.
8. Verdier-Sévrain S, Bonté F. Skin hydration: a review on its molecular mechanisms. *J. Cosm. Dermat.* 2007;5:75–82.
9. Nagai S, Matsuno J. Enzyme hydrolyzing N-long chain acyl-L-aspartic acids from *Mycobacterium smegmatis*: Purification and specificity of the enzyme, and the effect of alkaline metal ions on its activity. *J Biochem.* 1964;56:465–76. doi:10.1093/oxfordjournals.jbchem.a128018.
10. Natsch A, Gfeller H, Gyax P, Schmid J, Acuna G. A specific bacterial aminoacylase cleaves odorant precursors secreted in the human axilla. *J. Biol. Chem.* 2003;278:5718–27. doi:10.1074/jbc.M210142200.
11. Battista N, Bari M, Bisogno T. N-Acyl amino acids: Metabolism, molecular targets, and role in biological processes. *Biomol.* 2019. doi:10.3390/biom9120822.

12. Tan B, Yu YW, Monn MF, Hughes HV, O'Dell DK, Walker JM. Targeted lipidomics approach for endogenous N-acyl amino acids in rat brain tissue. *J Chromatogr. B.* 2009;877:2890–4. doi:10.1016/j.jchromb.2009.01.002.
13. Lin H, Long JZ, Roche AM, Svensson KJ, Dou FY, Chang MR, et al. Discovery of hydrolysis-resistant isoindoline N-acyl amino acid analogues that stimulate mitochondrial respiration. *J Medic Chem.* 2018;61:3224–30. doi:10.1021/acs.jmedchem.8b00029.
14. Yan H-D, Ishihara K, Serikawa T, Sasa M. Activation by N-acetyl-L-aspartate of acutely dissociated hippocampal neurons in rats via metabotropic glutamate receptors. *Epilepsia.* 2003;44:1153–9. doi:10.1046/j.1528-1157.2003.49402.x.
15. Morland C, Nordengen K. N-Acetyl-aspartyl-glutamate in brain health and disease. *Int J Mol Sci.* 2022. doi:10.3390/ijms23031268.
16. Bhandari S, Bisht KS, Merkler DJ. The biosynthesis and metabolism of the N-acylated aromatic amino acids: N-acylphenylalanine, N-acyltyrosine, N-acyltryptophan, and N-acylhistidine. *Front Mol Biosc.* 2021;8:801749. doi:10.3389/fmolb.2021.801749.
17. Kim JT, Terrell SM, Li VL, Wei W, Fischer CR, Long JZ. Cooperative enzymatic control of N-acyl amino acids by PM20D1 and FAAH. *Elife* 2020. doi:10.7554/eLife.55211.
18. Takehara M, Yoshimura I, Takizawa K, Yoshida R. Surface active N-acylglutamate: I. Preparation of long chain N-acylglutamic acid. *J. Am. Oil Chem. Soc.* 1972;49:157. doi:10.1007/BF02633785.
19. Leoncini A, Huskens J, Verboom W. Preparation of diglycolamides *via* Schotten–Baumann approach and direct amidation of esters. *Synlett.* 2016;27:2463–6. doi:10.1055/s-0035-1561495.
20. McGovern PE, Zhang J, Tang J, Zhang Z, Hall GR, Moreau RA, et al. Fermented beverages of pre- and proto-historic China. *PNAS.* 2004;101:17593–8. doi:10.1073/pnas.0407921102.
21. Greco E, El-Aguizy O, Ali MF, Foti S, Cunsolo V, Saletti R, Ciliberto E. Proteomic analyses on an ancient Egyptian cheese and biomolecular evidence of brucellosis. *Anal Chem.* 2018;90:9673–6. doi:10.1021/acs.analchem.8b02535.
22. Farag MA, Elmassry MM, Baba M, Friedman R. Revealing the constituents of Egypt's oldest beer using infrared and mass spectrometry. *Sci Rep.* 2019;9:16199. doi:10.1038/s41598-019-52877-0.
23. Heckmann CM, Paradisi F. Looking back: A short history of the discovery of enzymes and how they became powerful chemical tools. *ChemCatChem.* 2020;12:6082–102. doi:10.1002/cctc.202001107.

#### 4. References

24. Pauling L. Nature of forces between large molecules of biological interest. *Nature*. 1948;161:707–9.
25. Koshland DE. Application of a theory of enzyme specificity to protein synthesis. *Proc. Natl. Acad. Sci. USA*. 1958;44:98–104.
26. Bhanja Dey T, Banerjee R. Purification, biochemical characterization and application of  $\alpha$ -amylase produced by *Aspergillus oryzae* IFO-30103. *Biocatalysis and Agricultural Biotechnology*. 2015;4:83–90. doi:10.1016/j.bcab.2014.10.002.
27. Pandey A, Webb C, Soccol CR, Larroche C, editors. *Enzyme Technology*. New York: Springer; 2006.
28. Maurer K-H. Detergent proteases. *Curr. Opin. Biotechnol*. 2004;15:330–4. doi:10.1016/j.copbio.2004.06.005.
29. Anastas P, Eghbali N. Green chemistry: principles and practice. *Chem Soc Rev*. 2010;39:301–12. doi:10.1039/b918763b.
30. Koreishi M, Kawasaki R, Imanaka H, Imamura K, Takakura Y, Nakanishi K. Efficient N $\epsilon$ -lauroyl-L-lysine production by recombinant  $\epsilon$ -lysine acylase from *Streptomyces mobaraensis*. *J Biotechnol*. 2009;141:160–5. doi:10.1016/j.jbiotec.2009.03.008.
31. Bourkaib MC, Delaunay S, Framboisier X, Hôtel L, Aigle B, Humeau C, et al. N-acylation of L-amino acids in aqueous media: Evaluation of the catalytic performances of *Streptomyces ambofaciens* aminoacylases. *Enzyme Microb. Technol*. 2020;137:109536. doi:10.1016/j.enzmictec.2020.109536.
32. Wardenga R, Lindner HA, Hollmann F, Thum O, Bornscheuer U. Increasing the synthesis/hydrolysis ratio of aminoacylase 1 by site-directed mutagenesis. *Biochimie*. 2010;92:102–9. doi:10.1016/j.biochi.2009.09.017.
33. Takakura Y, Asano Y. Purification, characterization, and gene cloning of a novel aminoacylase from *Burkholderia* sp. strain LP5\_18B that efficiently catalyzes the synthesis of N-lauroyl-L-amino acids. *Biosci Biotechnol Biochem*. 2019;83:1964–73. doi:10.1080/09168451.2019.1630255.
34. Chibata I. Industrial application of immobilized enzyme systems. *Pure & Appl Chem*. 1978;50:667–75. doi:10.1351/PAC197850070667.
35. Kimura Y, Kobayashi Y, Adachi S, Matsuno R. Aminoacylase-catalyzed synthesis of N-acyl amino acid from fatty acid or its ethyl ester and amino acid. *Biochem. Eng. for* 2001. 1992:109–11. doi:10.1007/978-4-431-68180-9\_27.
36. Rawlings ND, Barrett AJ, Thomas PD, Huang X, Bateman A, Finn RD. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison

- with peptidases in the PANTHER database. *Nucleic Acids Res.* 2018;46:D624-D632. doi:10.1093/nar/gkx1134.
37. Rawlings ND, Bateman A. How to use the MEROPS database and website to help understand peptidase specificity. *Protein Sci.* 2021;30:83–92. doi:10.1002/pro.3948.
38. Rawlings ND, Barrett AJ. Chapter 77 - Introduction: Metallopeptidases and their clans. In: *Handbook of proteolytic enzymes* (2013). p. 325–370. doi:10.1016/B978-0-12-382219-2.00077-6.
39. Rawlings ND, Morton FR. The MEROPS batch BLAST: a tool to detect peptidases and their non-peptidase homologues in a genome. *Biochimie.* 2008;90:243–59. doi:10.1016/j.biochi.2007.09.014.
40. Le Coq J, An H-J, Lebrilla C, Viola RE. Characterization of human aspartoacylase: the brain enzyme responsible for Canavan disease. *Biochemistry.* 2006;45:5878–84. doi:10.1021/bi052608w.
41. Newman D, Abuladze N, Scholz K, Dekant W, Tsuprun V, Ryazantsev S, et al. Specificity of aminoacylase III-mediated deacetylation of mercapturic acids. *Drug Metab Dispos.* 2007;35:43–50. doi:10.1124/dmd.106.012062.
42. Botelho TO, Guevara T, Marrero A, Arêde P, Fluxà VS, Reymond J-L, et al. Structural and functional analyses reveal that *Staphylococcus aureus* antibiotic resistance factor HmrA is a zinc-dependent endopeptidase. *J Biol Chem.* 2011;286:25697–709. doi:10.1074/jbc.M111.247437.
43. Natsch A, Emter R. The specific biochemistry of human axilla odour formation viewed in an evolutionary context. *Philos Trans R Soc Lond B Biol Sci.* 2020;375:20190269. doi:10.1098/rstb.2019.0269.
44. Bertini I, Luchinat C, Monnanni R. Zinc enzymes. *J. Chem. Educ.* 1985;62:924–7. doi:10.1021/ed062p924.
45. Auld DS. Chapter 78 - Catalytic mechanisms for metallopeptidases. In: *Handbook of proteolytic enzymes* (2013). p. 370–396. doi:10.1016/B978-0-12-382219-2.00078-8.
46. Nocek BP, Gillner DM, Fan Y, Holz RC, Joachimiak A. Structural basis for catalysis by the mono- and dimetalated forms of the dapE-encoded N-succinyl-L,L-diaminopimelic acid desuccinylase. *J Mol Biol.* 2010;397:617–26. doi:10.1016/j.jmb.2010.01.062.
47. Nocek B, Reidl C, Starus A, Heath T, Bienvenue D, Osipiuk J, et al. Structural evidence of a major conformational change triggered by substrate binding in DapE enzymes: Impact on the catalytic mechanism. *Biochemistry.* 2018;574–84. doi:10.1021/acs.biochem.7b01151.s001.

#### 4. References

48. Bienvenue DL, Gilner DM, Davis RS, Bennett B, Holz RC. Substrate specificity, metal binding properties, and spectroscopic characterization of the DapE-encoded N-Succinyl-L,L-diaminopimelic acid desuccinylase from *Haemophilus influenzae*. *Biochemistry*. 2003;42:10756–63. doi:10.1021/bi034845+.
49. Schnackerz KD, Dobritzsch D. Amidohydrolases of the reductive pyrimidine catabolic pathway purification, characterization, structure, reaction mechanisms and enzyme deficiency. *Biochim Biophys Acta*. 2008;1784:431–44. doi:10.1016/j.bbapap.2008.01.005.
50. Nocek B, Starus A, Makowska-Grzyska M, Gutierrez B, Sanchez S, Jedrzejczak R, et al. The dimerization domain in DapE enzymes is required for catalysis. *PLoS ONE*. 2014;9:e93593. doi:10.1371/journal.pone.0093593.
51. Liu Z, Zhen Z, Zuo Z, Wu Y, Liu A, Yi Q, Li W. Probing the catalytic center of porcine aminoacylase 1 by site-directed mutagenesis, homology modeling and substrate docking. *J. Biochem*. 2006;139:421–30. doi:10.1093/jb/mvj047.
52. Lindner HA, Alary A, Boju LI, Sulea T, Ménard R. Roles of dimerization domain residues in binding and catalysis by aminoacylase-1. *Biochemistry*. 2005;44:15645–51. doi:10.1021/bi051180y.
53. Needleman SB, Wunsch CD. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol*. 1970;48:443–53. doi:10.1016/0022-2836(70)90057-4.
54. Jozic D, Bourenkow G, Bartunik H, Scholze H, Dive V, Henrich B, et al. Crystal structure of the dinuclear zinc aminopeptidase PepV from *Lactobacillus delbrueckii* unravels its preference for dipeptides. *Structure*. 2002;10:1097–106. doi:10.1016/s0969-2126(02)00805-5.
55. Girish TS, Gopal B. Crystal structure of *Staphylococcus aureus* metallopeptidase (Sapep) reveals large domain motions between the manganese-bound and apo-states. *J Biol Chem*. 2010;285:29406–15. doi:10.1074/jbc.M110.147579.
56. Greenblatt, H. M. et al. *Streptomyces griseus* aminopeptidase: X-ray crystallographic structure at 1.75 Å resolution. *J. Mol. Biol*. 1997. doi:10.1006/jmbi.1996.0729.
57. Gilboa R, Greenblatt HM, Perach M, Spungin-Bialik A, Lessel U, Wohlfahrt G, et al. Interactions of *Streptomyces griseus* aminopeptidase with a methionine product analogue: a structural study at 1.53 Å resolution. *Acta Crystallogr. D Biol. Crystallogr*. 2000;56:551–8. doi:10.1107/s0907444900002420.

58. Martí-Arbona R, Fresquet V, Thoden JB, Davis ML, Holden HM, Raushel FM. Mechanism of the reaction catalyzed by isoaspartyl dipeptidase from *Escherichia coli*. *Biochemistry*. 2005;44:7115–24. doi:10.1021/bi050008r.
59. Martí-Arbona R, Thoden JB, Holden HM, Raushel FM. Functional significance of Glu-77 and Tyr-137 within the active site of isoaspartyl dipeptidase. *Bioorg Chem*. 2005;33:448–58. doi:10.1016/j.bioorg.2005.10.002.
60. Xiang DF, Patskovsky Y, Xu C, Fedorov AA, Fedorov EV, Sisco AA, et al. Functional identification and structure determination of two novel prolidases from cog1228 in the amidohydrolase superfamily. *Biochemistry*. 2010;49:6791–803. doi:10.1021/bi100897u.
61. D'Ambrosio C, Talamo F, Vitale RM, Amodeo P, Tell G, Ferrara L, Scaloni A. Probing the dimeric structure of porcine aminoacylase 1 by mass spectrometric and modeling procedures. *Biochemistry*. 2003;42:4430–43. doi:10.1021/bi0206715.
62. Lindner HA, Lunin VV, Alary A, Hecker R, Cygler M, Ménard R. Essential roles of zinc ligation and enzyme dimerization for catalysis in the aminoacylase-1/M20 family. *Journal Biol Chem*. 2003;278:44496–504. doi:10.1074/jbc.M304233200.
63. Lindner HA, Täfler-Naumann M, Röhm K-H. N-acetylamino acid utilization by kidney aminoacylase-1. *Biochimie*. 2008;90:773–80. doi:10.1016/j.biochi.2007.12.006.
64. Wada E, Handa M, Imamura K, Sakiyama T, Adachib S, Matsunob R, Nakanishia K. Enzymatic synthesis of N-acyl-L-amino acids in a glycerol-water system using acylase I from pig kidney. *JAOCS*. 2002;79:41–46. doi:10.1007/s11746-002-0432-7.
65. Ferjancic-Biagini A, Giardina T, Reynier M, Puigserver A. Hog kidney and intestine aminoacylase-catalyzed acylation of L-methionine in aqueous media. *Biocatalysis and Biotransformation*. 1997;15:313–23. doi:10.3109/10242429709003197.
66. Wardenga R, Hollmann F, Thum O, Bornscheuer U. Functional expression of porcine aminoacylase 1 in *E. coli* using a codon optimized synthetic gene and molecular chaperones. *Appl Microbiol Biotechnol*. 2008;81:721–9. doi:10.1007/s00253-008-1716-7.
67. Long JZ, Svensson KJ, Bateman LA, Lin H, Kamenecka T, Lokurkar IA, et al. The secreted enzyme PM20D1 regulates lipidated amino acid uncouplers of mitochondria. *Cell*. 2016;166:424–35. doi:10.1016/j.cell.2016.05.071.
68. Bourkaib MC, Delaunay S, Framboisier X, Humeau C, Guilbot J, Bize C, et al. Enzymatic synthesis of N-10-undecenoyl-phenylalanine catalysed by aminoacylases from *Streptomyces ambofaciens*. *Process Biochem*. 2020;99:307–15. doi:10.1016/j.procbio.2020.09.009.

#### 4. References

69. Koreishi M, Nakatani Y, Ooi M, Imanaka H, Imamura K, Nakanishi K. Purification, characterization, molecular cloning, and expression of a new aminoacylase from *Streptomyces mobaraensis* that can hydrolyze N-(middle/long)-chain-fatty-acyl-L-amino acids as well as N-short-chain-acyl-L-amino acids. *Biosci Biotechnol Biochem*. 2009;73:1940–7. doi:10.1271/bbb.90081.
70. Natsch A. What makes us smell: The biochemistry of body odour and the design of new deodorant ingredients. *Chimia*. 2015;69:414. doi:10.2533/chimia.2015.414.
71. Ishikawa K, Ishida H, Matsui I, Kawarabayasi Y, Kikuchi H. Novel bifunctional hyperthermostable carboxypeptidase/aminoacylase from *Pyrococcus horikoshii* OT3. *Appl Environ Microbiol*. 2001;67:673–9. doi:10.1128/AEM.67.2.673-679.2001.
72. Sakanyan V, Desmarez L, Legrain C, Charlier D, Mett I, Kochikyan A, et al. Gene cloning, sequence analysis, purification, and characterization of a thermostable aminoacylase from *Bacillus stearothermophilus*. *Appl Environ Microbiol*. 1993;59:3878–88. doi:10.1128/aem.59.11.3878-3888.1993.
73. Jamdar SN, Are VN, Navamani M, Kumar S, Nagar V, Makde RD. The members of M20D peptidase subfamily from *Burkholderia cepacia*, *Deinococcus radiodurans* and *Staphylococcus aureus* (HmrA) are carboxydipeptidases, primarily specific for Met-X dipeptides. *Arch Biochem Biophys*. 2015;587:18–30. doi:10.1016/j.abb.2015.10.003.
74. Koreishi M, Kawasaki R, Imanaka H, Imamura K, Nakanishi K. A novel  $\epsilon$ -lysine acylase from *Streptomyces mobaraensis* for synthesis of N $\epsilon$ -acyl-L-lysines. *JAOCS*. 2005;82:631–7. doi:10.1007/s11746-005-1121-2.
75. Chibata I, Tosa T, Ishikawa T. On the substrate specificity of the bacterial  $\epsilon$ -acylase. *Arch Biochem Biophys*. 1964;104:231–7. doi:10.1016/S0003-9861(64)80008-4.
76. Padayatty JD, van Kley H. Specificity of  $\epsilon$ -peptidase of *Achromobacter pestifer* EA. *Arch Biochem Biophys*. 1967;120:296–302. doi:10.1016/0003-9861(67)90242-1.
77. Takakura Y, Nakanishi K, Suzuki S, Nio N, Koreishi M, Imamura T, Imanaka H. N $\epsilon$ -acyl-L-lysine specific aminoacylase (N $\epsilon$ -アシル-L-リジン特異的アミノアシラーゼ, in Japanese) - JP2012039878A (2009): Japanese Patent Application.
78. Koreishi M, Zhang D, Imanaka H, Imamura K, Adachi S, Matsuno R, Nakanishi K. A novel acylase from *Streptomyces mobaraensis* that efficiently catalyzes hydrolysis/synthesis of capsaicins as well as N-acyl-L-amino acids and N-acyl-peptides. *J Agr Food Chem*. 2006;54:72–8. doi:10.1021/jf052102k.
79. Koreishi M, Tani K, Ise Y, Imanaka H, Imamura K, Nakanishi K. Enzymatic synthesis of beta-lactam antibiotics and N-fatty-acylated amino compounds by the acyl-transfer

- reaction catalyzed by penicillin V acylase from *Streptomyces mobaraensis*. *Biosci Biotechnol Biochem*. 2007;71:1582–6. doi:10.1271/bbb.70052.
80. Liljeblad A, Kallio P, Vainio M, Niemi J, Kanerva LT. Formation and hydrolysis of amide bonds by lipase A from *Candida antarctica*; exceptional features. *Org Biomol Chem*. 2010;8:886–95. doi:10.1039/b920939p.
81. Jaeger K-E, Eggert T. Lipases for biotechnology. *Curr Opin Biotechnol*. 2002;13:390–7. doi:10.1016/s0958-1669(02)00341-5.
82. Kidwai M, Poddar R, Mothsra P. N-acylation of ethanolamine using lipase: a chemoselective catalyst. *Beilstein J Org Chem*. 2009;5:10. doi:10.3762/bjoc.5.10.
83. Dettori L, Jelsch C, Guiavarc'h Y, Delaunay S, Framboisier X, Chevalot I, Humeau C. Molecular rules for selectivity in lipase-catalysed acylation of lysine. *Process Biochemistry*. 2018;74:50–60. doi:10.1016/j.procbio.2018.07.021.
84. Kua GKB, Nguyen GKT, Li Z. Enzyme engineering for high-yielding amide formation: lipase-catalyzed synthesis of N-acyl glycines in aqueous media. *Angew Chem Int Ed Engl*. 2023;62:e202217878. doi:10.1002/anie.202217878.
85. Yang YB, Hu HL, Chang MC, Li H, Tsai YC. Purification and characterization of L-aminoacylase from *Alcaligenes denitrificans* DA181. *Biosci Biotechnol Biochem*. 1994;58:204–5. doi:10.1271/bbb.58.204.
86. LeClere S, Tellez R, Rampey RA, Matsuda SPT, Bartel B. Characterization of a family of IAA-amino acid conjugate hydrolases from *Arabidopsis*. *J Biol Chem*. 2002;277:20446–52. doi:10.1074/jbc.M111955200.
87. Bode ML, van Rantwijk F, Sheldon RA. Crude aminoacylase from *Aspergillus* sp. is a mixture of hydrolases. *Biotechnol Bioeng*. 2003;84:710–3. doi:10.1002/bit.10828.
88. Gentzen I, Löffler H-G, Schneider F. Aminoacylase from *Aspergillus oryzae*. Comparison with the pig kidney enzyme. *Z. Naturforsch*. 1980;35:544–50. doi:10.1515/znc-1980-7-804.
89. Jean Leclerc, Leo Benoiton. Further studies on  $\epsilon$ -lysine acylase. The  $\omega$ -N-acyl-diamino acid hydrolase activity of avian kidney. *Can. J. Biochem*. 1968;46:471–5. doi:10.1139/o68-071.
90. Kempf B, Bremer E. A novel amidohydrolase gene from *Bacillus subtilis* cloning: DNA-sequence analysis and map position of *amhX*. *FEMS Microbiol Lett*. 1996;141:129–37. doi:10.1111/j.1574-6968.1996.tb08374.x.

#### 4. References

91. Cho H-Y, Tanizawa K, Tanaka H, Soda K. Thermostable aminoacylase from *Bacillus thermoglucosidius*: purification and characterization. *Agric. Biol. Chem.* 1987;51:2793–800.
92. Ozaki K, Wetter LR. Some enzymic properties of a partially purified aminoacylase obtained from a Polish variety of rapeseed (*Brassica campestris* L.). *Can J Biochem Physiol.* 1961;39:843–53. doi:10.1139/o61-084.
93. Hani EK, Chan VL. Expression and characterization of *Campylobacter jejuni* benzoylglycine amidohydrolase (Hippuricase) gene in *Escherichia coli*. *J Bacteriol.* 1995;177:2396–402. doi:10.1128/jb.177.9.2396-2402.1995.
94. Xiang DF, Patskovsky Y, Xu C, Meyer AJ, Sauder JM, Burley SK, et al. Functional identification of incorrectly annotated prolidases from the amidohydrolase superfamily of enzymes. *Biochemistry.* 2009;48:3730–42. doi:10.1021/bi900111q.
95. Hsu S-K, Lo H-H, Kao C-H, Lee D-S, Hsu W-H. Enantioselective synthesis of L-homophenylalanine by whole cells of recombinant *Escherichia coli* expressing L-aminoacylase and N-acylamino acid racemase genes from *Deinococcus radiodurans* BCRC12827. *Biotechnol Prog.* 2006;22:1578–84. doi:10.1021/bp0601241.
96. Javid-Majd F, Blanchard JS. Mechanistic analysis of the *argE*-encoded N-acetylmethionine deacetylase. *Biochemistry.* 2000;39:1285–93. doi:10.1021/bi992177f.
97. Dion M, Loussouarn F, Batisse N, Rabiller C, Sakanyan V. Use of the overexpressed *Bacillus stearothermophilus* aminoacylase for the resolution of D,L-amino acids in conventional and non-conventional media. *Biotechnology Letters.* 1995:905–10. doi:10.1007/BF00127424.
98. Kuhns EH, Seidl-Adams I, Tumlinson JH. A lepidopteran aminoacylase (L-ACY-1) in *Heliothis virescens* (Lepidoptera: Noctuidae) gut lumen hydrolyzes fatty acid-amino acid conjugates, elicitors of plant defense. *Insect Biochem Mol Biol.* 2012;42:32–40. doi:10.1016/j.ibmb.2011.10.004.
99. Kaul R, Gao GP, Balamurugan K, Matalon R. Cloning of the human aspartoacylase cDNA and a common missense mutation in Canavan disease. *Nat Genet.* 1993;5:119–23. doi:10.1038/ng1093-118.
100. Herga S, Berrin J-G, Perrier J, Puigserver A, Giardina T. Identification of the zinc binding ligands and the catalytic residue in human aspartoacylase, an enzyme involved in Canavan disease. *FEBS Lett.* 2006;580:5899–904. doi:10.1016/j.febslet.2006.09.056.

101. Curley P, van Sinderen D. Identification and characterisation of a gene encoding aminoacylase activity from *Lactococcus lactis* MG1363. FEMS Microbiol Lett. 2000;183:177–82. doi:10.1111/j.1574-6968.2000.tb08954.x.
102. Sz wajcer E, Szewczuk A, Mordarski N. Aminoacylase from *Micrococcus agilis*. Acta Biochim Pol. 1980;27:123–34.
103. Ryazantsev S, Abuladze N, Newman D, Bondar G, Kurtz I, Pushkin A. Structural characterization of dimeric murine aminoacylase III. FEBS Lett. 2007;581:1898–902. doi:10.1016/j.febslet.2007.03.088.
104. Nagai S. Enzymatic hydrolysis of N-palmitoyl-amino acids by *Mycobacterium avium*. J Biochem. 1961:428–33.
105. Matsuno J, Nagai S. Amidohydrolases for N-short and long chain acyl-L-amino acids from *Mycobacteria*. J Biochem. 1972:269–79. doi:10.1093/oxfordjournals.jbchem.a129906.
106. Lugay JC, Aronson JN. Palo Verde (*Parkinsonia aculeata* L.) seed aminoacylase. Biochim Biophys Acta. 1969;191:397–414. doi:10.1016/0005-2744(69)90259-9.
107. Fukuda H, Iwade S, Kimura A. A new enzyme: Long acyl aminoacylase from *Pseudomonas diminuta*. J Biochem. 1982:1731–8. doi:10.1093/oxfordjournals.jbchem.a133865.
108. Hayashi J, Ichiki Y, Kanda A, Takagi K, Wakayama M. Identification, characterization, and cloning of a novel aminoacylase, L-pipecolic acid acylase from *Pseudomonas* species. J Gen Appl Microbiol. 2021;67:186–94. doi:10.2323/jgam.2020.12.001.
109. Story SV, Grunden AM, Adams MW. Characterization of an aminoacylase from the hyperthermophilic archaeon *Pyrococcus furiosus*. J Bacteriol. 2001;183:4259–68. doi:10.1128/JB.183.14.4259-4268.2001.
110. Tanimoto K, Higashi N, Nishioka M, Ishikawa K, Taya M. Characterization of thermostable aminoacylase from hyperthermophilic archaeon *Pyrococcus horikoshii*. FEBS J. 2008;275:1140–9. doi:10.1111/j.1742-4658.2008.06274.x.
111. Mori N, Enokibara S, Yamaguchi Y, Kitamoto Y, Ichikawa Y. Partial purification and substrate specificity of acylamino acid-releasing enzyme from *Rhodotorula glutinis*. Agric Biol Chem. 1990;54:263–5. doi:10.1271/bbb1961.54.263.
112. Bourkaib MC. Doctoral thesis: Procédé enzymatique intensifié et durable de N-acylation d'acides aminés en milieux éocompatibles : caractérisation, cinétiques et immobilisation de nouveaux biocatalyseurs (in French). Nancy, France; 2020.

#### 4. References

113. Dettori L, Ferrari F, Framboisier X, Paris C, Guiavarc'h Y, Hôtel L, et al. An aminoacylase activity from *Streptomyces ambofaciens* catalyzes the acylation of lysine on  $\alpha$ -position and peptides on N-terminal position. *Eng Life Sci.* 2018;18:589–99. doi:10.1002/elsc.201700173.
114. Koreishi M, Asayama F, Imanaka H, Imamura K, Kadota M, Tsuno T, Nakanishi K. Purification and characterization of a novel aminoacylase from *Streptomyces mobaraensis*. *Biosci Biotechnol Biochem.* 2005;69:1914–22. doi:10.1271/bbb.69.1914.
115. German AB, Neklyudov AD. Isolation and characterization of immobilized aminoacylase from *Streptoverticillium olivoreticuli*. *Appl Biochem Microbiol.* 2001;37:55–8.
116. Occhipinti E, Bec N, Gambirasio B, Baietta G, Martelli PL, Casadio R, et al. Pressure and temperature as tools for investigating the role of individual non-covalent interactions in enzymatic reactions *Sulfolobus solfataricus* carboxypeptidase as a model enzyme. *Biochim Biophys Acta.* 2006;1764:563–72. doi:10.1016/j.bbapap.2005.12.007.
117. Sommaruga S, Galbiati E, Peñaranda-Avila J, Brambilla C, Tortora P, Colombo M, Prosperi D. Immobilization of carboxypeptidase from *Sulfolobus solfataricus* on magnetic nanoparticles improves enzyme stability and functionality in organic media. *BMC Biotechnol.* 2014. doi:10.1186/1472-6750-14-82.
118. Pittelkow, S, Lindner H, Röhm K-H. Human and porcine aminoacylase I overproduced in a Baculovirus expression vector system: Evidence for structural and functional identity with enzymes isolated from kidney. *Protein Expr Purif.* 1998;12:269–76. doi:10.1006/prep.1997.0816.
119. Toogood HS, Hollingsworth EJ, Brown RC, Taylor IN, Taylor SJ, McCague R, Littlechild JA. A thermostable L-aminoacylase from *Thermococcus litoralis*: cloning, overexpression, characterization, and applications in biotransformations. *Extremophiles.* 2002;6:111–22. doi:10.1007/s007920100230.
120. Parker BM, Taylor IN, Woodley JM, Ward JM, Dalby PA. Directed evolution of a thermostable L-aminoacylase biocatalyst. *J Biotechnol.* 2011;155:396–405. doi:10.1016/j.jbiotec.2011.07.029.
121. Sørensen HP. Towards universal systems for recombinant gene expression. *Microb Cell Fact.* 2010;9:27. doi:10.1186/1475-2859-9-27.
122. Rosano GL, Ceccarelli EA. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front Microbiol.* 2014;5:172. doi:10.3389/fmicb.2014.00172.

123. Baneyx F. Recombinant protein expression in *Escherichia coli*. *Curr Opin Biotechnol.* 1999;10:411–21. doi:10.1016/S0958-1669(99)00003-8.
124. Turner P, Holst O, Karlsson EN. Optimized expression of soluble cyclomaltodextrinase of thermophilic origin in *Escherichia coli* by using a soluble fusion-tag and by tuning of inducer concentration. *Protein Expr Purif.* 2005;39:54–60. doi:10.1016/j.pep.2004.09.012.
125. Studier FW. Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif.* 2005;41:207–34. doi:10.1016/j.pep.2005.01.016.
126. Martínez-Alonso M, García-Fruitós E, Ferrer-Miralles N, Rinas U, Villaverde A. Side effects of chaperone gene co-expression in recombinant protein production. *Microb Cell Fact.* 2010;9:64. doi:10.1186/1475-2859-9-64.
127. Hoffmann F, Rinas U. Roles of heat-shock chaperones in the production of recombinant proteins in *Escherichia coli*. *Adv Biochem Eng Biotechnol.* 2004;89:143–61. doi:10.1007/b93996.
128. Villaverde A, Carrió MM. Protein aggregation in recombinant bacteria: biological role of inclusion bodies. *Biotechnology Letters.* 2003;25:1385–95.
129. Eberhardt F, Aguirre A, Menzella HG, Peiru S. Strain engineering and process optimization for enhancing the production of a thermostable steryl glucosidase in *Escherichia coli*. *J Ind Microbiol Biotechnol.* 2017;44:141–7. doi:10.1007/s10295-016-1866-z.
130. Sakikawa C, Taguchi H, Makino Y, Yoshida M. On the maximum size of proteins to stay and fold in the cavity of GroEL underneath GroES. *J Biol Chem.* 1999;274:21251–6. doi:10.1074/jbc.274.30.21251.
131. Anja Hoffmann, Bernd Bukau, Günter Kramer. Structure and function of the molecular chaperone Trigger Factor. *Biochim Biophys Acta.* 2010:650–61. doi:10.1016/0006-3002(62)90265-2.
132. Liu C-P, Perrett S, Zhou J-M. Dimeric trigger factor stably binds folding-competent intermediates and cooperates with the DnaK-DnaJ-GrpE chaperone system to allow refolding. *J Biol Chem.* 2005;280:13315–20. doi:10.1074/jbc.M414151200.
133. Payne WJ. Studies on bacterial utilization of uronic acids III: Induction of oxidative enzymes in a marine isolate. *J. Bacteriology.* 1958;76:301–7.
134. Eagon RG. *Pseudomonas natriegens*, a marine bacterium with a generation time of less than 10 minutes. *Am Soc Microbiol J.* 1961:736–7.

#### 4. References

135. Hoff J, Daniel B, Stukenberg D, Thuronyi BW, Waldminghaus T, Fritz G. *Vibrio natriegens*: an ultrafast-growing marine bacterium as emerging synthetic biology chassis. *Env Microbiol*. 2020;22:4394–408. doi:10.1111/1462-2920.15128.
136. Hoffart E, Grenz S, Lange J, Nitschel R, Müller F, Schwentner A, et al. High Substrate Uptake Rates Empower *Vibrio natriegens* as Production Host for Industrial Biotechnology. *Appl Environ Microbiol* 2017. doi:10.1128/AEM.01614-17.
137. Aiyar SE, Gaal T, Gourse RL. rRNA promoter activity in the fast-growing bacterium *Vibrio natriegens*. *J Bacteriol*. 2002;184:1349–58. doi:10.1128/JB.184.5.1349-1358.2002.
138. Maida I, Bosi E, Perrin E, Papaleo MC, Orlandini V, Fondi M, et al. Draft genome sequence of the fast-growing bacterium *Vibrio natriegens* strain DSMZ 759. *Genome Announc*. 2013;1:1–2. doi:10.1128/genomeA.00648-13.
139. Lee HH, Ostrov N, Wong BG, Gold MA, Khalil AS, Church GM. Functional genomics of the rapidly replicating bacterium *Vibrio natriegens* by CRISPRi. *Nat Microbiol*. 2019;4:1105–13. doi:10.1038/s41564-019-0423-8.
140. Weinstock MT, Heseck ED, Wilson CM, Gibson DG. *Vibrio natriegens* as a fast-growing host for molecular biology. *Nat Methods*. 2016;13:849–51. doi:10.1038/nmeth.3970.
141. Xu J, Dong F, Wu M, Tao R, Yang J, Wu M, et al. *Vibrio natriegens* as a pET-compatible expression host complementary to *Escherichia coli*. *Front Microbiol*. 2021;12:627181. doi:10.3389/fmicb.2021.627181.
142. Sun Y, Xu J, Zhou H, Zhang H, Wu J, Yang L. Recombinant protein expression chassis library of *Vibrio natriegens* by fine-tuning the expression of T7 RNA polymerase. *ACS Synth Biol*. 2023;12:555–64. doi:10.1021/acssynbio.2c00562.
143. Schleicher L, Muras V, Claussen B, Pfannstiel J, Blombach B, Dibrov P, et al. *Vibrio natriegens* as host for expression of multisubunit membrane protein complexes. *Front Microbiol*. 2018;9:2537. doi:10.3389/fmicb.2018.02537.
144. Stadler KA, Becker W, Darnhofer B, Birner-Gruenberger R, Zangger K. Overexpression of recombinant proteins containing non-canonical amino acids in *Vibrio natriegens*: p-azido-L-phenylalanine as coupling site for 19F-tags. *Amino Acids*. 2022;54:1041–53. doi:10.1007/s00726-022-03148-2.
145. Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA. *Practical Streptomyces genetics*. Norwich, England: Innes; 2000.

146. Hamed MB, Anné J, Karamanou S, Economou A. *Streptomyces* protein secretion and its application in biotechnology. *FEMS Microbiol Letters* 2018. doi:10.1093/femsle/fny250.
147. Valverde JR, Gullón S, Mellado RP. Modelling the metabolism of protein secretion through the Tat route in *Streptomyces lividans*. *BMC Microbiol.* 2018;18:59. doi:10.1186/s12866-018-1199-3.
148. Sevillano L, Vijgenboom E, van Wezel GP, Díaz M, Santamaría RI. New approaches to achieve high level enzyme production in *Streptomyces lividans*. *Microb Cell Fact.* 2016;15:28. doi:10.1186/s12934-016-0425-7.
149. Berini F, Marinelli F, Binda E. Streptomycetes: Attractive hosts for recombinant protein production. *Front Microbiol.* 2020;11:1958. doi:10.3389/fmicb.2020.01958.
150. Youshko MI, van Langen LM, Sheldon RA, Švedas VK. Application of aminoacylase I to the enantioselective resolution of  $\alpha$ -amino acid esters and amides. *Tetrahedron: Asymmetry.* 2004;15:1933–6. doi:10.1016/j.tetasy.2004.05.018.
151. Rawlings ND, Salvesen G, editors. *Handbook of proteolytic enzymes*: Academic Press; 2013.
152. Boyen A, Charlier D, Charlier J, Sakanyan V, Mett I, Glansdorff N. Acetylornithine deacetylase, succinyldiaminopimelate desuccinylase and carboxypeptidase G2 are evolutionarily related. *Gene.* 1992;116:1–6. doi:10.1016/0378-1119(92)90621-u.
153. Rowsell S, Pauptit RA, Tucker AD, Melton RG, Blow DM, Brick P. Crystal structure of carboxypeptidase G2, a bacterial enzyme with applications in cancer therapy. *Structure.* 1997;5:337–47. doi:10.1016/s0969-2126(97)00191-3.
154. Scherr N, Nguyen L. *Mycobacterium* versus *Streptomyces* - we are different, we are the same. *Curr Opin Microbiol.* 2009;12:699–707. doi:10.1016/j.mib.2009.10.003.
155. Chiaradia L, Lefebvre C, Parra J, Marcoux J, Burlet-Schiltz O, Etienne G, et al. Dissecting the mycobacterial cell envelope and defining the composition of the native mycomembrane. *Sci. Reports.* 2017;7:12807. doi:10.1038/s41598-017-12718-4.
156. Gillner DM, Bienvenue DL, Nocek BP, Joachimiak A, Zachary V, Bennett B, Holz RC. The dapE-encoded N-succinyl-L,L-diaminopimelic acid desuccinylase from *Haemophilus influenzae* contains two active-site histidine residues. *J Biol Inorg Chem.* 2009;14:1–10. doi:10.1007/s00775-008-0418-z.
157. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol.* 2011;7:539. doi:10.1038/msb.2011.75.

#### 4. References

158. Illergård K, Ardell DH, Elofsson A. Structure is three to ten times more conserved than sequence - A study of structural response in protein cores. *Proteins*. 2009;77:499–508. doi:10.1002/prot.22458.
159. Bertoline LMF, Lima AN, Krieger JE, Teixeira SK. Before and after AlphaFold2: An overview of protein structure prediction. *Front Bioinform*. 2023;3:1120370. doi:10.3389/fbinf.2023.1120370.
160. Guex N, Peitsch MC, Schwede T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. *Electrophoresis*. 2009;30 Suppl 1:S162-73. doi:10.1002/elps.200900140.
161. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucl Acids Res*. 2018;46:W296-W303. doi:10.1093/nar/gky427.
162. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling. *Electrophoresis*. 1997:2714–23.
163. Dalton JAR, Jackson RM. An evaluation of automated homology modelling methods at low target template sequence similarity. *Bioinformatics*. 2007;23:1901–8. doi:10.1093/bioinformatics/btm262.
164. Vitkup D, Melamud E, Moulton J, Sander C. Completeness in structural genomics. *Nat Struct Biol*. 2001;8:559–66. doi:10.1038/88640.
165. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc*. 2015;10:845–58. doi:10.1038/nprot.2015.053.
166. Yang J, Zhang Y. Protein structure and function prediction using I-TASSER. *Curr Protoc Bioinformatics*. 2015;52:5.8.1-5.8.15. doi:10.1002/0471250953.bi0508s52.
167. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596:583–9. doi:10.1038/s41586-021-03819-2.
168. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: Making Protein folding accessible to all. *Nat Meth*. 2022. doi:10.1038/s41592-022-01488-1.
169. Lin Y-F, Cheng C-W, Shih C-S, Hwang J-K, Yu C-S, Lu C-H. MIB: Metal Ion-Binding Site Prediction and Docking Server. *J Chem Inf. Model*. 2016;56:2287–91. doi:10.1021/acs.jcim.6b00407.

170. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol.* 2019;37:420–3. doi:10.1038/s41587-019-0036-z.
171. Teufel F, Almagro Armenteros JJ, Johansen AR, Gíslason MH, Pihl SI, Tsirigos KD, et al. SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nat Biotechnol.* 2022;40:1023–5. doi:10.1038/s41587-021-01156-3.
172. Stevens AO, He Y. Benchmarking the accuracy of AlphaFold 2 in loop structure prediction. *Biomolecules* 2022. doi:10.3390/biom12070985.
173. Muschallik L, Molinnus D, Jablonski M, Kipp CR, Bongaerts J, Pohl M, et al. Synthesis of  $\alpha$ -hydroxy ketones and vicinal (*R,R*)-diols by *Bacillus clausii* DSM 8716<sup>T</sup> butanediol dehydrogenase. *RSC Adv.* 2020;10:12206–16. doi:10.1039/d0ra02066d.
174. Haddadin FT, Harcum SW. Transcriptome profiles for high-cell-density recombinant and wild-type *Escherichia coli*. *Biotechnol Bioeng.* 2005;90:127–53. doi:10.1002/bit.20340.
175. Dvorak P, Chrast L, Nikel PI, Fedr R, Soucek K, Sedlackova M, et al. Exacerbation of substrate toxicity by IPTG in *Escherichia coli* BL21(DE3) carrying a synthetic metabolic pathway. *Microb Cell Fact.* 2015;14:201. doi:10.1186/s12934-015-0393-3.
176. Francis DM, Page R. Strategies to optimize protein expression in *E. coli*. *Curr Prot Protein Sci.* 2010;Chapter 5:Unit 5.24.1-29. doi:10.1002/0471140864.ps0524s61.
177. Hausjell J, Weissensteiner J, Molitor C, Halbwirth H, Spadiut O. *E. coli* HMS174(DE3) is a sustainable alternative to BL21(DE3). *Microb Cell Fact.* 2018;17:169. doi:10.1186/s12934-018-1016-6.
178. Zhong C, Wei P, Zhang Y-HP. Enhancing functional expression of codon-optimized heterologous enzymes in *Escherichia coli* BL21(DE3) by selective introduction of synonymous rare codons. *Biotechnol Bioeng.* 2017;114:1054–64. doi:10.1002/bit.26238.
179. Sørensen HP, Mortensen KK. Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. *Microb Cell Fact.* 2005;4:1. doi:10.1186/1475-2859-4-1.
180. Ferrer M, Chernikova TN, Yakimov M. M., Golyshin P. N., Timmis K. N. Chaperonins govern growth of *Escherichia coli* at low temperatures. *Nat Biotechnol.* 2003;21:1266–7. doi:10.1038/nbt1103-1266.
181. Ferrer M, Chernikova TN, Timmis KN, Golyshin PN. Expression of a temperature-sensitive esterase in a novel chaperone-based *Escherichia coli* strain. *Appl Env Microbiol.* 2004;70:4499–504. doi:10.1128/AEM.70.8.4499-4504.2004.

#### 4. References

182. Quirós LM, Aguirrezabalaga I, Olano C, Méndez C, Salas JA. Two glycosyltransferases and a glycosidase are involved in oleandomycin modification during its biosynthesis by *Streptomyces antibioticus*. *Mol Microbiol.* 1998;28:1177–85. doi:10.1046/j.1365-2958.1998.00880.x.
183. Bibb MJ, Janssen GR, Ward JM. Cloning and analysis of the promoter region of the erythromycin resistance gene (*ermE*) of *Streptomyces erythraeus*. *Gene.* 1988;38:215–26. doi:10.1016/0378-1119(85)90220-3.
184. 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. doi:10.1021/jf030490p.
185. Work E. Reaction of ninhydrin in acid solution with straight-chain amino acids containing two amino groups and its application to the estimation of  $\alpha$ -diaminopimelic acid. *Biochem J.* 1957:416–23.
186. Tanimoto K, Higashi N, Nishioka M, Ishikawa K, Taya M. Characterization of thermostable aminoacylase from hyperthermophilic archaeon *Pyrococcus horikoshii*. *FEBS J.* 2008;275:1140–9. doi:10.1111/j.1742-4658.2008.06274.x.
187. Yang YB, Hsiao KM, Li H, Yano H, Tsugita A, Tsai YC. Characterization of D-aminoacylase from *Alcaligenes denitrificans* DA181. *Biosci Biotechnol Biochem.* 1992;56:1392–5. doi:10.1271/bbb.56.1392.
188. Kołodziejczak-Radzimska A, Ciesielczyk F, Jesionowski T. A novel biocatalytic system obtained via immobilization of aminoacylase onto sol–gel derived  $ZrO_2 \cdot SiO_2$  binary oxide material: physicochemical characteristic and catalytic activity study. *Adsorption.* 2019;25:855–64. doi:10.1007/s10450-019-00085-7.
189. Li J, Zhao Z, Mo T, Wang L, Li P. Immobilization of aminoacylase on electrospun nanofibrous membrane for the resolution of DL-theanine. *J Mol Catal B: Enzymatic.* 2015;116:24–8. doi:10.1016/j.molcatb.2015.03.003.
190. Kosáry J, Sisák CS, Szajani B, Boross L. Acylation of amino acids by aminoacylase in non-conventional media. *Biocatalysis.* 1994;11:329–37. doi:10.3109/10242429409008865.
191. Zhang Y, Fu Y, Zhou S, Kang L, Li C. A straightforward ninhydrin-based method for collagenase activity and inhibitor screening of collagenase using spectrophotometry. *Anal Biochem.* 2013;437:46–8. doi:10.1016/j.ab.2013.02.030.

192. Standara S, Drdák M, Veselá M. Amino acid analysis: Reduction of ninhydrin by sodium borohydride. *Nahrung*. 1999;410–3. doi:10.1002/(SICI)1521-3803(19991201)43:6<410::AID-FOOD410>3.0.CO;2-1.
193. Rosen H. A modified ninhydrin calorimetric analysis for amino acids. *Arch Biochem Biophys*. 1957;10–5. doi:10.1016/0003-9861(57)90241-2.
194. Pitts LJ, Pallot MG, Jones P. A method for analysing amino acids and reagent for use with the same - EP 2735876 A1 (2014): European Patent Application.
195. Cayot P, Tainturier G. The quantification of protein amino groups by the trinitrobenzenesulfonic acid method: A reexamination. *Anal Biochem*. 1997;249:184–200. doi:10.1006/abio.1997.2161.
196. Weiß HM, Palm JG, Röhm K-H. Thermostable aminoacylase from *Bacillus stearothermophilus*: Significance of the metal center for catalysis and protein stability. *Biol Chem. Hoppe-Seyler*. 1995;643–9. doi:10.1515/bchm3.1995.376.11.643.
197. Stocker P, Brunel JM, Rezende L de, do Amaral AT, Morelli X, Roche P, et al. Aminoacylase 1-catalysed deacetylation of bioactives epoxides mycotoxin-derived mercapturates; 3,4-epoxyprococenes as models of cytotoxic epoxides. *Biochimie*. 2012;94:1668–75. doi:10.1016/j.biochi.2012.01.006.
198. Shintani Y, Fukuda H, Okamoto N, Murata K, Kimura A. Isolation and characterization of N-long chain acyl aminoacylase from *Pseudomonas diminuta*. *J Biochem*. 1984;637–43. doi:10.1093/oxfordjournals.jbchem.a134879.
199. Haley EE. Purification and properties of a  $\beta$ -aspartyl peptidase from *Escherichia coli*. *J Biol Chem*. 1968;243:5748–52. doi:10.1016/S0021-9258(18)91928-9.
200. Heese D, Berger S, Röhm KH. Nuclear magnetic relaxation studies of the role of the metal ion in Mn<sup>2+</sup>-substituted aminoacylase I. *Eur J Biochem*. 1990;188:175–80. doi:10.1111/j.1432-1033.1990.tb15385.x.
201. Esposito A, Comez L, Cinelli S, Scarponi F, Onori G. Influence of glycerol on the structure and thermal stability of lysozyme: a dynamic light scattering and circular dichroism study. *J Phys Chem B*. 2009;113:16420–4. doi:10.1021/jp906739v.
202. Meng F-G, Hong Y-K, He H-W, Lyubarev AE, Kurganov BI, Yan Y-B, Zhou H-M. Osmophobic effect of glycerol on irreversible thermal denaturation of rabbit creatine kinase. *Biophys J*. 2004;87:2247–54. doi:10.1529/biophysj.104.044784.
203. Lindner HA, Alary A, Wilke M, Sulea T. Probing the acyl-binding pocket of aminoacylase-1. *Biochemistry*. 2008;47:4266–75. doi:10.1021/bi702156h.

#### 4. References

204. Woodcock NH, Hammond BH, Ralyea RD, Boor KJ. Short communication: N $\alpha$ -lauroyl-L-arginine ethylester monohydrochloride reduces bacterial growth in pasteurized milk. *J Dairy Sci.* 2009;92:4207–10. doi:10.3168/jds.2009-2150.
205. Ruckman SA, Rocabayera X, Borzelleca JF, Sandusky CB. Toxicological and metabolic investigations of the safety of N- $\alpha$ -lauroyl-L-arginine ethyl ester monohydrochloride (LAE). *Food Chem Toxicol.* 2004;42:245–59. doi:10.1016/j.fct.2003.08.022.
206. Infante Martinez-Pardo MR, Contijoch Mestres A, Erra Serrabasa P. Process for the synthesis of cationic surfactants obtained from the condensation of fatty acids with esterified dibasic aminoacids, and their application as antimicrobial agents - EP0749960 A1 (1995): European Patent Application.
207. Wardenga R. Rekombinante Expression und Design der Aminoacylase 1 für die Synthese von N-Acyl-Aminosäuren. Inaugural disseration, Universität Greifswald (in German). 2008.

## 5. Abbreviations

AAS	amino acid surfactants
ATP	adenosine triphosphate
CMC	critical micelle concentration
EC	Enzyme Commission
EDTA	ethylenediaminetetraacetic acid
ELSD	evaporative light scattering detector
<i>et al.</i>	et alia
GOI	gene of interest
HPLC	high-performance liquid chromatography
IB	inclusion body
IPTG	isopropyl $\beta$ -D-1-thiogalactopyranoside
Kcx	carboxylated lysine
kDa	kilo Dalton
LAE	O-ethyl-N-lauroyl arginine
LB	Lysogeny Broth
MALDI-TOF	matrix-assisted laser desorption/ionization-time-of-flight
MIB	Metal Ion-Binding Site Prediction and Docking Server
MS	mass spectrometry
MWCO	molecular weight cut-off
NCBI	National Center for Biotechnology Information
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PDB	Protein Data Bank
pI	isoelectric point
SDS	sodium lauryl sulfate
SLES	sodium lauroyl ether sulfate
SLP	small linear plasmids
sp.	species
TB	Terrific Broth
TF	Trigger Factor
THF	tetrahydrofuran
Tris	(tris(hydroxymethyl)aminomethane
UV	ultraviolet light (detector)
Xaa	arbitrary amino acid

For amino acids, the standard one-letter code was used.

## 5. Abbreviations

### Abbreviations for enzymes:

Acy2	human aminoacylase-2 or aspartoacylase
Acy3	murine aminoacylase-3
ArgE	acetylornithine deacetylase from <i>Escherichia coli</i>
ASPA	human aspartoacylase
CalB	lipase B from <i>Candida antarctica</i>
Cc2672	carboxypeptidase from <i>Caulobacter crescentus</i>
CPG2	carboxypeptidase G2 from <i>Pseudomonas</i> sp. RS-16
CsAga	glutamine-aminoacylase <i>Corynebacterium striatum</i> Ax20
ELA	$\epsilon$ -lysine acylase
FAAH	fatty acid amide hydrolase
hAcy1	human aminoacylase-1
HiDapE	succinyl-diaminopimelate desuccinylase from <i>Haemophilus influenzae</i>
HmrA	peptidase from <i>Staphylococcus aureus</i>
IADA	isoaspartyl dipeptidase from <i>Escherichia coli</i>
ILR1	indole-3-acetic acid hydrolase from <i>Arabidopsis thaliana</i>
L-ACY-1	aminoacylase 1 from <i>Heliothis virescens</i>
LpipACY	L-pipecolic acid acylase from <i>Pseudomonas</i> sp. AK2
MsAA	$\alpha$ -aminoacylase from <i>Mycolicibacterium smegmatis</i> MKD 8
MsELA	$\epsilon$ -lysine aminoacylase from <i>Mycolicibacterium smegmatis</i> MKD 8
pAcy1	porcine aminoacylase-1 from <i>sus scrofa</i>
PepV	peptidase V from <i>Lactobacillus delbrueckii</i>
PM20D1	extracellular N-acyl-amino acid hydrolase
PmAcy	aminoacylase from <i>Paraburkholderia monticola</i> DSM 100849
SamAA	-aminoacylase from <i>Streptomyces ambofaciens</i>
SamELA	$\epsilon$ -lysine aminoacylase from <i>Streptomyces ambofaciens</i>
Sapep	metallopeptidase from <i>Staphylococcus aureus</i>
ScELA	$\epsilon$ -lysine aminoacylase from <i>Streptomyces coelicolor</i>
SgAA	aminoacylase from <i>Streptomyces griseus</i> DSM 40236 <sup>T</sup>
SgAP	aminopeptidase from <i>Streptomyces griseus</i>
SgELA	$\epsilon$ -lysine aminoacylase from <i>Streptomyces griseus</i> DSM 40236 <sup>T</sup>
Sgx9260b	prolidase from <i>Paraburkholderia phytofirmans</i>
Sgx9260c	prolidase from <i>Burkholderia lata</i>
SmAA	$\alpha$ -aminoacylase from <i>Streptomyces mobaraensis</i> IFO 13819
SmELA	$\epsilon$ -lysine aminoacylase from <i>Streptomyces mobaraensis</i> IFO13819
TliACY	Aminoacylase from <i>Thermococcus litoralis</i>

## 6. Acknowledgements

My thanks go especially to Prof. Dr. P. Siegert, not only for giving me this interesting, explorative scientific topic within the research project "LipoPep", but also for the support and independence I was granted in my work. You were always open for discussions and helped me keep a clear vision of my goals. I would also like to thank Prof. Dr. J. Bongaerts for the collaboration and discussions when the subject became microbiological again. Not least, I thank Prof. Dr. K-E. Jaeger for kindly taking over the examination.

I would like to thank all my colleagues for the great working atmosphere, especially my two Fabians, who have always been trusty colleagues and lab mates. I would also like to take this opportunity to also thank all my bachelor and master students for their active support and hard work within this project. I also thank my academic partner's at TH Köln, especially Prof. Dr. U. Schörken and T. Jolmes, and the industrial partners of this project for their collaboration.

I would also like to express my gratitude to Prof. Dr. I. Chevalot and Prof. Dr. P. Siegert for making my research stay in France possible and for the great integration into the team and the work of the LRGP. I thank all its members for the great experience, especially Yann and Catherine for the great supervision, guidance, and discussions.

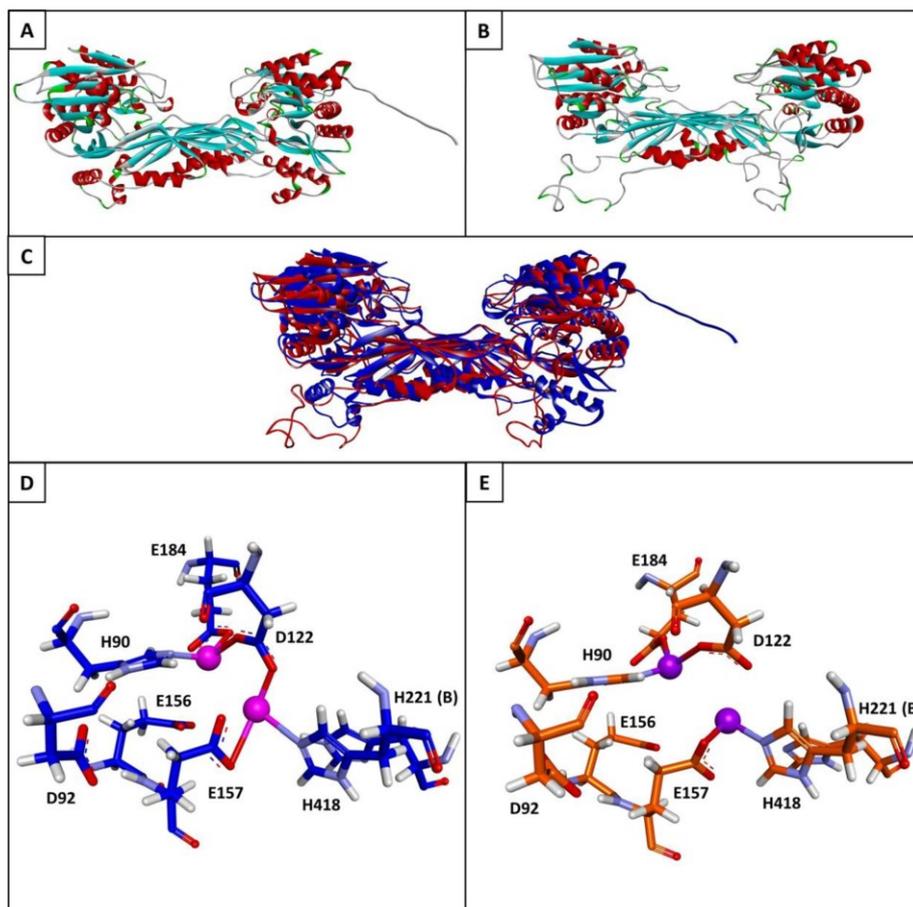
I thank all my friends, my family, and my partner Carina for supporting me on this way. You could always lift me up in bad times and were always there to celebrate my achievements. I'm looking forward to exciting times ahead.

7. Supplementary material

## **7. Supplementary material**

### Supplementary materials

AlphaFold-generated structures and homology models  
for the aminoacylases investigated in this thesis



**Fig. S1: Predicted protein structures of SgAA.**

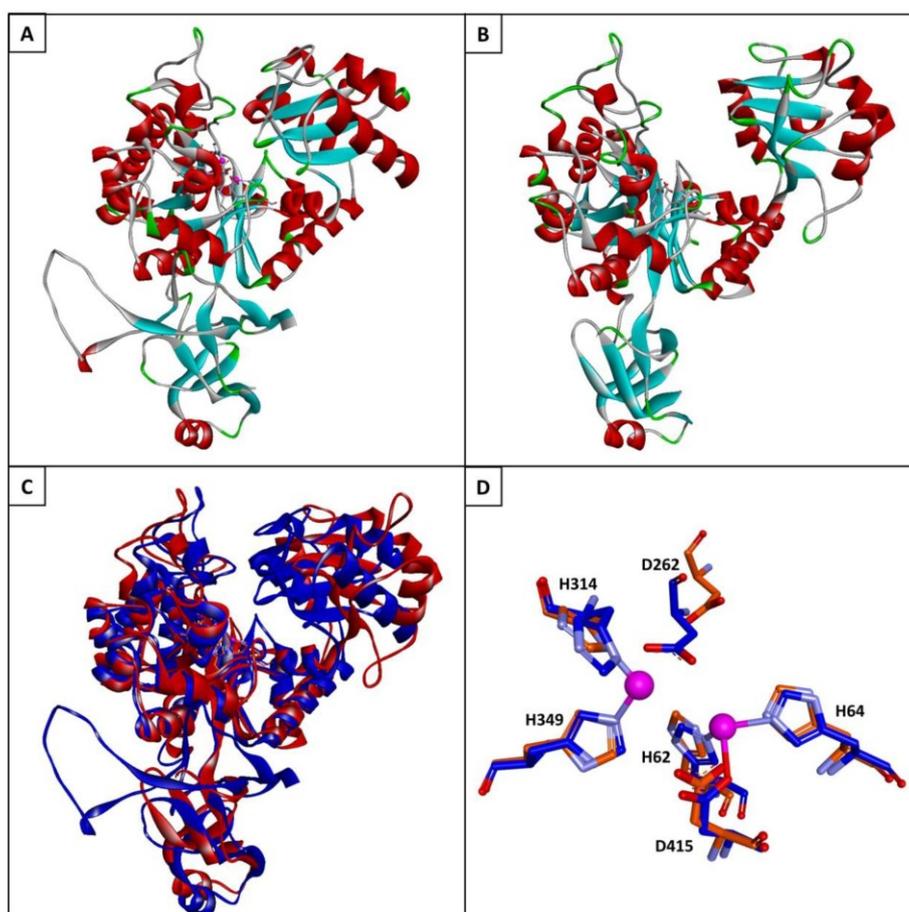
(A) AlphaFold-generated structure.

(B) Swiss-Model-generated structure. The Swiss-Model structure is truncated at the N-terminus and starts with A14.

(C) Overlay of AlphaFold structure (blue) and Swiss-Model structure (red).

(D) Active site of the AlphaFold-generated structure. The carbon atoms are shown in blue color. The zinc ions are shown as magenta balls.

(E) Active site of the Swiss-Model-generated structure. The carbon atoms are shown in orange color. The zinc ions are shown as purple balls.



**Fig. S2: Predicted protein structures of MsELA.**

(A) AlphaFold-generated structure.

(B) Swiss-Model-generated structure based on PDB 3ICJ. The Swiss-Model structure is truncated at both termini and starts at T2 and ends with E486 (instead of G530).

(C) Overlay of AlphaFold structure (blue) and Swiss-Model structure (red)

(D) Overlay of active site of the AlphaFold-generated (blue carbon atoms) and Swiss-Model (orange carbon atoms) active sites. The zinc ions added to the AlphaFold structure are shown as magenta balls.

## Supplementary materials for chapter 2

Journal name:

**Applied Microbiology and Biotechnology**

Title:

**Novel recombinant aminoacylase from *Paraburkholderia monticola* capable of N-acyl-amino acid synthesis**

Author's names:

Gerrit Haeger<sup>a</sup>, Tristan Jolmes<sup>b</sup>, Sven Oyen<sup>a</sup>, Karl-Erich Jaeger<sup>c,d</sup>, Johannes Bongaerts<sup>a</sup>, Ulrich Schörken<sup>b</sup>, and Petra Siegert<sup>a</sup>

Addresses:

<sup>a</sup>Institute of Nano- and Biotechnologies, Aachen University of Applied Sciences, 52428 Jülich, Germany

<sup>b</sup>TH Köln University of Applied Sciences - Leverkusen Campus, Faculty of Applied Natural Sciences, Leverkusen, Germany

<sup>c</sup>Institute of Molecular Enzyme Technology, Heinrich Heine University Düsseldorf, 52425 Jülich, Germany

<sup>d</sup>Institute of Bio- and Geosciences IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

Author for correspondence:

Petra Siegert; Heinrich-Mussmann-str. 1, 52428, Jülich, Germany. E-mail address: siegert@fh-aachen.de (P. Siegert), Tel.: +49 241 6009 53124

## Supplementary

Protein sequence of PmAcy (N-terminal Strep-tag and linker underlined):

MWSHPQFEKSGMLSLVLPGLSQAQSTLPPAPPAPKPVLFNFRFLFDGKSMTLRDGLYMVVEGNSISQLG  
 QGQPASVEGKTLVDCGGKVMMPGLIDMHWALLAALPIQVILQSDIAFVHLAASAEAERTLLRGFTTI  
 RDAGGPFALKQAIDSGMISGPRIPYPSGAMITTTGGHGDFRPLTDLPTSSQVTQGERDGGFAIADTE  
 DEMRVVRREQFIQGATQIKLVGSGGVSTPRSPLDMLTFTEKQLRAAVETAADWGTYYVLSHAYTPEAVQ  
 RSVAGAQCIEHGHLMDDKTAALMAKNGTWLSTQPFISEEDVGPLAPQSREKFLVAVAGTDNAFRLAR  
 KHGKVAFGTDLLEFSQAIATRQGTMLTHMKRWYSPAELGMATGTNGQLLALTGKRNPYPRRLGVLEE  
 GAYADLLLVDGNPLENLDLIANPEQNLRIVMKDGKFKYKNTLKA

Codon -optimized sequence with N-terminal strep-tag (as ordered from GeneArt, Thermo Fisher); GenBank: OR188138:

**ATGTGGTCACATCCGCAGTTTGAAAAAGCGGTATG**CTGAGCCTGGTTCTGCCTGGTCTGAG  
 CCAGGCACAGAGCACCCCTGCCCTCCGGCACCCTGCAAAACCGGTCTGTTTACCAATTTTC  
 GTCTGTTTGTATGGCAAAGCATGACCCTGCGTGATGGTCTGTATATGGTTGTTGAAGGTAAT  
 AGCATTAGCCAGTTAGGTCAGGGTCAGCCTGCAAGCGTGGAAGGTAACCCCTGGTTGATTG  
 TGGTGGTAAAGTTATGATGCCAGGCTGATGATGATATGCATTGGCATGCACTGCTGGCAGCAC  
 TGCCGATTCAGGTTATCTGCAGAGCGATATTGCATTTGTTTCATCTGGCAGCAAGCGCAGAA  
 GCAGAACGTACCCTGCTGCGTGGTTTTACCACAATTCGTGATGCCGGTGGTCCGAGCTTGC  
 ACTGAAACAGGCAATTGATAGCGGTATGATTAGCGGTCCGCTATTTATCCGAGCGGTGCAA  
 TGATTACCACAACCGGTGGTCATGGTGATTTTCGTCCGCTGACCGATCTGCCTCGTACCAGC  
 AGCCAGGTTACCCAGGGTGAACGTGATGGTGGTTTTGCAATTGCCGATACCGAAGATGAAAT  
 GCGTGTTCGTGTGCGTGAACAGTTTTATTAGGGTGCAACACAGATTAACCTGGTTGGTAGCG  
 GTGGTGTAGCACACCGCGTAGTCCGCTGGATATGCTGACCTTTACCGAAAAACAGCTGCGT  
 GCAGCAGTTGAAACCGCAGCAGATTGGGGCACCTATGTTCTGAGCCATGCATATACACCGGA  
 AGCAGTTAGCGTAGCGTTGCAGCCGGTGCACAGTGTATTGAACATGGTCATCTGATGGATG  
 ATAAAAACAGCAGCACATGATGGCCAAAAATGGCACCTGGCTGAGCACCCAGCCGTTTATTAGC  
 GAAGAAGATGTTGGTCCGCTGGCACCAGCGAGCCGTGAAAAATTTCTGGAAGTTGTTCAGG  
 CACCGATAATGCATTTCTGCTGGCACGTAAACATGGTATTAAGTTGCATTTGGTACGGATC  
 TGCTGTTTAGCCAGGCCATTGCAACCCGTGAGGCACCATGCTGACCCATATGAAACGTTGG  
 TATAGCCCTGCCGAAGCATTAGGTATGGCAACCGGCACCAATGGTCAGCTGCTGGCAGTGC  
 CGGTAAACGTAATCCGTATCCGCGTCTGTTGGTGTCTGGAAGAAGGTGCATACGCCGATT  
 TACTGCTGGTTGATGGTAACCCGCTGGAAAACTGGATCTGATTGCAAAATCCGGAACAGAAT  
 CTGCGTATTGTGATGAAAGATGGCAAGTTCTATAAGAACACCCGTAAAGCA**AGCGGTTGGAG  
 CCATCTCAGTTCGAGAAATGA**

Primer sequences for cloning:

Table S1: Primer sequences used for amplification and cloning of *pmAcy Ntag*. BsaI-sites are highlighted in bold font.

No.	Name	Sequence 5'→3'
P1	NTag for:	<b>GGTCTCC</b> CATGTGGTCACATCCGCAGTTTGAAAAAG
P2	noTag for:	<b>GGTCTCC</b> CATGCTGAGCCTGGTTCTGCC
P3	Ctag rev:	<b>GGTCTCT</b> CTCATTCTCGAACTGAGGATGGCTC
P4	noTag rev:	<b>GGTCTCT</b> CTCATGCTTTCAGGGTGTCTTATAGAACTTGC

MALDI-TOF analysis:

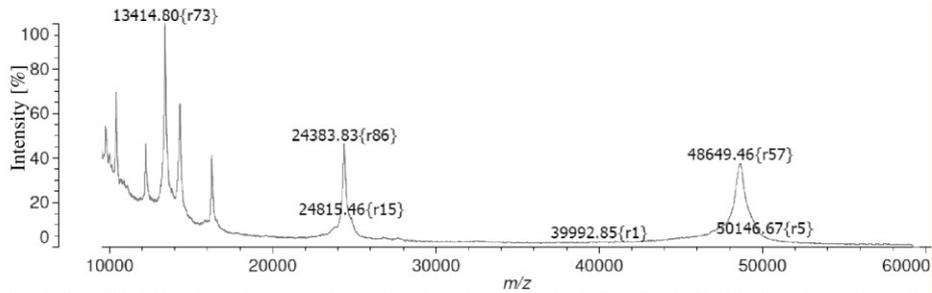


Figure S1: MALDI-TOF mass spectrum of PmAcY NTag (present in 100 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM ZnCl<sub>2</sub>); The relative intensity is plotted against the m/z-values. The numbers above the peaks indicate the measured m/z-value.

Native PAGE:

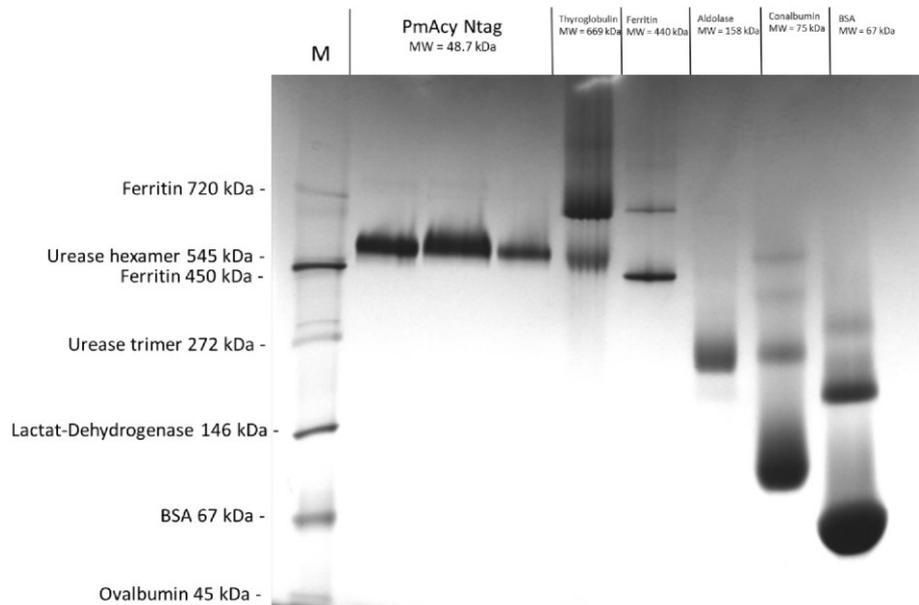


Figure S2: Blue native PAGE of PmAcY NTag and reference proteins. Lane 1: SERVA Native Marker, Liquid Mix for BN/CN (Cat.No. 39219.01); Lane 2-4: PmAcY NTag (three independent preparations); Lane 5: thyroglobulin, MW = 669 kDa; Lane 6: ferritin, MW = 440 kDa; Lane 7: aldolase, MW = 158 kDa; Lane 8: conalbumin, MW = 75 kDa; Lane 9: BSA, 67 kDa. Reference proteins from lane 5-8: Proteins from gel filtration calibration kit (High molecular weight, GE healthcare), BSA was from PanReac AppliChem).

## 7. Supplementary material

Michaelis-Menten plot:

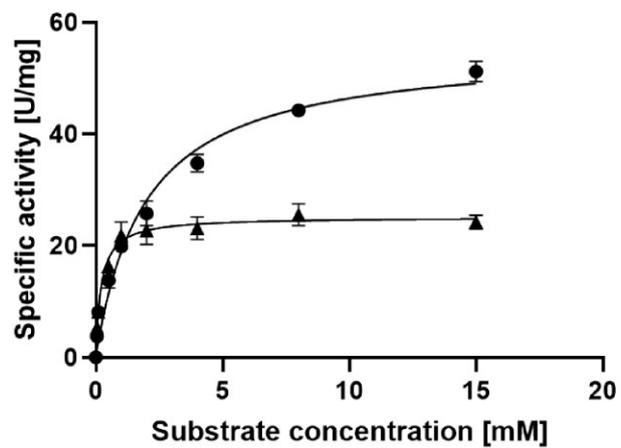
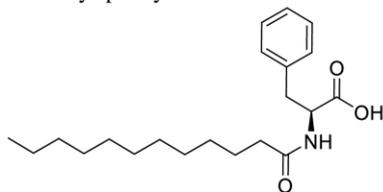


Figure S3: Michaelis-Menten hydrolysis kinetics of PmAcy NTag against lauroyl-alanine (●) and lauroyl-phenylalanine (▲). Non-linear fit to generate Michaelis-Menten plot was conducted with GraphPad Prism 8.

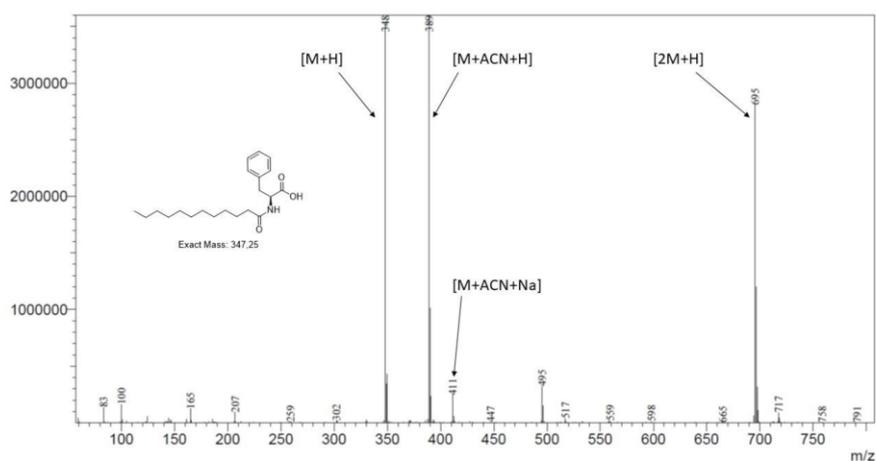
Analytical Data:  
*N*-lauroyl-phenylalanine



LC-MS: (EI, 70 eV):  $m/z$  [%] = 695 (2M+H, 80), 411 (M+ACN+Na, 10), 389 (M+ACN+H, 100), 348 (M+H, 100).

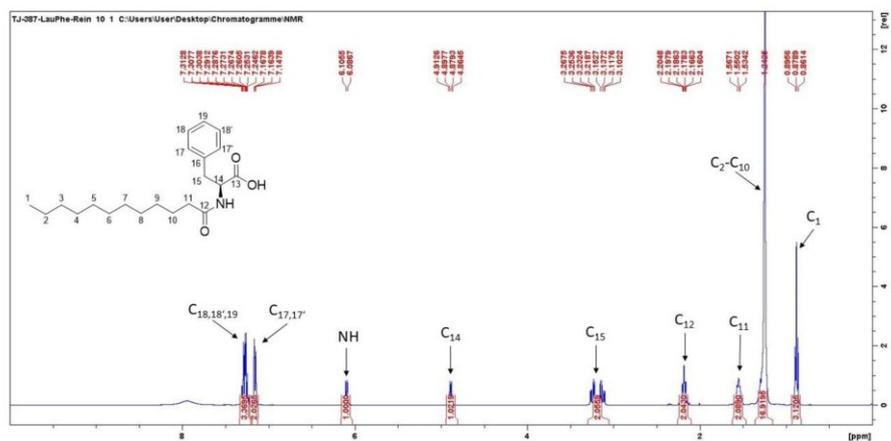
$^1\text{H-NMR}$  ( $\text{CDCl}_3$ , 400 MHz)  $\delta$  = 0.88 (t,  $^3\text{J} = 7$  Hz, 3H), 1.19 – 1.36 (m, 17 H), 1.48 – 1.59 (m, 2 H), 2.182 (dt,  $^2\text{J} = 3.2$  Hz,  $^3\text{J} = 7.2$  Hz, 2 H), 3.09 – 3.28 (m, 2 H), 4.88 (q,  $^3\text{J} = 5.9$  Hz, 1 H), 6.09 (d,  $^3\text{J} = 7.5$  Hz, 1 H), 7.13 – 7.18 (m, 2 H), 7.22 – 7.33 (m, 3 H).

$^{13}\text{C-NMR}$  ( $\text{CDCl}_3$ , 100 MHz)  $\delta$  = 14.14, 22.70, 25.59, 29.16, 29.32, 29.35, 29.48, 29.63, 29.93, 36.44, 37.27, 53.14, 127.22, 128.64, 129.38, 135.69, 174.14, 174.62.

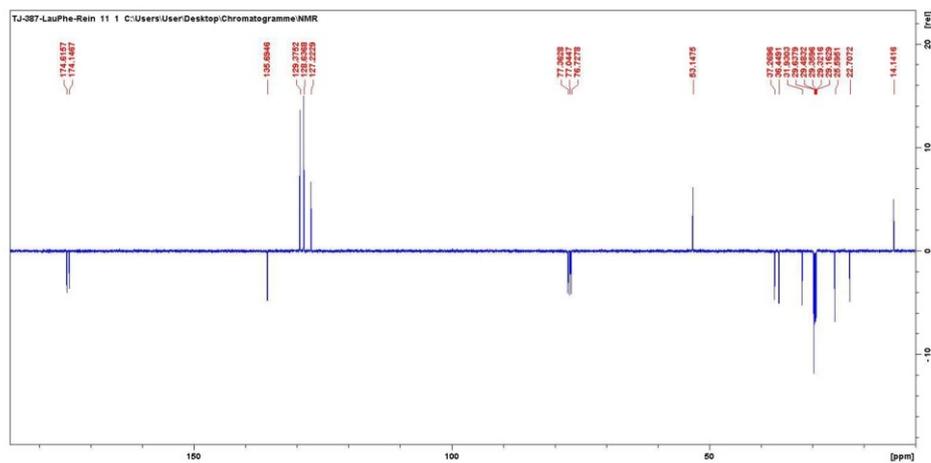


**Figure S4:** Mass-spectrum of *N*-lauroyl-phenylalanine, produced with PmAcy, isolated by acidic precipitation with 5 N HCl.

## 7. Supplementary material



**Figure S5:**  $^1\text{H-NMR}$ -spectrum of N-lauroyl-phenylalanine, produced with PmAcy, isolated by acidic precipitation with 5 N HCl.



**Figure S6:**  $^{13}\text{C-NMR}$ -spectrum of N-lauroyl-phenylalanine, produced with PmAcy, isolated by acidic precipitation with 5 N HCl.

### Supplementary materials for chapter 3

Journal name:

**Microbial Cell Factories**

Title:

**Chaperone assisted recombinant expression of a mycobacterial aminoacylase in *Vibrio natriegens* and *Escherichia coli* capable of N-lauroyl-L-amino acid synthesis**

Author's names:

Gerrit Haeger<sup>a</sup>, , Jessika Wirges<sup>a</sup>, Nicole Tanzmann<sup>a</sup>, Sven Oyen<sup>a</sup>, Tristan Jolmes<sup>b</sup>, Karl-Erich Jaeger<sup>c,d</sup>, Ulrich Schörken<sup>b</sup>, Johannes Bongaerts<sup>a</sup>, and Petra Siegert<sup>a</sup>

Addresses:

<sup>a</sup>Institute of Nano- and Biotechnologies, Aachen University of Applied Sciences, 52428 Jülich, Germany

<sup>b</sup>TH Köln University of Applied Sciences - Leverkusen Campus, Faculty of Applied Natural Sciences, Leverkusen, Germany

<sup>c</sup>Institute of Molecular Enzyme Technology, Heinrich Heine University Düsseldorf, 52425 Jülich, Germany

<sup>d</sup>Institute of Bio- and Geosciences IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

Author for correspondence:

Petra Siegert; Heinrich-Mussmann-str. 1, 52428, Jülich, Germany. E-mail address: siegert@fh-aachen.de (P. Siegert), Tel.: +49 241 6009 53124

## 7. Supplementary material

**Supplementary Material 1: MsAA protein sequence, primer sequences, gel electrophoresis results, chemically synthesized acyl-amino acids, MS spectrum for lauroyl-methionine.**

**Chaperone assisted recombinant expression in *V. natriegens* and *E. coli* and characterization of a mycobacterial aminoacylase capable of N-lauroyl-L-methionine synthesis**

Protein sequence of MsAA NTag (N-terminal Strep-tag and linker underlined):

MWSHPQFEKSGMVRMVTVTVSAASADEVVLDLVSALIRFDTSNITGDPATTKGEAECAHWVAQQLEEVGYETETEYVES  
GAPGRGNVFARLRGADPSRGALMVHGLDVPVPAEPADWSVHFFSGAVKDGYPVWGRGAVDMKDMVGMTLAVARHFK  
RAGIVPVRDLVFAFVADEEHGGTYGADWLNNRDLFEGVTEAIGEVGGFSLTVPRKDGGERRLYLIIETAEKGLS  
WMRLTARGRAGHGSVMHDDNAVTAIAGAVDRLRHEFFLVLSPAVEEFLTAVAEETGYTFDPNSPDLEGTIAKLG  
GVARIVSATLRDTANPTMLKAGYKANVI PAVAEAMI DCRVLPGRKEAFEREVDELIGPDVTRSVERDLPSYETSF  
DGDLDVAMNASVLTLDPEARIVPYMLSAGTDAKSFQRLGIRCFGFAPLRLPPDLDFAAALFHGVDERVFDALQFG  
AGVLEHFLQNC

Codon-optimized sequence with N- and C-terminal strep-tag shown in bold letters (as ordered from GeneArt, Thermo Fisher)

**ATGTGGTCACATCCGCAGTTTGAAAAAGCGGTATG**TTTCGTATGGTTACCGTTACCGTGAGCGCAGC  
AAGCGCAGATGAAGTTGTTGATCTGGTTAGCGCACTGATTCGTTTGGATACCAGCAATACCGGTGATC  
CGGCAACCACCAAAGGTGAAGCAGAATGTGCCCATTTGGGTTGCACAGCAGCTGGAAGAGGTTGGTTAT  
GAAACCGAATATGTTGAAAGCGGTGCACCTGGTCGTGGTAATGTTTTGCACGCTCGCGTGGTGCAGA  
TCCGAGCCGTGGTGCACCTGATGGTTCATGGTCATCTGGATGTTGTTCCGGCAGAACCGGCAGATTGGA  
GCGTTCATCCGTTTAGCGGTGCAGTTAAAGATGGTTATGTTTGGGGTCGTGGTGCCGTTGATATGAAA  
GATATGGTTGGTATGACCCTGGCAGTTGCACGTCATTTTTAAACGTGCAGGATTTGTTCCGCCTCGTGA  
TCTGGTGTTCGATTTGTTGCCGATGAAGAACATGGTGGCACCTATGGTGCCGATTGGCTGGTAAATA  
ATCGTCCGGACCTGTTTGAAGTGTACCGAAGCAATTGGTGAAGTTGGTGGTTTTAGCCTGACCGTT  
CCGCGTAAAGATGGCGGTGAACGTCGTCTGTATCTGATTGAAACCGCAGAAAAAGGCTGAGCTGGAT  
GCGTCTGACCGCACGTGGTCGTGCAGGTCATGGTAGCATGGTGCATGATGATAATGCAGTTACCGCAA  
TTGCCGGTTCGGTTGATCGTCTGGGTCGTATGAATTTCCGCTGGTCTGAGTCCGGCAGTTGAAGAA  
TTTCTGACAGCAGTTGCAGAAGAACCGGTTATACCTTTGATCCGAATAGTCCGGATCTGGAAGGCAC  
CATTGCAAAACTTGGTGGTGTGCACGATATGTTAGCGCAACCTGCGTGATACCGCAAATCCGACCA  
TGCTGAAAGCAGGTTATAAAGCAAATGTGATTCGGCAGTGGCAGAAGCAATGATTGATTGTCGTGTT  
CTGCCTGGTCGTAAAGAAGCATTTGAACGCGAAGTTGATGAACTGATTGGTCCGGATGTTACCCGTAG  
CTGGGAACGTGATCTGCCGAGCTATGAAACCAGCTTTGATGGCGATCTGGTTGATGCAATGAATGCAA  
GCGTTCGACCCTGGATCCGGAAGCACGCATTGTTCCGTATATGCTGAGCGCAGGCACCGATGCAAAA  
AGCTTTCAGCGTCTGGGTATTCGTTGTTTTGGTTTTGCACCGCTGCGTCTGCCACCTGATCTGGATTT  
TGCAGCACTGTTTCATGGTGTGGATGAACGTGTTCCGGTTGATGCCCTGCAGTTTGGTGCCGGTGTTC  
TGGAACATTTCCCTGCAGAATTGT**AGCGGTTGGAGCCATCCTCAGTTCGAGAAATGA**

Primer sequences for cloning:

Table S1: Primer sequences used for amplification and cloning of *msAA*.

Bsal-sites are highlighted in bold font.

No.	Name	Sequence 5'→3'
P1	NTag for:	<b>GGTCTCCCATGTGGTCACATCCGCAGTTTGAAAAAAG</b>
P2	noTag for:	<b>GGTCTCCCATGGTTCGTATGGTTACCGTTACCGTG</b>
P3	CTag rev:	<b>GGTCTCTCTCATTCTCGAACTGAGGATGGCTC</b>
P4	noTag rev:	<b>GGTCTCTCTCAACAATTCTGCAGGAAATGTTCCAGAAC</b>

Table S2: Purity and HPLC-retention time of N-lauroyl-amino acids synthesized in this study.

Synthesis was conducted via Schotten-Baumann-method utilizing caproyl chloride (C8:0), lauroyl chloride (C12:0) and palmitoyl chloride (C16:0).

Amino acid	Purity [%]	$t_{ret}$ [min]
N-lauroyl-L-alanine	94	13.3
N-palmitoyl-L-alanine	96	17.0
N-lauroyl-L-arginine	98	8.7
N-lauroyl-L-aspartic acid	89	8.5
N-lauroyl-L-cysteine	99	13.6
N-lauroyl-L-glutamic acid	97	11.2
N-lauroyl-L-glycine	99	12.8
N-lauroyl-L-histidine	60	8.6
N-lauroyl-L-isoleucine	98	14.3
N-lauroyl-L-leucine	99	14.3
N-lauroyl-L-methionine	99	13.8
N-lauroyl-L-phenylalanine	80	14.2
N-lauroyl-L-serine	94	18.6
N <sub>α</sub> -lauroyl-L-threonine	68	14.4
N <sub>α</sub> -lauroyl-L-tryptophan	98	13.8
N-lauroyl-L-tyrosine	97	14.1
N-lauroyl-L-valine	99	13.9
N <sub>α</sub> -lauroyl-L-glutamine	87	11.3
N <sub>α</sub> -palmitoyl-L-glutamine	82	14.6
N <sub>α</sub> -caproyl-L-glutamine	98	7.7

## 7. Supplementary material

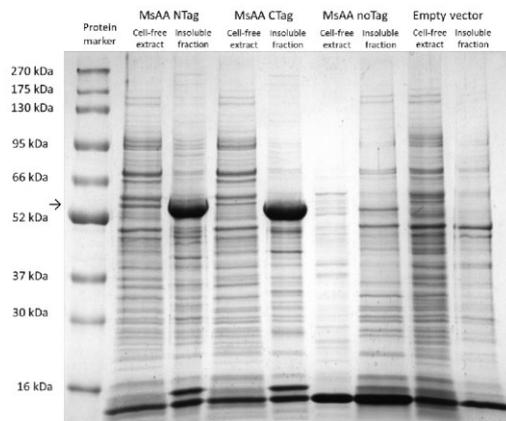


Figure S1: SDS-PAGE of expression of all MsAA tag-variants from *V. natriegens* Vmax™ at 37 °C.

Lane 1: protein marker (BlueEasy Prestained Protein Marker, Nippon Genetics); lane 2: cell-free extract with MsAA NTag; lane 3: insoluble fraction with MsAA NTag; lane 4: cell-free extract with MsAA CTag; lane 5: insoluble fraction with MsAA CTag; lane 6: cell-free extract with MsAA noTag; lane 7: insoluble fraction with MsAA noTag; lane 8: cell-free extract of empty vector control; lane 9: insoluble fraction of empty vector control

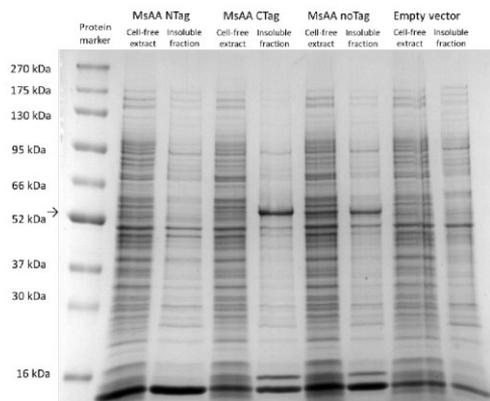


Figure S2: SDS-PAGE of expression of all MsAA tag-variants from *V. natriegens* Vmax™ at 20 °C.

Lane 1: protein marker (BlueEasy Prestained Protein Marker, Nippon Genetics); lane 2: cell-free extract with MsAA NTag; lane 3: insoluble fraction with MsAA NTag; lane 4: cell-free extract with MsAA CTag; lane 5: insoluble fraction with MsAA CTag; lane 6: cell-free extract with MsAA noTag; lane 7: insoluble fraction with MsAA noTag; lane 8: cell-free extract of empty vector control; lane 9: insoluble fraction of empty vector control

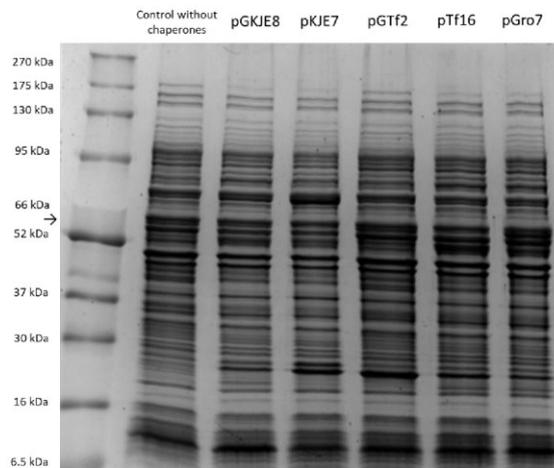


Figure S3: SDS-PAGE of soluble fractions from MsAA expression from *V. natriegens* Vmax™ with chaperone co-expression.

Lane 1: protein marker (BlueEasy Prestained Protein Marker, Nippon Genetics); lane 2: cell-free extract of empty vector control without chaperone co-expression; lane 3: cell-free extract of MsAA NTag expression without chaperone co-expression; lane 4: cell-free extract of MsAA NTag expression with GroEL/S and DnaK/J/GrpE chaperone co-expression from pGKJE8; lane 5: cell-free extract of MsAA NTag expression with DnaK/J/GrpE chaperone co-expression from pKJE7; lane 6: cell-free extract of MsAA NTag expression with GroEL/S and Trigger factor chaperone co-expression from pGTF2; lane 7: cell-free extract of MsAA NTag expression with Trigger factor chaperone co-expression from pTf16; lane 8: cell-free extract of MsAA NTag expression with GroEL/S chaperone co-expression from pGro7

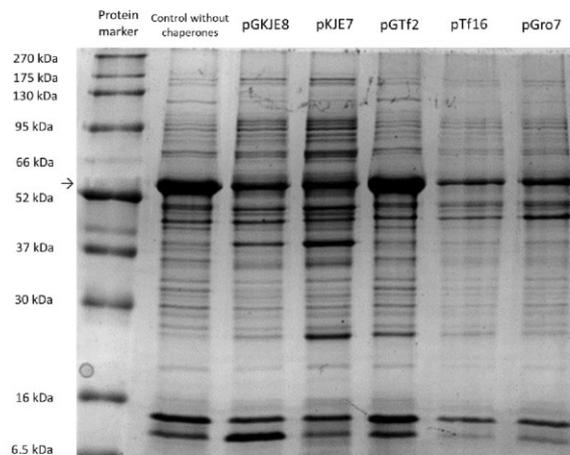


Figure S4: SDS-PAGE of insoluble fractions from MsAA from *V. natriegens* Vmax™ with chaperone co-expression.

Lane 1: protein marker (BlueEasy Prestained Protein Marker, Nippon Genetics); lane 2: insoluble fraction of empty vector control without chaperone co-expression; lane 3: insoluble fraction of MsAA NTag expression without chaperone co-expression; lane 4: insoluble fraction of MsAA NTag expression with GroEL/S and DnaK/J/GrpE chaperone co-expression from pGKJE8; lane 5: insoluble fraction of MsAA NTag expression with DnaK/J/GrpE chaperone co-expression from pKJE7; lane 6: insoluble fraction of MsAA NTag expression with GroEL/S and Trigger factor chaperone co-expression from pGTF2; lane 7: insoluble fraction of MsAA NTag expression with Trigger factor chaperone co-expression from pTf16; lane 8: insoluble fraction of MsAA NTag expression with GroEL/S chaperone co-expression from pGro7

## 7. Supplementary material

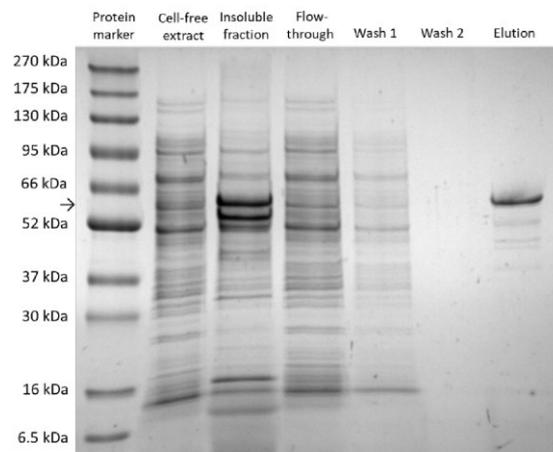


Figure S5: SDS-PAGE of MsAA NTag overexpression and purification from *V. natriegens* Vmax™ without chaperone co-expression.

Lane 1: Protein marker (BlueEasy Prestained Protein Marker, Nippon Genetics); lane 2: cell-free extract with MsAA NTag; lane 3: insoluble fraction with MsAA NTag; lane 4: Flow-through from purification; lane 5: first wash fraction; lane 6: second wash fraction; lane 7: Elution of MsAA NTag.

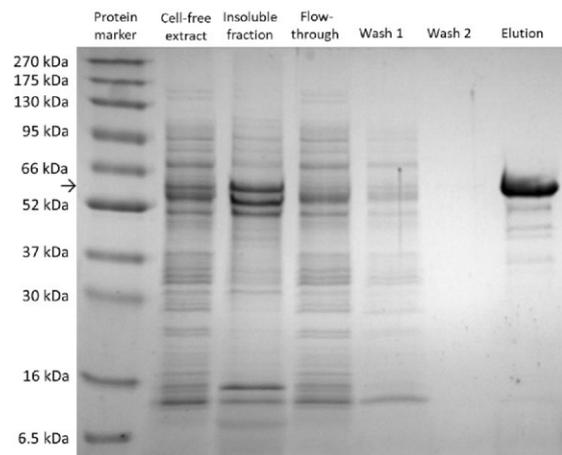


Figure S6: SDS-PAGE of MsAA NTag overexpression and purification from *V. natriegens* Vmax™ with GroEL/S co-expression.

Lane 1: Protein marker (BlueEasy Prestained Protein Marker, Nippon Genetics); lane 2: cell-free extract with MsAA NTag; lane 3: insoluble fraction with MsAA NTag; lane 4: Flow-through from purification; lane 5: first wash fraction; lane 6: second wash fraction; lane 7: Elution of MsAA NTag.

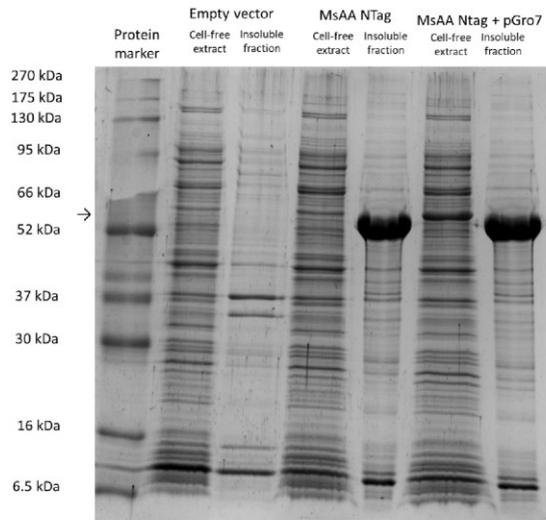


Figure S7: SDS-PAGE of IPTG-induced expression of MsAA NTag from *E. coli* BL21 (DE3) at 30 °C.

Lane 1: protein marker (BlueEasy Prestained Protein Marker, Nippon Genetics); lane 2: cell-free extract of empty vector control; lane 3: insoluble fraction of empty vector control; lane 4: cell-free extract with MsAA NTag; lane 5: insoluble fraction with MsAA NTag; lane 6: cell-free extract with MsAA NTag and GroEL/S co-expression from pGro7; lane 7: insoluble fraction with MsAA NTag and GroEL/S co-expression from pGro7.

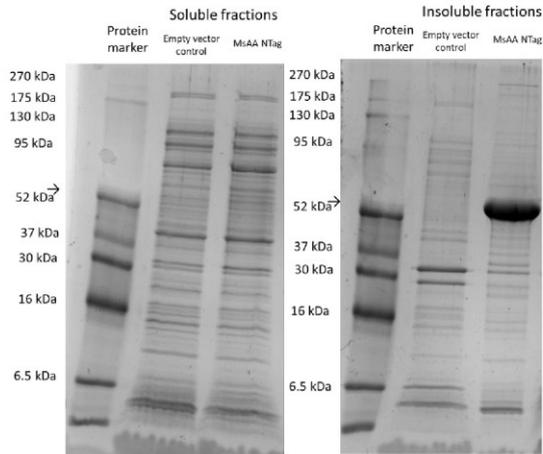


Figure S8: SDS-PAGE of IPTG-induced expression of MsAA NTag from *E. coli* Tuner™ (DE3) at 30 °C.

Lane 1: protein marker (BlueEasy Prestained Protein Marker, Nippon Genetics); lane 2: cell-free extract of empty vector control; lane 3: cell-free extract with MsAA NTag; lane 4: protein marker; lane 5: insoluble fraction of empty vector control; lane 6: insoluble fraction with MsAA NTag

## 7. Supplementary material

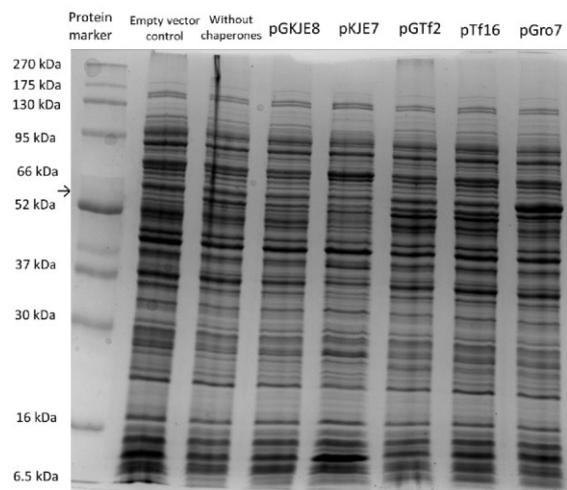


Figure S9: SDS-PAGE of soluble fractions from MsAA expression from *E. coli* BL21 (DE3) with chaperone co-expression

Lane 1: protein marker (BlueEasy Prestained Protein Marker, Nippon Genetics); lane 2: cell-free extract of empty vector control without chaperone co-expression; lane 3: cell-free extract of MsAA NTag expression without chaperone co-expression; lane 4: cell-free extract of MsAA NTag expression with GroEL/S and DnaK/J/GrpE chaperone co-expression from pGKJE8; lane 5: cell-free extract of MsAA NTag expression with DnaK/J/GrpE chaperone co-expression from pKJE7; lane 6: cell-free extract of MsAA NTag expression with GroEL/S and Trigger factor chaperone co-expression from pGTf2; lane 7: cell-free extract of MsAA NTag expression with Trigger factor chaperone co-expression from pTf16; lane 8: cell-free extract of MsAA NTag expression with GroEL/S chaperone co-expression from pGro7

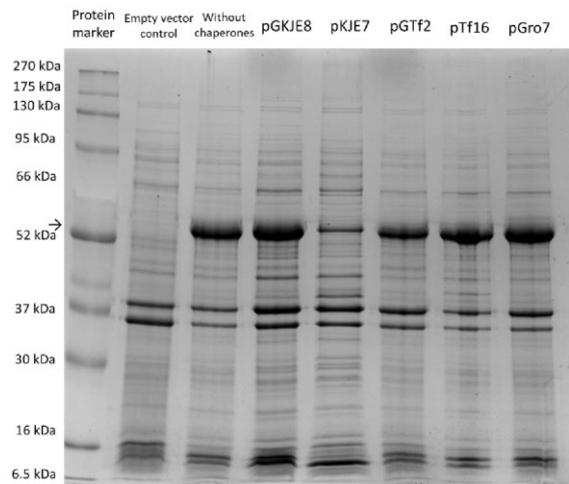


Figure S10: SDS-PAGE of insoluble fractions from MsAA expression from *E. coli* BL21 (DE3) with chaperone co-expression

Lane 1: protein marker (BlueEasy Prestained Protein Marker, Nippon Genetics); lane 2: insoluble fraction of empty vector control without chaperone co-expression; lane 3: insoluble fraction of MsAA NTag expression without chaperone co-expression; lane 4: insoluble fraction of MsAA NTag expression with GroEL/S and DnaK/J/GrpE chaperone co-expression from pGKJE8; lane 5: insoluble fraction of MsAA NTag expression with DnaK/J/GrpE chaperone co-expression from pKJE7; lane 6: insoluble fraction of MsAA NTag expression with GroEL/S and Trigger factor chaperone co-expression from pGTF2; lane 7: insoluble fraction of MsAA NTag expression with Trigger factor chaperone co-expression from pTf16; lane 8: insoluble fraction of MsAA NTag expression with GroEL/S chaperone co-expression from pGro7

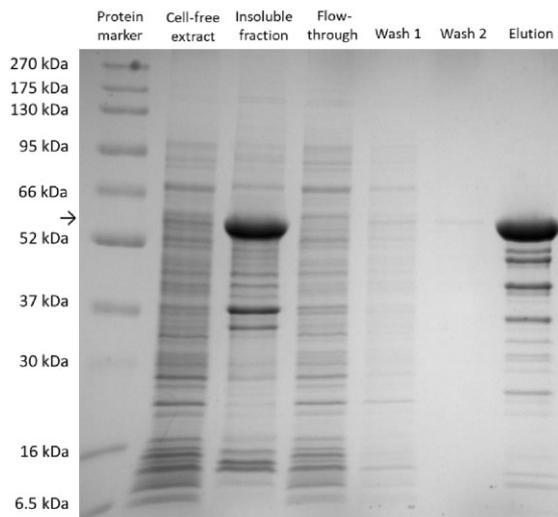


Figure S11: SDS-PAGE of MsAA NTag overexpression and purification from *E. coli* BL21 (DE3) without chaperone co-expression.

Lane 1: Protein marker (BlueEasy Prestained Protein Marker, Nippon Genetics); lane 2: cell-free extract with MsAA NTag; lane 3: insoluble fraction with MsAA NTag; lane 4: Flow-through from purification; lane 5: first wash fraction; lane 6: second wash fraction; lane 7: Elution of MsAA NTag.

## 7. Supplementary material

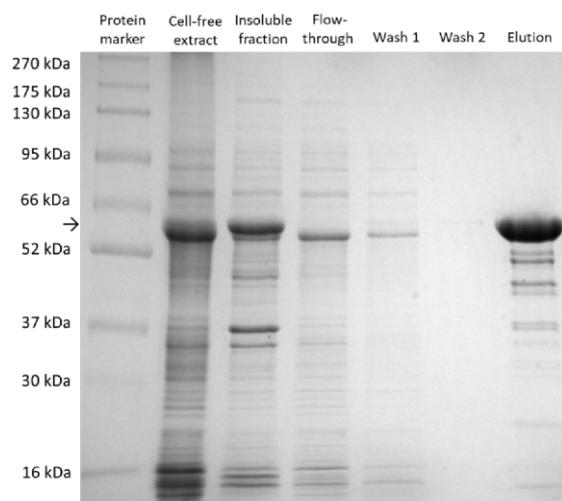


Figure S12: SDS-PAGE of MsAA NTag overexpression and purification from *E. coli* BL21 (DE3) with GroEL/S co-expression.

Lane 1: Protein marker (BlueEasy Prestained Protein Marker, Nippon Genetics); lane 2: cell-free extract with MsAA NTag; lane 3: insoluble fraction with MsAA NTag; lane 4: Flow-through from purification; lane 5: first wash fraction; lane 6: second wash fraction; lane 7: Elution of MsAA NTag.

### MALDI-TOF analysis:

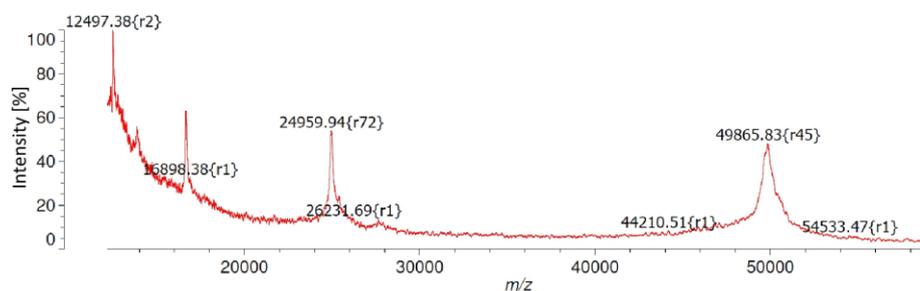


Figure S13: MALDI-TOF mass spectrum of purified MsAA NTag

The relative intensity is plotted against the  $m/z$ -values. The numbers above the peaks indicate the measured  $m/z$ -value.

Native PAGE:

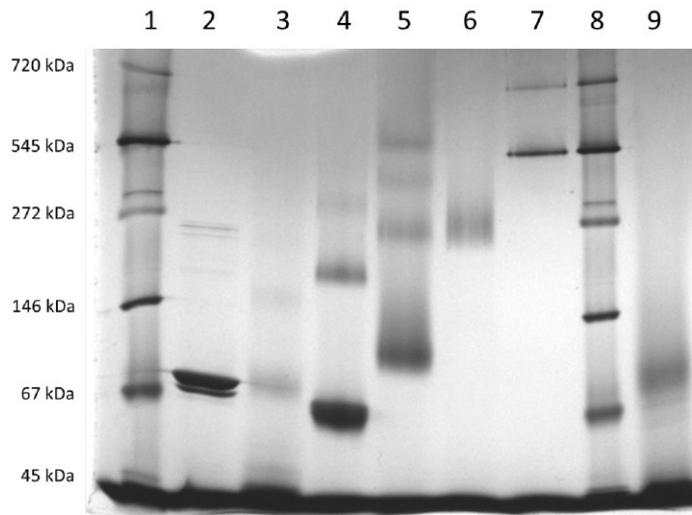


Figure S14: Native PAGE of MsAA and reference proteins.

lane 1: protein marker (SERVA Native Marker, Liquid Mix for BN/CN); lane 2: MsAA; lane 3: ovalbumin (44 kDa); lane 4: BSA (67 kDa); lane 5: conalbumin (75 kDa); lane 6: aldolase (158 kDa); lane 7: ferritin (440 kDa); lane 8: protein marker; lane 9: CsAga (Alpha-glutamine aminoacylase from *Corynebacterium striatum* Ax20; 96 kDa dimeric/48 kDa monomeric; heterologously expressed; unpublished data)

Isoelectric focusing:

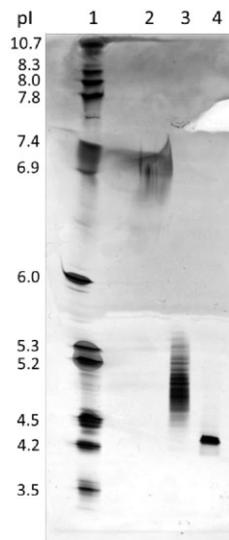


Figure S15: Isoelectric focusing for pI analysis of MsAA NTag.

Lane 1: protein marker (IEF Marker 3-10, Liquid Mix, Serva); lane 2 and 3: other protein samples; lane 4: MsAA

## 7. Supplementary material

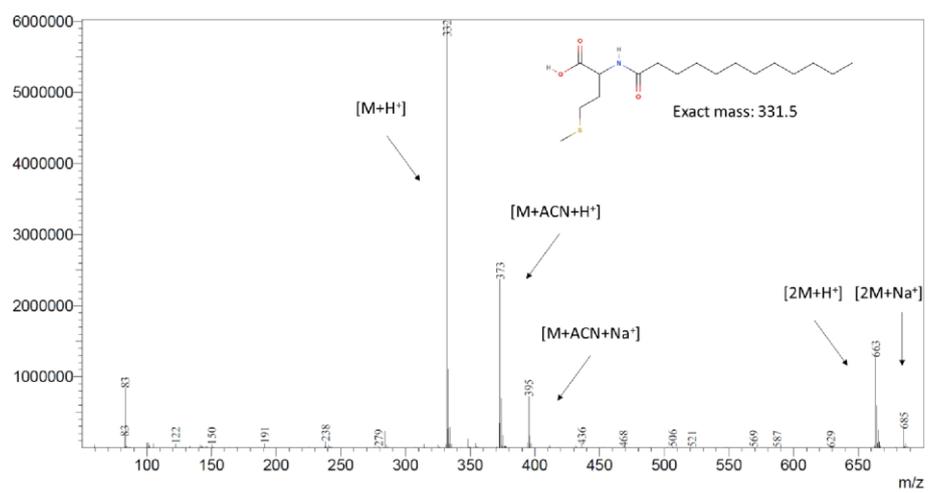


Figure S16: Mass-spectrum of N-lauroyl-L-methionine produced by MsAA

## Supplementary materials for chapter 4

Journal name:

**ChemBioChem**

Title:

**Biocatalytic potential of MsAA aminoacylase for synthesis of N-acyl-L-amino acids in aqueous media**

Author's names:

Gerrit Haeger<sup>a</sup>, Jessika Wirges<sup>a</sup>, Laureline Gennesseaux<sup>b</sup>, Yann Guiavarc'h<sup>b</sup>, Catherine Humeau<sup>b</sup>, Cédric Paris<sup>c</sup>, Isabelle Chevalot<sup>b</sup>, Karl-Erich Jaeger<sup>d,e</sup>, Johannes Bongaerts<sup>a</sup>, Petra Siegert<sup>a</sup>

Addresses:

<sup>a</sup>Institute of Nano- and Biotechnologies, Aachen University of Applied Sciences, 52428 Jülich, Germany

<sup>b</sup>Université de Lorraine, CNRS, Laboratoire Réactions et Génie des Procédés (LRGP), F-54000, Nancy, France

<sup>c</sup>Université de Lorraine, CNRS, Laboratoire d'Ingénierie des Biomolécules (LIBio), F-54000, Nancy, France

<sup>d</sup>Institute of Molecular Enzyme Technology, Heinrich Heine University Düsseldorf, 52425 Jülich, Germany

<sup>e</sup>Institute of Bio- and Geosciences IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

Author for correspondence:

Petra Siegert; Heinrich-Mussmann-str. 1, 52428, Jülich, Germany. E-mail address: siegert@fh-aachen.de (P. Siegert), Tel.: +49 241 6009 53124

## 7. Supplementary material

### MS/MS-spectra of synthesized acyl-methionines:

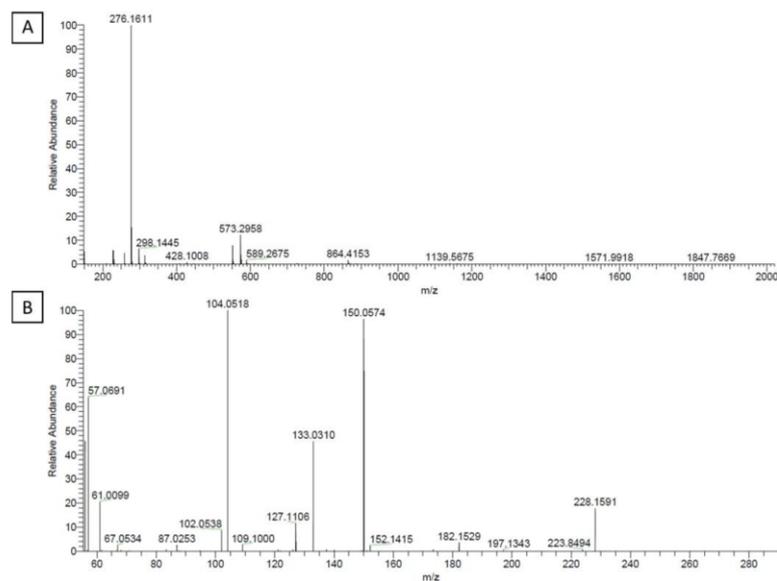


Figure S1: N-Capryloyl-L-methionine (N-octanoyl-L-methionine; molecular weight = 275.4 g/mol). (A) MS<sup>1</sup> and (B) MS<sup>2</sup> spectra.

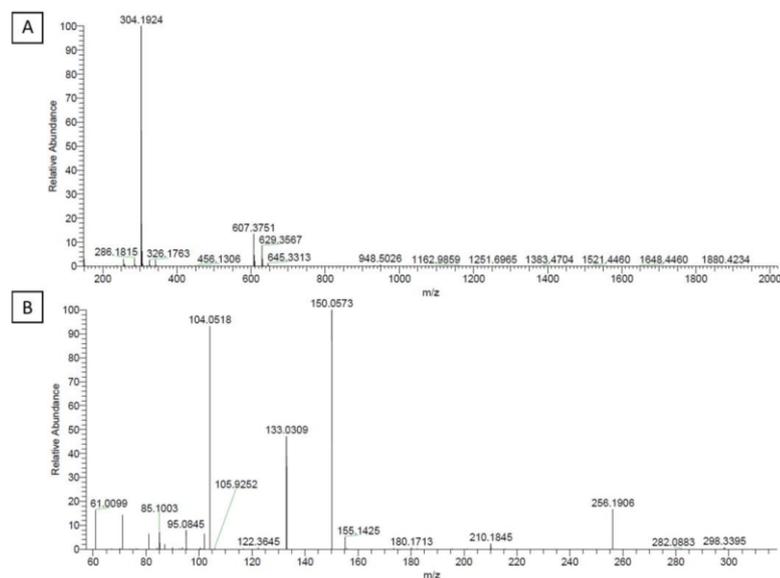


Figure S2: N-Decanoyl-L-methionine (molecular weight = 303.5 g/mol). (A) MS<sup>1</sup> and (B) MS<sup>2</sup> spectra.

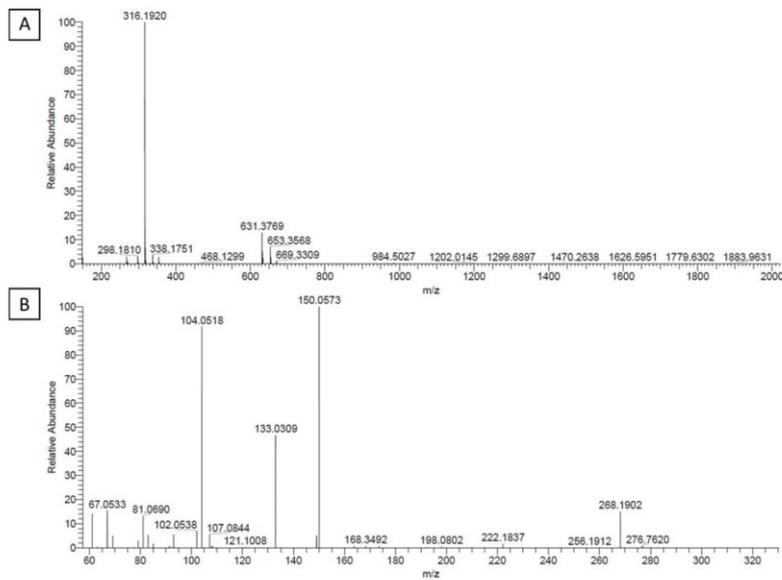


Figure S3: N-10-Undecenyl-L-methionine (molecular weight = 315.5 g/mol). (A) MS<sup>1</sup> and (B) MS<sup>2</sup> spectra..

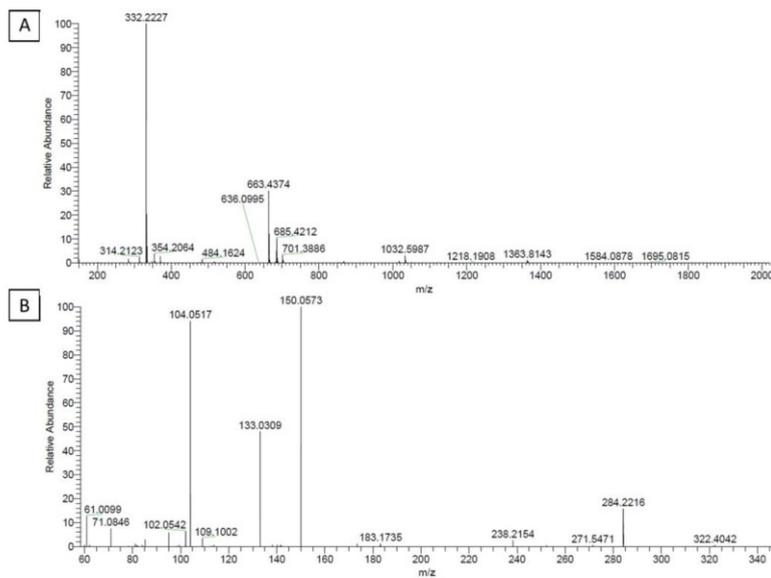


Figure S4: N-lauroyl-L-methionine (molecular weight = 331.5 g/mol). (A) MS<sup>1</sup> and (B) MS<sup>2</sup> spectra.

## 7. Supplementary material

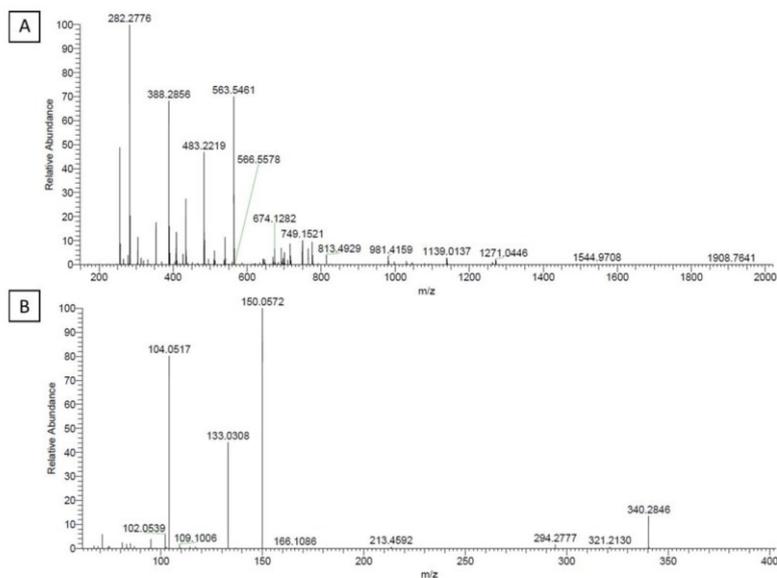


Figure S5: N-Palmitoyl-L-methionine (molecular weight = 387.6 g/mol). (A) MS<sup>1</sup> and (B) MS<sup>2</sup> spectra.

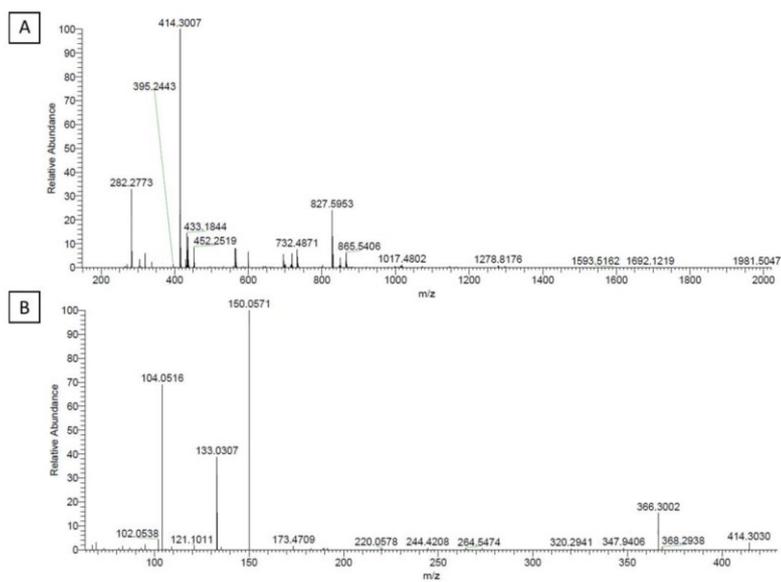


Figure S6: N-Oleoyl-L-methionine (molecular weight = 413.7 g/mol). (A) MS<sup>1</sup> and (B) MS<sup>2</sup> spectra.

MS/MS-spectra of synthesized lauroyl-amino acids:

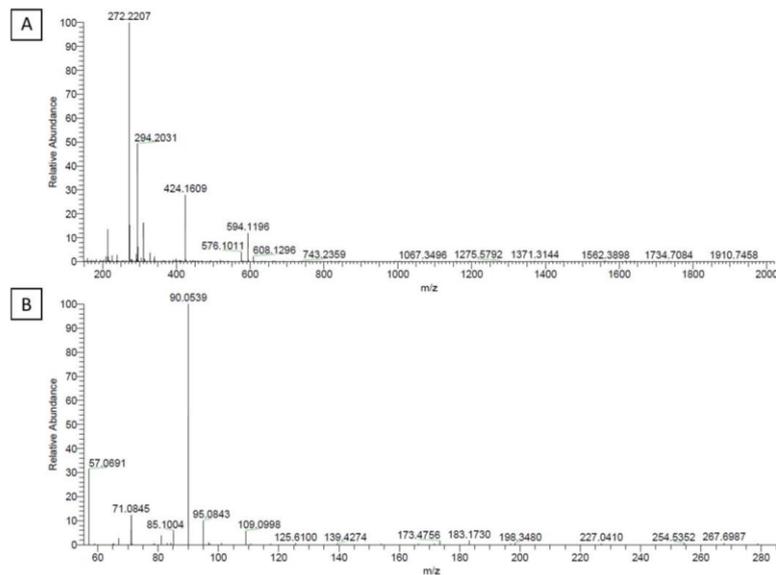


Figure S7: N-lauroyl-L-alanine (molecular weight = 271.4 g/mol). (A) MS<sup>1</sup> and (B) MS<sup>2</sup> spectra.

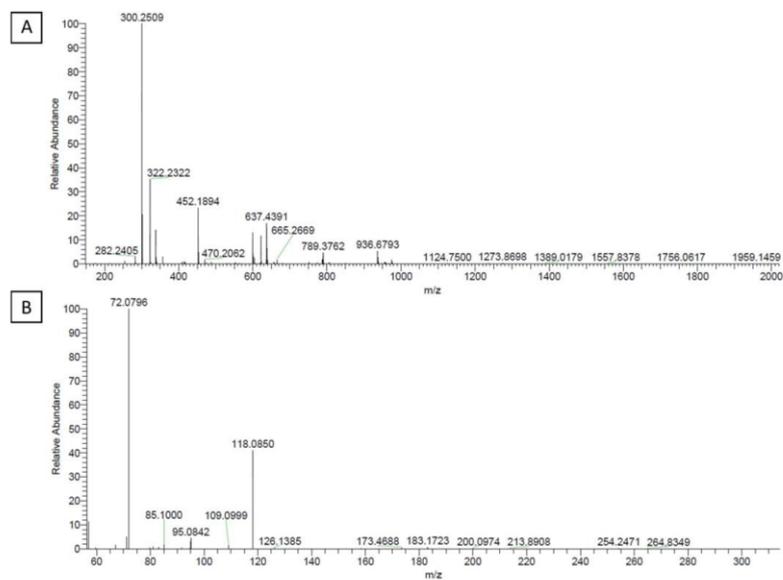


Figure S8: N-lauroyl-L-valine (molecular weight = 299.4 g/mol). (A) MS<sup>1</sup> and (B) MS<sup>2</sup> spectra.

## 7. Supplementary material

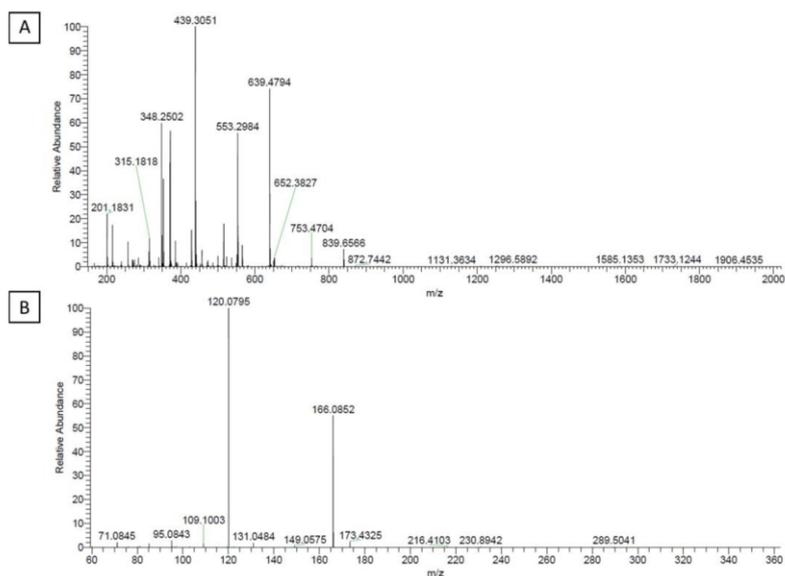


Figure S9: N-lauroyl-L-phenylalanine (molecular weight = 347.5 g/mol). (A) MS<sup>1</sup> and (B) MS<sup>2</sup> spectra.

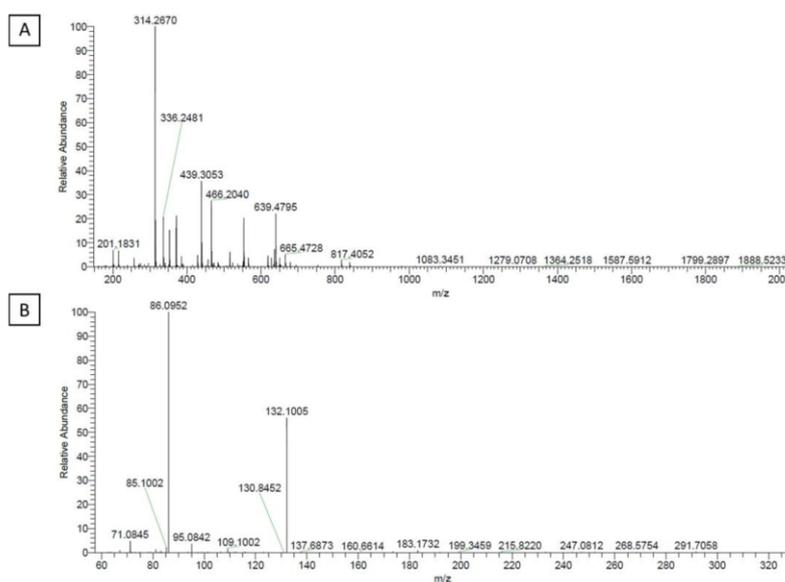


Figure S10: N-lauroyl-L-leucine (molecular weight = 313.5 g/mol). (A) MS<sup>1</sup> and (B) MS<sup>2</sup> spectra.

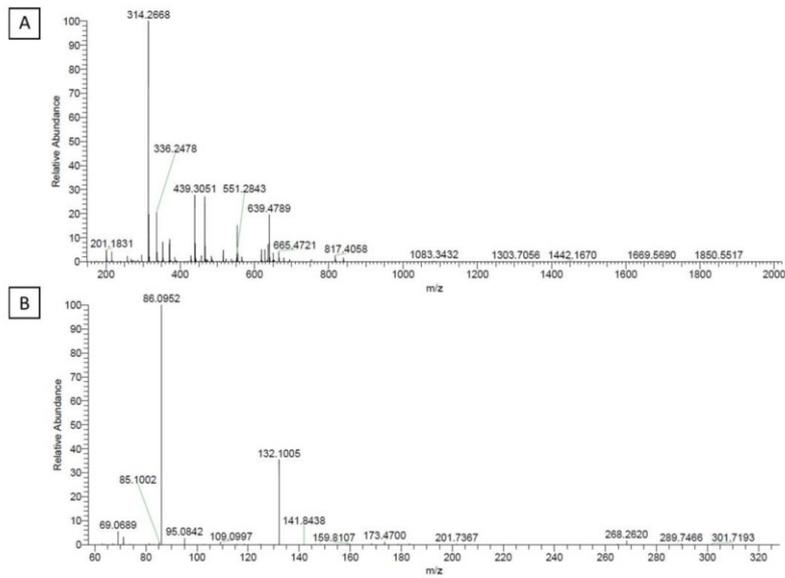


Figure S11: N-lauroyl-L-isoleucine (molecular weight = 313.5 g/mol). (A) MS<sup>1</sup> and (B) MS<sup>2</sup> spectra.

**Supplementary materials for chapter 5**

Journal name:

**FEBS Open Bio**

Title:

**Novel aminoacylases from *Streptomyces griseus* DSM 40236 and their recombinant production in *Streptomyces lividans***

Author's names:

Gerrit Haeger<sup>a</sup>, Johanna Probst<sup>a</sup>, Karl-Erich Jaeger<sup>b,c</sup>, Johannes Bongaerts<sup>a</sup>, Petra Siegert<sup>a\*</sup>

Addresses:

<sup>a</sup>Institute of Nano- and Biotechnologies, Aachen University of Applied Sciences, 52428 Jülich, Germany

<sup>b</sup>Institute of Molecular Enzyme Technology, Heinrich Heine University Düsseldorf, 52425 Jülich, Germany

<sup>c</sup>Institute of Bio- and Geosciences IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

Author for correspondence:

Petra Siegert; Heinrich-Mussmann-Str. 1, 52428, Jülich, Germany. E-mail address: siegert@fh-aachen.de (P. Siegert), Tel.: +49 241 6009 53124

Supplementary material to “Novel aminoacylases from *Streptomyces griseus* DSM 40236 and recombinant production in *Streptomyces lividans*”:

DNA and protein sequences for SgAA and SgELA, native PAGE, MS spectrum for lauroyl-methionine

SgAA NTag DNA sequence from *S. griseus* DSM 40236:

```
ATGTGGTCCCACCCGAGTTCGAGAAGTCCGGCATGAGCGAGAGCAGCACGGGCAGGGCCGGCGCCGG
CAGGGCCGAGCAGGAGGTCGTGACCTCTGTGTCGTGACCTGATCCGGATCGACACCAGCAACTACGGCG
ACCACTCGGGCCCCGGCGAGCGGCTCGCGGCCGAGTACGTGCGGAGAAAGCTCGCGGAGGTCGGCCTG
GAGCCCGCGATCTTCGAGTCGCACAAGGGACGCGCCTCCACCGTCGCCCGGATCGAGGGCGAGACCC
CTCCCGCCCGCGCTGCTGATCCACGGACACACCGACGTGTCGCCGGCCAAACGGCGGACTGGACGC
ACGACCCGTCTCCGGGAGATCGCGGACGGCTGCGTGTGGGGCCGGGGCGCGGTTGACATGAAGGAC
ATGGACGCGATGACCTCGCGGTGTCGGGAGCGGATGCGCAGCGGGCCGAAGCCCCCGCGGACAT
CGTGCTCGCCTTCTCGCGGACGAGGAGGGCGGGCACGTACGGGGCCCGCTATCTCGTGGACAACC
ACCCGGGCTCTTCGAGGGCGTCACCGAGGCGATCAGCGAGGTCGGCGGCTTCTCCTTACCGTCAAC
GAGAACCTCGCGGTGTATCTGGTGGAGACCGCCAGAAGGGCATGCACTGGATGAAGCTGACCGTGG
CGCACCCCGGACACGGCTCGATGATCCACAAGGACAACGCCATCACGGAGCTGTCGAGGCGGTCCG
GGCGGTGGGCCGGACAAGTCCCGGTGCGGGTCAACAAGACGCTGCGGCACCTTCTGGACGAGCTC
TCCGACGCGCTGGGCACCGAGCTGGACCCGGAGAACATGGACGAGACGCTCGCCAAGCTCGCGGCG
CGCCAAGCTCATCGGCGCCTCCCTCCAGAACACCGCCAACCCACGCAGCTCGGCGCGGCTACAAGG
TCAACGTATCCCGGGCCAGGCGACCGCCACGTGGACGGCCGGTACCTCCCGGGTACGAGGAGGAG
TTCTCGCGACCTGGACCGGATCCTCGGGCCCAAGTCCGGCGCGAGGACGTGCACGCGGACAAGGC
CCTGGAGACCAGTTCGACGGCGCGCTGGTCGACGCCATGCAGACCGCGCTGGTCGCGGAGACCCCA
TCGCCCGTCCCGTGCCTTACATGCTCTCGGCCGGCACCGACGCCAAGTCTTCGACGACCTGGGCAT
CGGGGCTTCGGCTTCGCCCGCTGAAGCTGCCCGGAGCTGGACTTCGCGGCGATGTTCCACGGCGT
CGACGAGCGCTCCCGTTCGACGGGCTGCAGTTCCGGCTGCGGGTGCACCGGTTTCATCGACCCT
CCTGA
```

Codon-optimized sequence of SgAA for *E. coli* expression with N- and C-terminal strep-tag shown in bold letters (as ordered from GeneArt, Thermo Fisher)

```
ATGTGGTCACATCCGAGTTTGAAAAAGCGGTATGAGCGAAAGCAGCACCGGTCGTGCCGGTGCAGG
TCGTGCAGAACAAAGAGTTGTTGATCTGTGTCGTGATCTGATTCGTATGATACAGCAATATATGGT
ATCATAGCGGTCCGGTGAAAGTCTGGCAGCAGAAATATGTTGCAGAAAAACTGGCAGAAAGTGGTCTG
GAACCGGTATTTTGAAGCCATAAAGTTCGTGCCAGCACCGTTGCACGTATGAAGGTGAAGATCC
GAGCCCTCCGGCACTGCTGATTCATGGTCATACCGATGTTGTTCCGGCAAATGCAGCAGATTTGGACCC
ATGATCCGTTTAGCGGTGAAATTCAGATGGTTGTGTTGGGGTCGTGGTGCAGTTGATATGAAAGAT
ATGGATGCAATGACCTGGCAGTTGTTGTCGTAACGTATGCGTAGCGGTCGTAAACCGCCTCGTGATAT
TGTTCTGGCATTCTGGCAGATGAAGAGGCAGGCGGTACATATGGTGCACGTTATCTGGTTGATAATC
ATCCGGGTCTGTTGAAGGTGTTACCGAAGCAATTAGCGAAGTTGGTGGTTTTAGCTTTACCGTGAAT
GAAAATCTGCGTCTGTATCTGGTGGAAACCGCACAGAAAGGTATGCATTGGATGAAACTGACCGTTGA
TGGCACCCGAGGTGATGGCAGCATGATTCATAAAGATAACCGGATTACCGAAGTGCAGCAAGCGGTGG
GCCCGCTGGGCCGCCATAAATTTCCGGTGCAGCTGACAAAACCCCTGCGCCATTTCTGGATGAACTG
AGCGATGCGCTGGGCACCGAACTGGATCCGGAAAACATGGATGAAACCTGGCGAAAAGTGGGCGGCAT
TGCGAAAAGTATGGCGCGAGCCTGCAGAACACCGCGAACCCGACCCAGCTGGGCGCGGGCTATAAAG
TGAACGTGATTCGGGGCCAGGCGACCGCGCATGTGGATGGCCGCTATCTGCCGGGCTATGAAGAAGAA
TTTCTGGCGGATCTGGATCGCATTTCTGGGCCGAAAGTGCAGCGCGAAGATGTGCATGCGGATAAAGC
GCTGGAAACCACCTTTGATGGCGCGCTGGTGGATGCGATGCAGACCGCGCTGGTGGCGGAAGATCCGA
TTGGCGCGCGGTGCCGTATATGCTGAGCGCGGGCACCGATGCGAAAAGCTTTGATGATCTGGGCATT
```

## 7. Supplementary material

CGCGGCTTTGGCTTTGCGCCGCTGAAACTGCCGCCGAACTGGATTTTGGCGGCATGTTTCATGGCGT  
GGATGAACCGGTGCCGGTGGATGGCCTGCAGTTTGGCGTGCAGTGGATCGCTTTATTGATCATA  
**GC****TCCGGATGGAGTCA****TCC****CAATTCGAAAAATGA**

SgAA protein sequence (Accession no. WP\_003970135.1) with N-terminal Strep-tag and linker  
underlined:

MWSHPQFEKSGMSESSTGRAGAGRAEQEVVDLCRDRLRIDTSNYGDHSGPGERLAAEYVAEKLAEVGLEPRIFES  
HKGRASTVARI EGEDFSRPALLI HGHTDVVPANAADWTHDPFSGELADGCVWGRGAVDMKMDAMT LAVVRERMR  
SGRKPPRDIVLAF LADEEAGGTYGARYLVDNHPGLFEGVTEAISEVGGFSFTV NENLRLYL VETAQKGMHWMKLT  
VDGTAGHGSMIHKDNAITELSEAVGRLGRHKFPVVRTKTLRHFLDELSDALGTELDPENMDET LAKLGGIAKLI G  
ASLQNTANPTQLGAGYKVNVI PGQATAHVDGRYLPGYEEEF LADLDRI LGPNVRREDVHADKALETTFD GALVDA  
MQTALVAEDPIARAVPYMLSAGTDAKSFDDLGRGFGFAPLKLPELDFAGMFHGVDERV PVDGLQFGVVRVLDRF  
IDHS

SgELA NTag DNA sequence from *S. griseus* DSM 40236:

ATGTGGTCCCACCCGAGTTCGAGAAGTCCGGCATGAGCCAGAGCACCGCCCCCAGAGCGCCCCGA  
ACACCGCACCGTGTGTGCGCGGTGGAGACGTCCACAGCCCCGCCGACCGTTTCGCCACCGCGATGG  
TCGTGCAACCGGGCATGTGCGCTGGGTCCGGTCCGAGGGGGCCGCCGACGCCTTCCGCGAGCGCGGTG  
GACGAGGTGGTGCACCTCGAAGGCGCCCTGGTCAACCCGGCGTTCACGGACGCCCATGTGCACACCAC  
CGCCACCGGCTGGCGCTGACGGGGCTCGACCTCTCCGGCGCCGCACCCTGTCCGAGGCCCTCGGCC  
TCGTCCGTGCGTTCGCGAAGGGGCGCTCCGCCGGGACGTTCTGCTCGGACACGGCTGGGACGCCGCC  
CGCTGGCCCCGAGCGGCCACCCCTCGCGCGCCGAGTTCGACGAGGCGGCCGGCGGCCGGCCCTGTA  
CCTGCCCGGATCGACGTGCACTCCGCGGTTCGTCACGACGGCCCTGCTCGACCTCGTCCCCGGCGTCA  
CCGCGATGACCGGCTACCACCCCGACGCTCCGCTCACCGGCGACGCCACCACCGGTTACGGCCCGCC  
GCCACAGCGCGTCCCGGCCGCCAGCGGGCGCCGCGCAGCGCGCCGCCCTCGACCACCGCCCTC  
CCTCGGCATCGGCAGCGTGCACGAGTGCGGGGGGCGGAGATCTCCGACGAGGAGGACTTACCTCGC  
TGCTCGCGCTCGCCGCGACCGGCCGGGGCCGCGCTCCTCGGCCTCTGGGCCGAGGAGATCGCGGAC  
GAGAAGGGCGCCCGGCATCCGCGAACTCGGCGGATCGGGCGGGCCGGCGACCTGTTCTGTCGACGG  
CTCGTTGGGCTCGCACACCGCTGCCTGCACCGGCCCTACGCGGACGACCCGCCACCGGCCCGCC  
ACCTGGACGCCCGCGGATCGCCGCCACGTCACCGCTGCACCGAGGCGGCCCTCCAGGCGGGCTTC  
CAGGCCATCGGCGACCGCGGTTCACCGCGTGGTGGACGGATCCGGCCGCCCGCGGAGGTGCTCGG  
CCTCGACCGCGTCCGGGCCGCCCGGCACCGCGTCAACACGCGGAGATGCTCACCCCGAGACGATCG  
CCGCTTCGCGCAACTGGGCCCTCACCGCTCCGTCAGCCCGCTTCGACCGCGCTGGGGCGGGCC  
GAGGGGATGTACGCCGAGCGCTCGGCGCGGAGCGGGCCGCCACGCTCAACCCCTACGCGCGCTGCT  
GCGGGCCGGCTGCCCTTGGCTTCGGCTCCGACAGCCCGGTACCCCGCTCGACCCCTGGGGCACGG  
TCCGCGCCCGCCACCACCGCACCCCGGAGCACCGCTTCGGTCCGCGCCGGGTTCACCGCGCAC  
ACCCGCGCGGCTGGCGGGCGTCCGGCCGACGACGCGGGCCCTCCTGGTGCCCGGGCCCCGGCCGA  
CTACGCCCTGGCGCACCGCGAACTCCTGGTCCAGGCCCGGACGACCGGGTCCGCCCGTGGTCCA  
CCGACCCCGGTCCGGCACGCCCGGCTGCCGGACCTCACCCCGGGGCCGACCTCCCCGTCTGCCTG  
CGGACCGTGGTCTCGGACAAACGGTCTACGTGCGACCGAACGAGTGA

Codon-optimized sequence of SgELA for *E. coli* expression with N- and C-terminal strep-tag shown  
in bold letters (as ordered from GeneArt, Thermo Fisher)

**ATGTGGTCA****CATCCG****CAGTTT****GAAAAA****AGCGGT**ATGAGCCAGAGCACCGCACCGCAGAGCGCACCGGA  
ACATCGTACCGTTCGCTGCGTGGTGGTGTATGTTTCATAGTCCGGCAGATCCGTTTGCAACCGCAATGG  
TTGTTGAACGTGGTATGTTGCATGGGTGGTAGCGAAGGTGCAGCAGATGCATTTGCAAGCGGTGTT  
GATGAAGTTGTTGATCTGGAAGTGCACCTGGTTACACCGGATTTACCGATGCACATGTTTCATACCAC  
CGCAACCGGTCTGGCACTGACAGGTCTGGATCTGAGCGGTGCACGTACCCTGAGCGAAGCACTGGGTT  
TAGTTCGTGATTTGCCAAAGTTCGTAGTCCCGGTGATGTGCTGTTAGGTTCATGGTTGGGATGCAGCA  
CGTTGGCCTGAACGTCGTATCCGAGCCGTGCAGAACTGGATGAAGCAGCCGGTGGTTCGTCACCTGTA  
TCTGCCTCGTATTGATGTGCATAGCGCAGTTGTGACCACCGCACTGCTGGATCTGGTTCGGGTGTTA  
CCGCAATGACCGGTTATCATCTGATGCACCGTACCGGTGATGCCATCATGCAGTTTCGTGCCGCA

GCACATAGCGCACTGCCTGCAGCACAGCGTGCAGCTGCTCAGCGTGCCGCACTGGATCATGCAGCAAG  
 CTTAGGTATTGGTAGCGTTCATGAATGTGGTGGTCCGAAATTAGTGATGAAGAAGATTTTACCAGCC  
 TGCTGGCACTGGCAGCAGATCGTCCGGGTCCGCGTGTTTTAGGTCTGTGGGCTGAAGAAATGCAGAT  
 GAAAAAGGTGCACGTCGTATTCGTGAACTGGGTGCAATTGGTGCAGCGGGTGACCTGTTTGTGATGG  
 TAGCCTGGGTAGCCATACCGCATGTCTGCATCGTCCGTATGCAGATGATCCGCATACCGGCACCGCAC  
 ATCTGGATGCAGCCCGTATTGCAGCACATGTTACCGCTGTACCGAAGCAGGTCTGCAAGCAGGTTTT  
 CATGCCATTGGTGTATGCCGAGTTACCGCAGTGGTGGATGGTATTTCGTGCAGCAGCCGAAGTTCTGGG  
 CTAGATCGTGTTCGCGCAGCACGTATCGTGTGAAATGCAGAAATGCTGACACCGGAAACCATTG  
 CAGCATTTCGCCAACTGGGTCTGACCGCAAGCGTTCAGCCTGCATTTGATGCAGCATGGGGTGGTCC  
 GAAGGTATGTATGCAGAACGTCTGGGTGCCGAACGTGCAGCAACCCGTAATCCGTATGCCGCACTGTT  
 ACGTGCCGGTGTTCGCGTGGCATTTGGTAGTGATAGTCCGGTGACACCGCTGGATCCGTGGGGCACCG  
 TACGTGCAGCGGCACATCATCGTACCCCTGAACATCGTGTAGCGTTCGTGCAGGTTTTACCGCACAT  
 ACCCGTGGTGGTTGGCGTGCAGTTGGTTCGTGATGATGCCGGTCTGCTGGTTCCTGGTGCACCGGCAGA  
 TTATGCAGTTTGGCGTACCGCAGAACTGCTGGTGCAGGCACCGGATGATCGTGTGCACGTTGGAGCA  
 CCGATCCGCGTAGCGGTACACCGGGTCTGCCGGATCTGACCCCTGGTGCCGATCTGCCTGTTTGTCTG  
 CGTACCGTGTTCGGGTGACACCGTTTTATGTTTCGTCGAATGAA**AGCGGTTGGAGCCATCCTCAGTT**  
**CGAGAAATGA**

SgELA protein sequence (Accession no. WP\_069631407.1) with N-terminal Strep-tag and linker

underlined:

MWSHPQFEKSGMSQSTAPQSAPEHRTVLLRGGDVHSPADPFATAMVVERGHVAVWGSEGAADAFASGVD  
 EVVDLEGALVTPAFTDAHVHTTATGLALTGLDLSGARTLSEALGLVRAFAKGRSAGDVLLGHGWDAAR  
 WPERRHPSRAELDEAAGGRALYLPRIDVHSAVVTTALLDLVPGVTAMTGYHPDAPLTGDAHHAHVAAA  
 HSALPAAQRAAAQRAALDHAASLGIGSVHECGGPEISDEEDFTSLLALAADRPGPRVLGLWAEIIDE  
 KGARRIRELGAIGAAGDLFVDGSLGSHTACLHRPYADDPHTGTAHLDAARIAAHVTACTEAGLQAGFH  
 AIGDAAVTAVVDGIRAAA<sup>EV</sup>LVGLDRVRAARHRVEHAEMLPETIAAF<sup>EL</sup>GLTASVQPAFDAAWGGPE  
 GMYAERLGAERAATLNPYAALLRAGVPLAFGSDSPVTPLDPWGTVRAAAHHRTP<sup>EH</sup>RVSVRAGFTAHT  
 RGGWRAVGRDDAGLLVPGAPADYAVWR<sup>TA</sup>ELLVQAPDDRVARWSTDP<sup>RS</sup>GTPLGLPDLTPGADLPVCLR  
 TVVLGQTVYVRPNE

## 7. Supplementary material

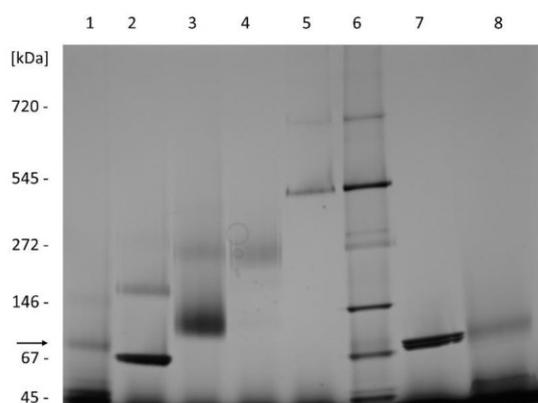


Figure S1: Native PAGE of SgAA and reference proteins.

lane 1: ovalbumin (44 kDa); lane 2: BSA (67 kDa); lane 3: conalbumin (75 kDa); lane 4: aldolase (158 kDa); lane 5: ferritin (440 kDa); lane 6: protein marker (SERVA Native Marker, Liquid Mix for BN/CN); lane 7: SgAA; lane 8: CsAga ( $\alpha$ -glutamine aminoacylase from *Corynebacterium striatum* Ax20; 96 kDa dimeric/48 kDa monomeric; heterologously expressed; unpublished data)

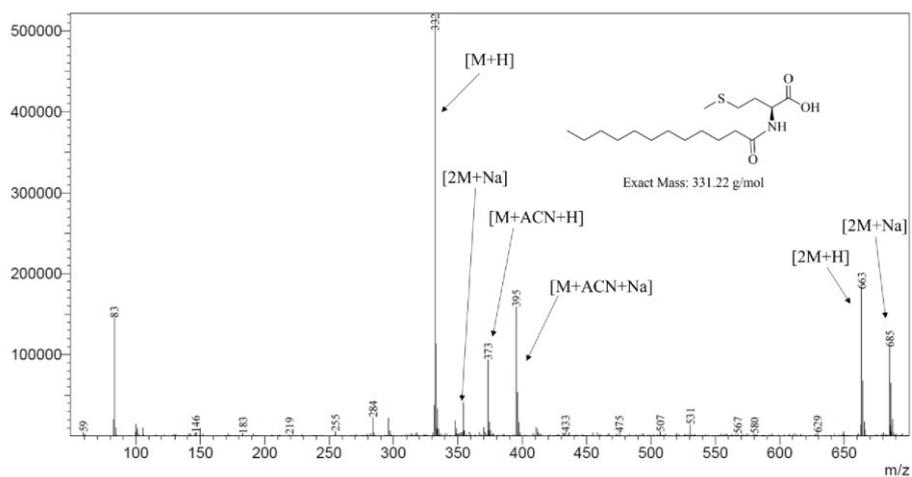


Figure S2: Mass-spectrum of N-lauroyl-L-methionine produced by SgAA.