Markus Plank BSc

Non-natural substrates for carboxylic acid reductase

MASTER'S THESIS

to achieve the university degree of

Master of Science

Master's degree programme: Biochemistry and Molecular Biomedical Sciences

submitted to

Graz University of Technology

Supervisor

O.Univ.-Prof. Dr.phil. Kurt Faber

Institute of Chemistry, Organic and Bioorganic Chemistry
University of Graz

Mag. Dr.rer.nat. Christoph Winkler
Institute of Chemistry, Organic and Bioorganic Chemistry, University of Graz

Graz, january 2016

AFFIDAVIT

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

8.1.2016
Date


Signature

„Wenn du nicht irrst, kommst du nicht zu Verstand.
Willst du entstehn, entsteh‘ auf eigne Hand!“
Mephistopheles
Johann Wolfgang von Goethe – Faust. Der Tragödie zweiter Teil

Acknowledgments

First and foremost I want to thank my parents, Renate and Leopold, for their support and for giving me the possibility to study. I also want to thank my brothers, Clemens and Bernhard, and Viktoria, the sister I never had.

I am grateful to my grandparents, Maria and Johann, and Christine, as well as to my uncle Hansi, for friendship and inspiration.

I want to express my gratitude to all my friends, wherever you are, for the impact you had on my life and for sharing a shorter or longer part of my path.

For giving me the opportunity to work in this group for the first time a few years ago and then again for my thesis, I want to thank Kurt Faber.

Additionally, I want to thank Wolfgang Kroutil and the rest of the Elk-Crew for providing a helpful and friendly working atmosphere, especially Horst for his help and patience throughout my work.

Special thanks to Christoph for his guidance and seemingly never-ending knowledge, but also for his positive spirit in the lab, as well as in the coffee room.

Abstract

Aldehydes are an important group of compounds with various applications in chemical industry. This makes chemoselective aldehyde synthesis a highly required task, however, the high reactivity of the aldehyde products often leads to undesired side reactions. Additionally, chemical synthesis usually relies on the use of toxic or hazardous compounds which require careful waste management, making these processes ecologically and economically disadvantageous.

Instead, biocatalysis is a useful alternative, as it is usually chemoselective and performed at ambient conditions. Carboxylic acid reductases for the reduction of carboxylic acids to their respective aldehydes have been identified from various eukaryotic and prokaryotic species. Most of the different classes of these enzymes rely on the use of ATP for the activation of the unreactive carboxylic acid functional group, which is expensive if ATP has to be applied in stoichiometric amounts on an industrial scale.

The *in vivo* catalytic cycle of the enzyme involves activation of the substrate, attachment to the enzyme, and reduction to the product. In this work, for an ATP-independent process, several non-natural, chemically activated substrates were used, so that they would get directly attached and reduced, circumventing the ATP-dependent activation step. Three different enzymes were chosen for recombinant expression in *E. coli* and investigation for activity on non-natural substrates: CAR from *Nocardia iowensis*, AAR from *Saccharomyces cerevisiae*, and the two-enzyme system of GriC and GriD from *Streptomyces griseus*. GriC and GriD, as well as CAR, could be expressed, however, none of these enzymes showed activity towards any of the non-natural substrates. Activity towards several carboxylic acids was shown for purified CAR, which was used for reaction engineering in regard to optimum reaction conditions. Additionally, an NADPH recycling system was established for this reaction.

Zusammenfassung

Aldehyde sind eine wichtige Stoffgruppe mit mannigfaltiger Anwendung in der chemischen Industrie. Dadurch besteht großer Bedarf an chemoselektiver Aldehydsynthese, die hohe Reaktionsfreudigkeit der Aldehydprodukte führt jedoch häufig zu unerwünschten Nebenreaktionen. Zusätzlich ist chemische Synthese üblicherweise auf toxische oder gefährliche Stoffe angewiesen, welche eine sorgfältige Entsorgung der Abfallprodukte erfordern, was diese Prozesse ökologisch und ökonomisch ungünstig macht.

Stattdessen ist Biokatalyse eine nützliche Alternative, da sie üblicherweise chemoselektiv ist und bei Umgebungsbedingungen durchgeführt wird. Carbonsäurereduktasen zur Reduktion von Carbonsäuren zu den entsprechenden Aldehyden wurden in verschiedenen eukaryotischen und prokaryotischen Spezies identifiziert. Die meisten der unterschiedlichen Enzymklassen sind auf ATP zur Aktivierung der unreaktiven funktionellen Gruppe der Carbonsäure angewiesen, wobei die stöchiometrische Verwendung von ATP in industriellen Größenordnungen teuer ist.

Der katalytische Zyklus des Enzyms beinhaltet *in vivo* die Aktivierung des Substrates, Bindung an das Enzym und Reduktion zum Produkt. Für einen ATP-unabhängigen Prozess wurden in dieser Arbeit verschiedene nicht-natürliche, chemisch aktivierte Substrate verwendet, die direkt gebunden und reduziert werden würden um den ATP-abhängigen Schritt zu umgehen. Drei verschiedene Enzyme wurden zur rekombinanten Expression in *E. coli* und Untersuchung auf nicht-natürliche Aktivität ausgewählt. CAR aus *Nocardia iowensis*, AAR aus *Saccharomyces cerevisiae* und das Zwei-Enzym-System GriC und GriD aus *Streptomyces griseus*. GriC und GriD, sowie CAR, konnten exprimiert werden, jedoch zeigte keines der Enzyme Aktivität mit einem nicht-natürlichen Substrat. Aufgereinigte CAR zeigte Aktivität mit einigen Carbonsäuren und wurde zur Optimierung der Reaktionsbedingungen eingesetzt. Zusätzlich konnte ein System zur NADPH-Wiederaufbereitung für diese Reaktion etabliert werden.

List of abbreviations

AAR	α-Aminoadipate reductase
Acetyl CoA	Acetyl coenzyme A
3,4-AHBA	3-Amino-4-hydroxybenzoic acid
3,4-AHBAL	3-Amino-4-hydroxybenzaldehyde
AMP	Adenosine monophosphate
ATP	Adenosine triphosphate
BDMP	Benzoyl dimethyl phosphate
BMP	Benzoyl methyl phosphate
CAR	Carboxylic acid reductase
CoASH	Coenzyme A
FMN	Flavin mononucleotide (oxidized form)
$FMNH_2$	Flavin mononucleotide (reduced form)
IPTG	Isopropyl β-D-thiogalactopyranoside
NAC	*N*-acetylcysteine
NAD^+	Nicotinamide adenine dinucleotide (oxidized form)
NADH	Nicotinamide adenine dinucleotide (reduced form)
$NADP^+$	Nicotinamide adenine dinucleotide phosphate (oxidized form)
NADPH	Nicotinamide adenine dinucleotide phosphate (reduced form)
PLP	Pyridoxal 5´- phosphate
Ppant	4'-Phosphopantetheine
PPi	Pyrophosphate
PPTase	Phosphopantetheine transferase

Content

1. Introduction

1.1 Chemical background

Aldehydes are an important and widely used class of chemical compounds with widespread industrial applications. Due to their high reactivity, short-chain aliphatic aldehydes are commonly used as starting materials in chemical synthesis while long-chain aliphatic aldehydes are building blocks for the preparation of alcohols, carboxylic acids, and amines.[1] Aromatic aldehydes are required as starting materials for the synthesis of pharmaceutical and agricultural chemicals, as well as for perfume industry. Efficient strategies for the production of aldehydes are thereby highly required.[2] However, the high reactivity of aldehydes makes their production a difficult task, especially on an industrial scale.

Non-redox synthesis of aldehydes:

The most important industrial process for the production of aldehydes is the hydroformylation of olefins. This process uses synthetic gas ($CO + H_2$) to react with a double bond, forming an aldehyde and increasing the chain length by one carbon atom.[1] The reaction requires a catalyst (e.g. cobalt compounds), high pressures, and high temperatures, which also influences the ratio of linear to branched aldehyde products.[3] Other industrially often implemented synthetic routes are aldol condensation[1] and redox reactions. Accordingly, there are two possibilities for reaching the oxidation state of aldehydes, either by oxidation of primary alcohols or by reduction of carboxylic acids, similar as in the present work.

Oxidation of primary alcohols:

Reduction of carboxylic acids requires strong reductive agents in order to overcome the energetic barrier that the COOH-group provides. Consequently, the produced aldehyde CHO-group that is much more reactive tends to be further reduced to the alcohol. Instead of trying to find strategies to stop the reaction at the aldehyde state, the most commonly used techniques simply focus on reoxidation of primary alcohols to the aldehyde.[4]

One of the earliest methods for medium- and large-scale applications, was the oxidation *via* the toxic and carcinogenic pyridinium chlorochromate. It enabled the chemoselective oxidation of primary and secondary alcohols to aldehydes and ketones, respectively, at mild conditions without overoxidation to carboxylic acids.[5] Another possibility is the so-called Swern oxidation, which uses DMSO, oxalyl chloride, and triethylamine. This method allows

the oxidation of a wide range of alcohols to their respective aldehydes at low temperatures, however, the side product dimethyl sulfide has a very foul odor.[6,7] Using Dess-Martin periodinane for oxidation of alcohols to aldehydes possesses several advantages over these methods, such as shorter reaction times, easier work-up, and no requirement of toxic or hazardous reagents.[8] Chemoselective, cheap, and fast oxidation of primary alcohols to aldehydes is also achieved by using catalytic amounts of 2,2,6,6-tetramethylpiperidine 1-oxyl (TEMPO). The *N*-oxoammonium salt of the catalyst serves as the actual oxidant which is recycled by sodium hypochlorite as the stoichiometric oxidant.[9] In general, overoxidation of the reactive aldehyde product can happen with strong oxidizing agents, however, overreduction of carboxylic acids to alcohols is even more difficult to avoid.

Reduction of carboxylic acids:

In order to chemoselectively reduce carboxylic acids to aldehydes without overreduction to the alcohol, three main strategies have been established. The most often employed strategy to overcome the energetic barrier of the carboxylic acid group is by performing a two step reduction. In the first step, the acid is activated by catalytic conversion into derivatives (e.g. by hydrosilylation), which are then more reactive towards the following reduction.[10] The direct reduction of carboxylic acids to aldehydes is much rarer and exists in two different approaches. In one such method, strong reducing agents such as bis(dialkylamino)aluminium hydrides or lithium in ethylamine are used in excess.[11,12] The other method proceeds *via* hydrogenation of carboxylic acids for the selective production of aldehydes by the use of metal oxide catalysts, such as ZrO_2 or Cr_2O_3.[13]

1.2 Enzymatic reactions

1.2.1 Enzyme classes

Instead of using chemical means for reduction with the mentioned difficulties, it was shown that the fungus *Polystictus versicolor* is able to reduce several aromatic acids to their corresponding aldehydes and alcohols.[14] Using another fungus, *Neurospora crassa*, it was possible to isolate the enzyme responsible for the reduction of aryl acids to aldehydes. This enzyme, initially named aryl-aldehyde oxidoreductase, required ATP and NADPH, with the molar ratios shown to be 1:1:1 for acid + ATP + NADPH to aldehyde + ADP + $NADP^+$.[15,16] The reaction mechanism was determined to proceed *via* the formation of an acyl-AMP intermediate, followed by reduction with NADPH to the aldehyde product.[17] Kato *et al.*

showed that the reduction of the carboxylic group of benzoic acid to benzyl alcohol was also possible by using the prokaryote *Nocardia asteroides* JCM3016 in whole cell catalysis.[18] The same group later purified the enzyme aryl-aldehyde dehydrogenase (EC 1.2.1.30) or carboxylic acid reductase (CAR) from *Nocardia asteroides* JCM3016 and showed that the mechanism for the reduction of benzoic acid to benzaldehyde also included the formation of a benzoyl-AMP intermediate and was therefore the same mechanism as elucidated for the isolated enzyme from the eukaryote *Neurospora crassa*.[16,19] A different strain of *Nocardia*, which was later named *Nocardia iowensis* sp. NRRL 5646, was also shown to reduce benzoic acid and derivatives.[20,21]

The biosynthesis of the amino acid lysine in lower eukaryotes such as *Saccharomyces cerevisiae* involves a similar enzyme, α-aminoadipate reductase (EC 1.2.1.31). The pathway (fig. 1) starts with α-ketoglutarate which forms homocitrate after reaction with coenzyme A. Homocitrate is then converted to homoisocitrate *via* the intermediate homoaconitate, and α-ketoadipate is formed in the next step. The following pyridoxal 5´-phosphate (PLP)-dependent reaction leads to α-aminoadipate, the substrate for α-aminoadipate reductase (AAR). Condensation of the reduction product, α-aminoadipate-δ-semialdehyde, with glutamate gives saccharopine, which is cleaved to α-ketoglutarate and the final product lysine by oxidative deamination.[22,23] The reduction step of AAR was shown early to require ATP, Mg^{2+}, and NADPH, in a similar way to CAR.[24]

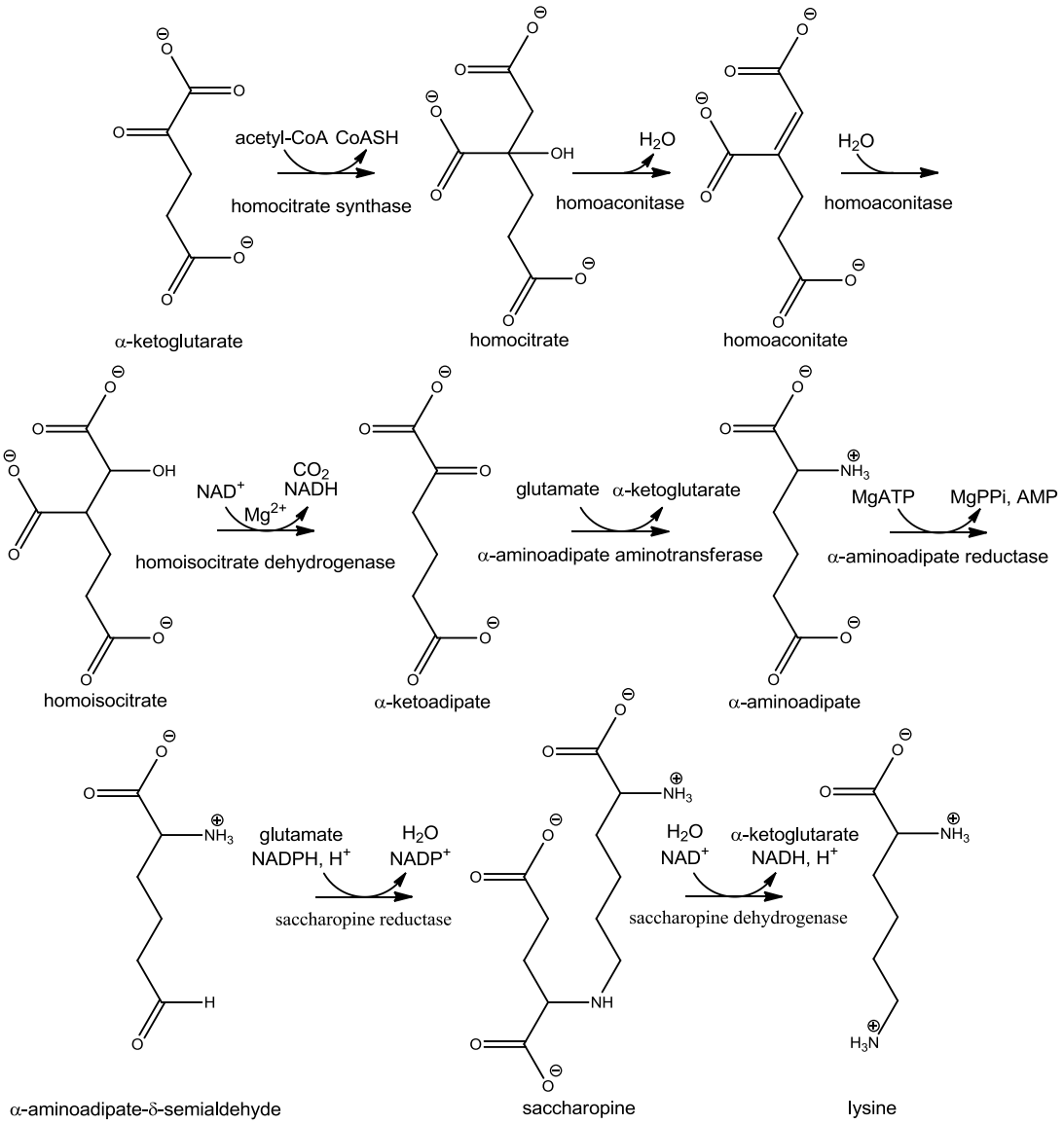


Figure 1: α-Aminoadipate pathway for the synthesis of lysine in fungi. AAR catalyzes the reduction of α-aminoadipate to α-aminoadipate-δ-semialdehyde.[23]

A different kind of carboxylic acid reduction activity is observed in the biosynthesis of grixazone, a yellow pigment consisting of grixazone A and grixazone B, in *Streptomyces griseus*.[25] In the first step of this pathway (fig. 2), 2-amino-4,5-dihydroxy-6-one-heptanoic acid-7-phosphate is synthesized by aldol condensation of aspartate-4-semialdehyde and dihydroxyacetone phosphate, catalyzed by GriI. Another enzyme coded in the grixazone biosynthesis gene cluster, GriH, then converts this intermediate to 3-amino-4-hydroxybenzoic acid (3,4-AHBA).[26] The following ATP- and NAD(P)H-dependent reduction of this aromatic carboxylic acid to 3-amino-4-hydroxybenzaldehyde (3,4-AHBAL) is catalyzed by the enzymes GriC and GriD.[27] Grixazone A is then synthesized from two molecules of 3,4-AHBAL with *N*-acetylcysteine (NAC) by GriE and GriF.[28]

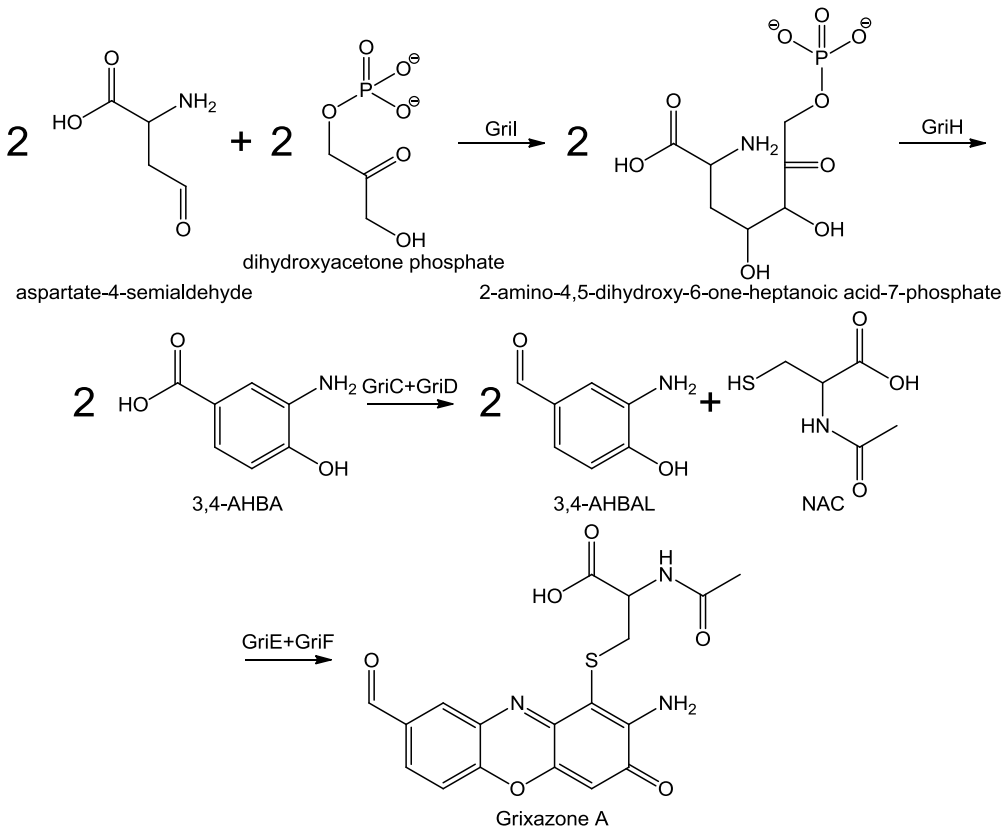


Figure 2: Grixazone biosynthesis in *S. griseus*. GriC and GriD catalyze the reduction of 3,4-AHBA to 3,4-AHBAL.[26,28]

ATP-dependent carboxylate reduction is also a part of the saframycin biosynthesis. Peptide synthetase SfmC from *Streptomyces lavendulae* uses ATP to bind the carboxylic acid residue of a tyrosin derivative, thereby forming a thioester. The thioester is then cleaved and reduced by using NAD(P)H to form an aldehyde intermediate, which eventually gives saframycin in a series of following enzymatic reactions.[29]

Involvement of another ATP-dependent class of carboxylate-reducing enzymes is found in bacterial bioluminescence. Bacterial luciferase catalyzes the oxidation of a long-chain fatty aldehyde to the corresponding acid, which results in light emission at 490 nm. In this reaction, $FMNH_2$ forms a 4a-hydroperoxide with molecular oxygen which then reacts with the aldehyde to give a flavin-aldehyde peroxide. After the carboxylic acid as the oxidation product is released, the excited 4a-hydroxy-substituted flavin remains and the reaction products H_2O and FMN are formed after light emission.[30] The luciferase is encoded by the *lux*AB genes, whereas the *lux*CDE genes code for a multienzyme complex consisting of a transferase, synthetase, and reductase. The transferase subunit cleaves fatty acyl-acyl carrier

protein, fatty acyl-CoA, or other donors and transfers the fatty acid to the synthetase. Substrate preference is shown for a chain length of 14 carbons.[31] The synthetase activates the fatty acid by using ATP to form acyl-AMP, which binds to a cysteine residue of the synthetase.[32] The acylated synthetase transfers the substrate to the reductase subunit, which catalyzes the NADPH-dependent reduction to the aldehyde.[33]

Reduction of activated carboxylic acids occurs in fatty acid biosynthesis as well. Fatty acyl-CoA/acyl carrier protein reductase from *Synechococcus elongatus* was used to produce long-chain fatty aldehydes for further production of fatty alcohols.[34]

Microbial reduction of aliphatic and aromatic carboxylic acids to their corresponding aldehydes and alcohols has also been observed under anaerobic conditions for several bacteria species. In some cases, the tungsten- or molybdenum-containing enzymes were purified, which also selectively reduced the carboxylic acids to aldehydes. Whole cell and isolated enzyme catalysis for the reduction of carboxylic acids to aldehydes or alcohols has also been applied with archea.[35]

Generally, coenzyme A- and metal cofactor-dependent enzymes have both been rarely used in research and large-scale biocatalysis. Coenzyme A-dependent enzymes are unpromising objects of research in this regard, due to the cost of coenzyme A and the substrate limitation to long-chain fatty acids, while metal cofactor-dependent enzymes additionally require intricate handling. Accordingly, these enzymes have been ruled out for further investigation in this work. Using ATP as a cofactor in enzymatic processes is costly as well when it has to be added in stoichiometric amounts without an ATP regeneration system. A crucial parameter for the catalytic efficacy of such cofactor regeneration systems is the total turnover number of the cofactor-dependent enzyme and methods to overcome limitations of low total turnover numbers have not yet been established for ATP regeneration.[36] Therefore, ATP-dependent enzymes were studied in this work with the idea of bypassing the need for ATP by the use of chemically activated substrates. The enzymes *N. iowensis* CAR, *S. cerevisiae* AAR, and *S. griseus* GriC and GriD were chosen for this task and further detail on the reaction mechanisms of these three enzyme classes is provided in the following section.

1.2.2 ATP-dependent enzymes

***Nocardia iowensis* sp. NRRL 5646 CAR:**

As shown in the previous chapter, enzymes from this family were found in different species and were used as a catalyst in whole cells, as well as after enzyme purification. The enzyme catalyzes the reduction of a wide range of aromatic carboxylic acids, such as benzoic acid derivatives, as well as aliphatic acids.[37] The enzyme shows slight stereoselectivity, e.g. ibuprofen was produced in 61.2% ee.[38] The disadvantage of whole cell catalysis is that the produced aldehydes are usually further reduced to the corresponding alcohols, which can be circumvented by only using the isolated enzyme, as shown in case of the reduction of vanillic acid to the aldehyde, vanillin.[39]

Phosphopantetheine transferase:

It was shown that the reduction of carboxylic acids required ATP for the activation of the acid *via* an acyl adenylate intermediate, and that the reduction to the aldehyde is NADPH-dependent.[16] However, the exact mechanism remained unclear until Venkitasubramanian *et al*.[40] showed that purified recombinant CAR in *Escherichia coli* was 50-fold less active than purified native CAR and found by sequence analysis that the amino acid sequence LGGDSLSA of CAR was similar to LGG(H/D)SL, the consensus sequence of the 4'-phosphopantetheine (Ppant)-binding site of acyl carrier proteins.[41]

When the recombinant CAR was incubated with coenzyme A (CoASH) and cell-free extracts (CFE) of *Nocardia iowensis* sp. NRRL 5646 some activity was regenerated. With cell-free extract from *E. coli*, no activity was restored. It was concluded that *Nocardia iowensis* expresses a necessary phosphopantetheine transferase (PPTase) for post-translational phosphopantetheinylation, while *E. coli* expresses an inactive or no PPTase at all. This means that *N. iowensis* CAR is an apoenzyme, requiring post-translational phosphopantetheinylation for conversion into an holoenzyme and to achieve maximum activity. Consequently, coexpression of recombinant CAR and PPTase in *E. coli* significantly improved enzyme activity, further supporting the idea of required PPTase. Labeling experiments confirmed that the acetylphosphopantetheine moiety was incorporated into recombinant CAR in the presence of *N. iowensis* PPTase and labeled acetyl CoA. Further evidence was provided by site-directed mutagenesis in which the conserved Ser^{689} was replaced by Alanine, which led to a completely inactive mutant.[40]

PPTases in general convert acyl carrier protein domains from inactive apo-forms to active holo-forms by attack of a conserved serine residue of the apoenzyme on the pyrophosphate of CoASH, thereby linking the 4-phosphopantetheinyl moiety of CoASH to the serine of the respective protein. Nucleophilic attack of this newly introduced thiol group on a substrate (e.g. aminoacyl-AMP) is required for acyl transfer of the substrate *via* the swinging arm of the phosphopantetheinyl residue. The acyl carrier protein domains to be activated are distinguished into type I peptide synthetase domains of multi-enzyme complexes and separate subunits of type II fatty acid multienzyme complex synthases. *Bacillus subtilis* PPTase, Sfp, is unspecific for activation between type I and II, while *Saccharomyces cerevisiae* PPTase, Lys 5, is specific for *Saccharomyces cerevisiae* AAR, Lys 2.[42] The reaction catalyzed by PPTase, in this case *N. iowensis* PPTase, is shown in figure 3.[40]

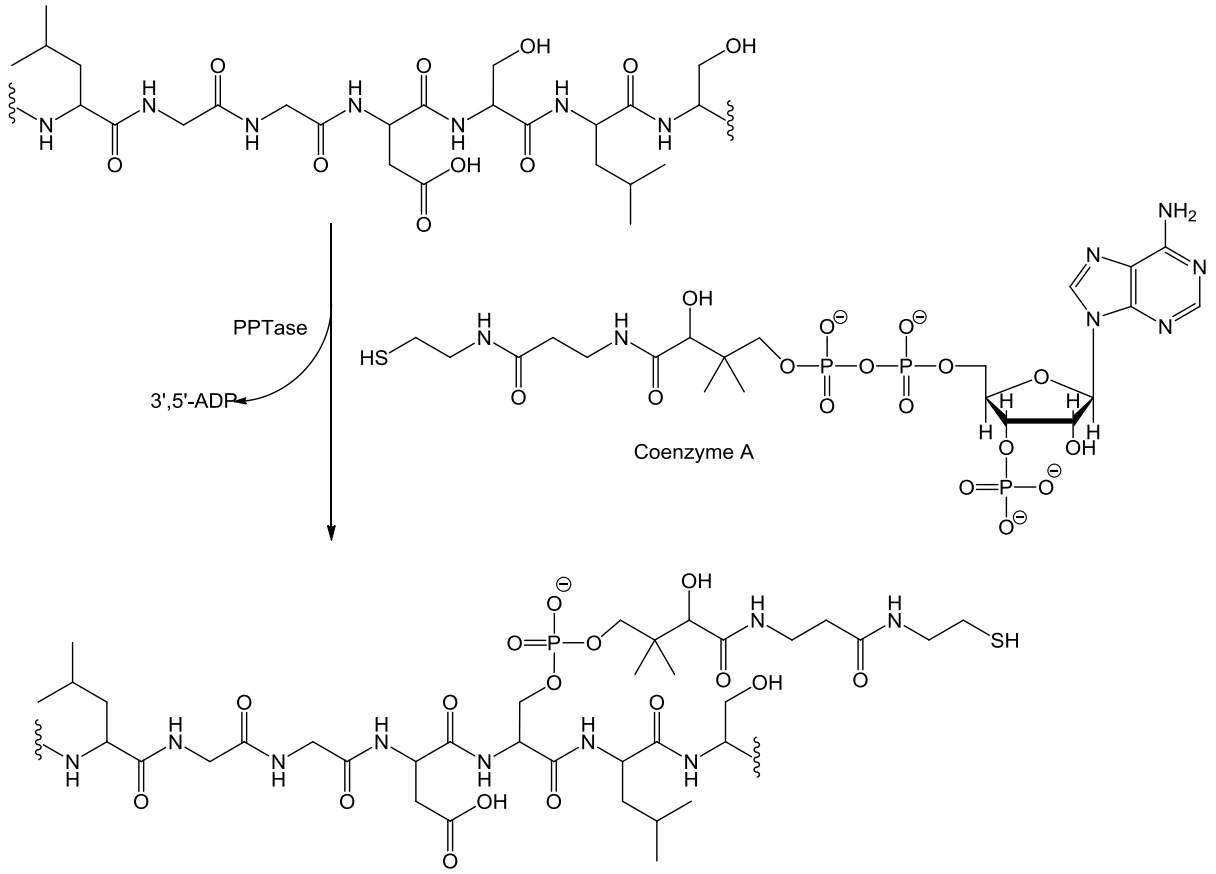


Figure 3: PPTase catalyzes the transfer of a phosphopantetheinyl moiety of CoASH to the serine of an apoenzyme (CAR).[40]

Mechanism:

In addition to the Ppant-binding site, BLAST analysis of the *car* gene revealed high homology of the *N*-terminal domain of the encoded enzyme with known AMP-binding sites, while the *C*-terminal domain showed high homology to NADPH-binding proteins. With these three domains, a mechanism was proposed that after adenylation of the substrate (e.g. benzoic acid), 4-phosphopantetheine provides a swinging arm, which forms a thioester with the acyl-AMP intermediate from the adenylating *N*-terminal domain. After switching to the reductive *C*-terminal domain, NADPH reduces the substrate to release the aldehyde product and to restore the thiol group for another catalytic cycle. The individual steps of the carboxylic acid reduction mechanism are shown in figure 4.[40]

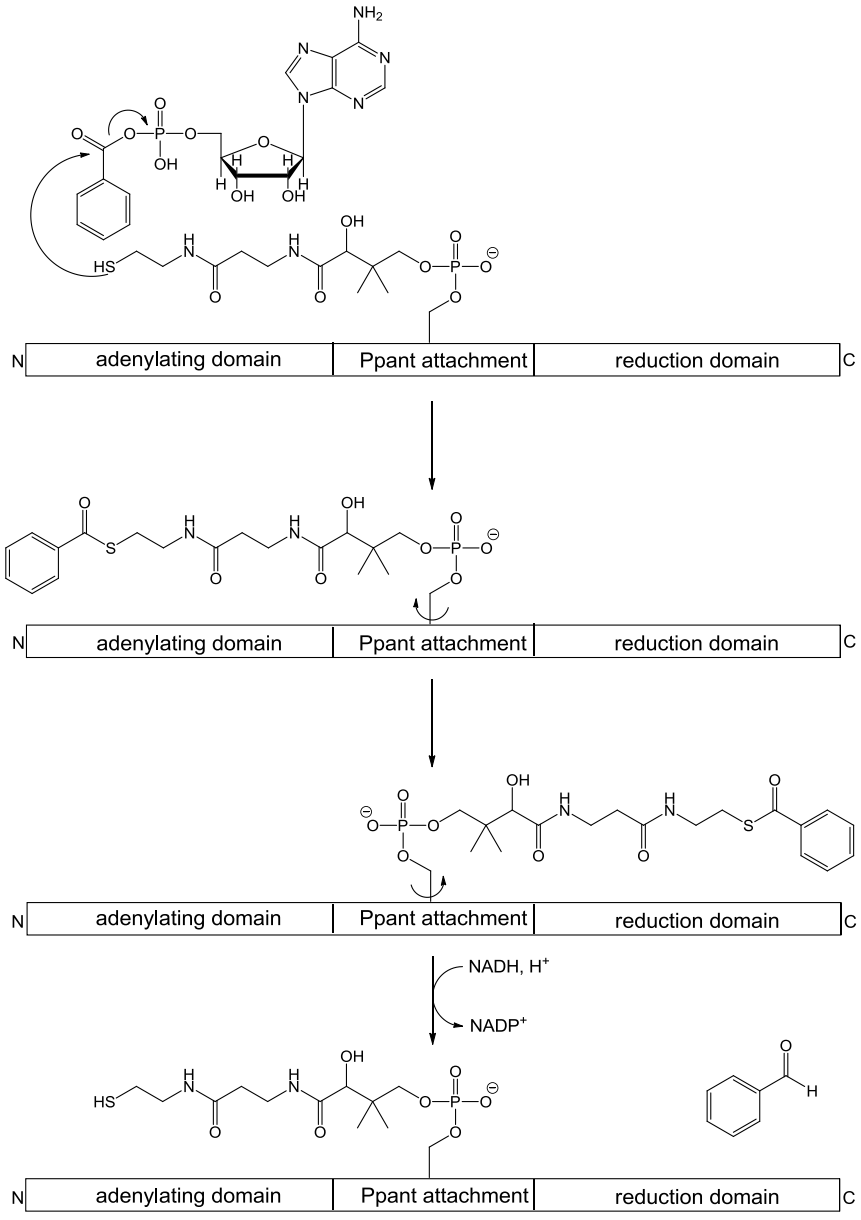


Figure 4: Reaction steps of benzoic acid reduction by holo-CAR.[40]

***Saccharomyces cerevisiae* AAR (Lys2):**

The catalytic cycle in the reduction of α-aminoadipate to α-aminoadipate-δ-semialdehyde was also shown to require ATP, Mg^{2+}, and NADPH.[24] Involvement of the *lys*2 and *lys*5 genes in *S. cerevisiae* in α-aminoadipate reduction was known, yet their exact role remained unclear.[43] Amino acid sequence comparison revealed that the Lys2 protein (AAR) contained an AMP-binding domain signature located in the first third of the protein sequence,[44] while the last third of the sequence revealed similarity to a NAD(P)H-binding motif.[45] Finally, sequence similarity analysis showed homology of Lys5 to PPTases and a Ppant-binding motif LGGHS with a conserved serine was found between the AMP-binding domain and the reductive domain of Lys2, therefore indicating that Lys5 acts as a PPTase required for conversion of apo-Lys2 to holo-Lys2. In the mechanism proposed, based on these findings, α-aminoadipoyl-AMP is formed in the adenylating domain of Lys2. Nucleophilic attack of the thiol group of the phosphopantetheinyl group then attaches the aminoadipoyl moiety to the enzyme. After reduction with NADPH, the α-aminoadipate-δ-semialdehyde product and holo-Lys2 ready for the next catalytic cycle are released.[42] The mechanism that Lys2 reduces the acyl-thioester of α-aminoadipate instead of acyl-AMP, as proposed earlier,[24] was confirmed by Ehmann *et al*.[46] The comparable mechanism for CAR is provided in figure 4. Phosphopantetheinylation of apo-Lys2 is achieved analogously to apo-CAR as shown in figure 3. Catalysis with this enzyme was performed using recombinant Lys2 and Lys5, overexpressed and purified from *E. coli*, however, no other substrate for this enzyme was described yet.[35]

***Streptomyces griseus* GriC and GriD:**

In contrast to *N. iowensis* CAR and *S. cerevisiae* AAR which catalyze this reaction as single enzymes, for the reduction of an aromatic carboxylic acid in the grixazone biosynthesis pathway, the interaction of two separate enzymes, GriC and GriD, is required. Amino acid sequence similarity was found between GriC and AMP-binding proteins, like the *N*-terminal AMP-binding domain of *N. iowensis* CAR, and between GriD and NAD(P)H-dependent aldehyde dehydrogenases. It was therefore suggested that CAR forms and acts as a GriC and GriD fusion protein.[40,47] The consequent reaction mechanism of GriC and GriD is shown in figure 5. In this, GriC and GriD form a carboxylic acid reductase complex, with GriC acting as an adenylating enzyme of 3,4-AHBA to give an acyl-AMP intermediate, which is then reduced to 3,4-AHBAL by GriD using NAD(P)H.[47]

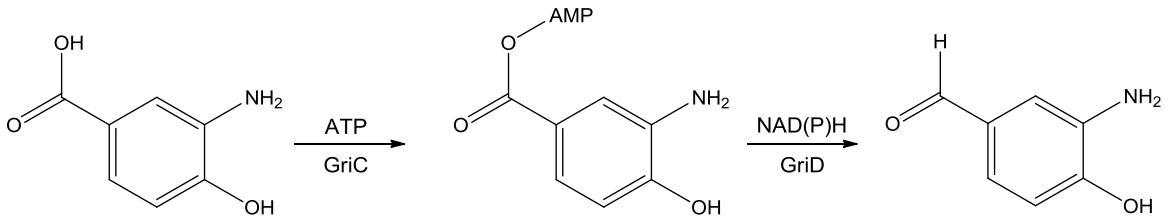


Figure 5: Reduction of 3,4-AHBA to 3,4-AHBAL by cooperative action GriC and GriD.[47]

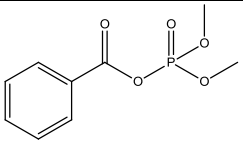
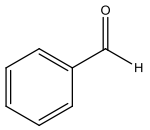
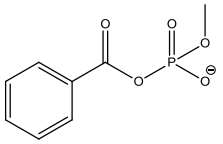
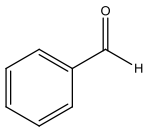
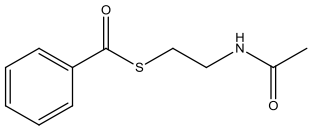
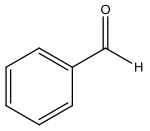
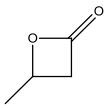
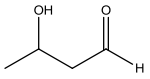
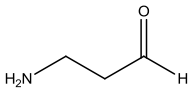
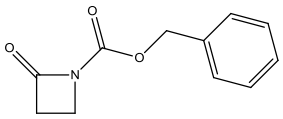
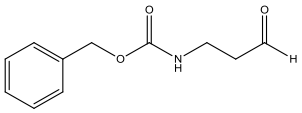
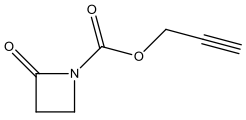
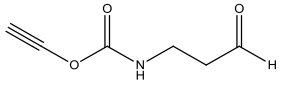
Substrate scope of this enzyme complex is limited, and reactivity was only shown for a few aromatic carboxylic acids. All of them contained a hydroxyl group in *para*-position, including the industrially relevant conversion of vanillic acid to vanillin. No activity was observed for reactions of the enzymes *in vitro*, due to possible stability problems of either enzyme. In this regard, it is also worth noting that the start codon of the *griD* gene and the stop codon of the *griC* gene overlap in the grixazone biosynthesis gene cluster in *S. griseus*. This not only suggests cotranscription and cotranslation of the genes, but it may also have effects on enzyme stability, solubility, and overall activity when the enzymes are recombinantly expressed from different plasmids.[47]

Non-natural substrates:

As detailed in the previous chapters, chemical and enzymatic reduction of carboxylic acids requires activation of the unreactive functional group. Enzymatically, this is achieved by using ATP, which is expensive on an industrial scale. In order to avoid the dependence on ATP, several non-natural substrates were investigated for their enzyme reactivity (table 1). Substrates **1a** and **2a** are derivatives of the natural substrate benzoic acid. As mixed anhydrides of carboxylic acid and phosphate, and mimics of the benzoyl-AMP intermediate, they were thought to be activated and reactive towards the Ppant moiety of the enzyme. Substrate **3a** was chosen because its *N*-acetylcysteine moiety resembles the Ppant swinging arm of holo-CAR, which is generated by CoASH. NAC-derivatives were also shown to be accepted substrates for the Ppant moiety of acyl carrier proteins.[48] Lactones are intramolecular esters with a high ring strain, depending on their ring size, which makes them chemically more reactive compared to underivatized carboxylic acids. Substrate **4a** was also an established electrophilic reaction partner for the Ppant thiol of acyl carrier proteins and therefore chosen as a substrate in this work.[49] As expected, a reactive advantage for thiol-reactivity compared to lactones was confirmed for the lactam substrates **6a** and **7a**, which are

derivatives of the simplest β-lactam **5a**.[50] Overall, these 7 substrates were thought to represent a broad overview and promising candidates for non-natural activity of carboxylic acid reductases.

Table 1: Non-natural substrates and their theoretical reduction products

Subtrate	Product
	
benzoyl dimethyl phosphate (**1a**)	benzaldehyde (**1b**)
	
benzoyl methyl phosphate (**2a**)	benzaldehyde (**1b**)
	
benzoyl-*N*-acetylcysteamine (**3a**)	benzaldehyde (**1b**)
	
β-butyrolactone (**4a**)	3-hydroxybutanal (**4b**)
	
azetidin-2-one (**5a**)	3-aminopropanal (**5b**)
	
benzyl-2-oxoazetidine-1-carboxylate (**6a**)	benzyl (3-oxopropyl)carbamate (**6b**)
	
prop-2-yn-1-yl 2-oxoazetidine-1-carboxylate (**7a**)	ethynyl (3-oxopropyl)carbamate (**7b**)

1.3 Motivation

Three enzyme classes of aryl-aldehyde dehydrogenases (from EC 1.2.1.30 and 1.2.1.31) were chosen for investigation in this work. Their respective genes were ordered for recombinant expression and purification, followed by detailed analysis of their behavior in biocatalytic reactions. For two of the chosen enzymes, *N. iowensis* CAR and *S. cerevisiae* AAR, coexpression of auxiliary enzymes was also necessary in order to obtain active holoenzymes. Immediate necessity for the purified enzymes was optimization of the reaction conditions described in literature, in regard to enzyme activity but also simplicity of the reaction setup. Main point of focus for this task was finding the ideal buffer composition, including the compounds required for enzyme activity (in ideal concentration) but excluding compounds which are not necessary or may even be impeding. Other improvable variables were the buffer system itself, pH, reaction temperature, and NADPH recycling system. In regard to the enzyme, storage conditions, stability, and enzyme concentration in biotransformations were to be determined. The natural enzyme substrate benzoic acid was used for these optimization screenings.

The main goal of this work was then to expand the substrate scope of the enzyme not only in regard to other carboxylic acids, but also towards activated carboxylic acid derivatives. As outlined in the introduction section, it was planned that the necessity of using expensive ATP in this reaction could be circumvented by using such chemically activated substrates. Due to the chemically highly unreactive nature of carboxylic acid moieties, activation is achieved *in vivo* by acyl-AMP formation. Executing the first part of this reaction mechanism *via* classical chemical synthesis, while still letting the enzyme perform the final part of the reaction provided a highly attractive goal for this work with multiple possible applications in biocatalytic industry and research. Organic synthesis of chemically activated carboxylic acid derivatives is an uncomplicated and cheap alternative to the use of ATP.

Candidates for potential reduction of non-natural substrates were thought to be different kinds of activated esters or anhydrides, e.g. phosphoric acid mixed anhydrides, thioesters, lactones, or lactams. In order to get as wide a range of diverse substrates as possible, 7 substrates of different groups were chosen as an overview (table 1). As chemically activated derivatives of benzoic acid, the natural substrate, benzoyl dimethyl phosphate (BDMP) and benzoyl methyl phosphate (BMP) as mixed anhydrides of carboxylic acid and phosphate were synthesized. Benzoyl-*N*-acetylcysteamine was chosen as a thioester that also resembles phosphopantetheinylated benzoic acid. β-Butyrolactone and azetidin-2-one were chosen as a simple lactone and lactam, respectively, while benzyl-2-oxoazetidine-1-carboxylate and prop-2-yn-1-yl 2-oxoazetidine-1-carboxylate were chosen as more complex derivatives of lactams with carboxylic acid esters.

2. Results and discussion

2.1 Aldehyde dehydrogenases

In preliminary tests, six different aldehyde dehydrogenases (ALDH) were tested on their ability to oxidize aldehydes to the corresponding carboxylic acids, hence the opposite direction of the reaction that was going to be examined in the later work. The tested substrates were benzaldehyde (**1b**) and its product benzoic acid (**1c**), and glyceraldehyde (**8b**) and its product glyceric acid (**8a**), respectively. The ALDHs were ALDH4 from *Saccharomyces cerevisiae*,[51] BALDH from *Acinetobacter sp.* ADP1,[52] Lreu0034 (succinate-semialdehyde dehydrogenase) from *Lactobacillus reuteri* DSM 20016,[53] NAD(P)H-dependent AldH (gamma-glutamyl-gamma-aminobutyraldehyde dehydrogenase from *Escherichia coli* str. K12 substr. MG1655,[54] Gox0499 from *Gluconobacter oxydans*,[55] and Gox1712 from *Gluconobacter oxydans*.[56]

Transformation, protein expression, and cell disruption were performed according to the experimental section, the obtained cell-free extract was then used in a photometric activity assay, results are shown in table 2.

Table 2: Change in substrate concentration according to absorption of NADH at 340 nm after 10 min reaction time by several aldehyde dehydrogenases oxidizing two different substrates. 10 mM starting concentration for both substrates, 2 mM starting concentration for NAD^+

Enzyme	Δ c Benzaldehyde (**1b**) [µmol/L*min]	Δ c Glyceraldehyde (**8b**) [µmol/L*min]
BALD	29	0.2
ALDH4	0.3	0.1
AldH	34	30
Lreu0034	4.3	1.4
Gox 0499	0.9	3.2
Gox 1712	0.1	22

The results showed that all enzymes are active with both substrates, with differences in activity between the enzymes, as well as between the substrates. AldH was shown to be the most active enzyme, with similarly high activities for both substrates, while ALDH4 was the slowest enzyme for both substrates. Enzymes BALD and Gox 1712 exhibited the largest discrepancy in activity towards both substrates.

2.2 Carboxylic acid reductase

2.2.1 Enzyme preparation and activity tests

The first overexpression of carboxylic acid reductase from *Nocardia iowensis* sp. NRRL 5646[21] in *E. coli* was conducted as described in the experimental section (3.3). After cell disruption in Tris buffer (50 mM, pH 7.5), *Nocardia iowensis* CAR was soluble in the cell-free extract (SDS-PAGE in figure 3) and used in a photometric activity assay with reaction buffer B and 50 µL CFE.

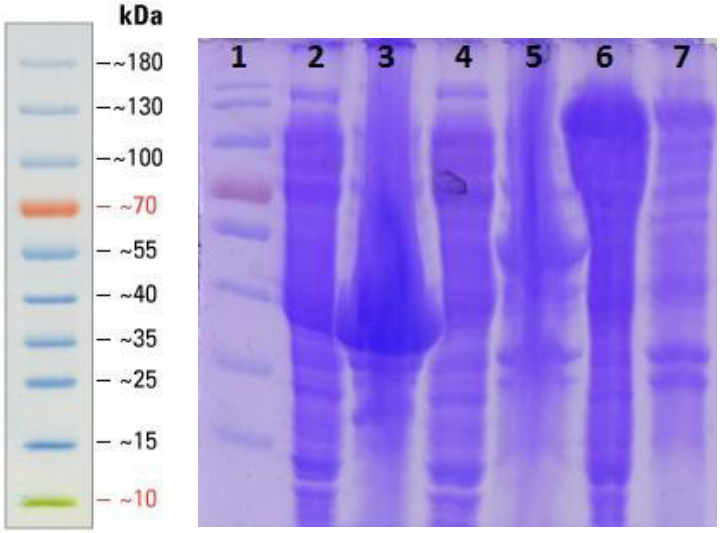


Figure 3: SDS-PAGE of GriC (39 kDa), GriD (48 kDa), and CAR (128 kDa) after cell disruption; 1: standard, 2: GriC lysate, 3: GriC pellet, 4: GriD lysate, 5: GriD pellet, 6: CAR lysate, 7: CAR pellet

One half of the cell lysate was stored at 4 °C and the other half was shock-frozen in liquid nitrogen and stored at -20 °C. The enzyme activity was determined qualitatively by the change in concentration of NADPH, which indicates the reduction of benzoic acid (**1c**) to benzaldehyde (**1b**). The photometric activity assay was repeated after 4 days with 50 µL and 150 µL cell-free extract in order to examine the enzyme stability and to find optimum storage conditions. As no significant difference of enzyme activity was observed, storage at – 20 °C was used for future applications. In the following assay, the capability of the enzyme to reduce a non-natural substrate was tested by using reaction buffer B with synthesized benzoyl methyl phosphate (**2a**, 5 mM) without the addition of ATP and $MgCl_2$. No activity was determined when using the non-natural substrate **2a** with CFE of *Nocardia iowensis* sp. NRRL 5646 CAR expressed without *Bacillus subtilis* PPTase.

Since CAR was described to require a prosthetic phosphopantetheine-group for full activity,[40] it was expressed together with *Bacillus subtilis* PPTase in a second step to increase activity. After a transformation of the *Bacillus subtilis* PPTase plasmid into self-made competent BL21 (DE3) cells containing the *Nocardia iowensis* sp. NRRL 5646 CAR plasmid, overexpression in TB-medium, harvesting and cell disruption in Tris buffer (50 mM, pH 7.5) were carried out as described in the experimental section. A photometric activity assay was performed with reaction buffer B and with substrate **2a** without the addition of ATP and $MgCl_2$. For the positive control, reaction buffer B including ATP, $MgCl_2$, and benzoic acid (**1c**) was used. Substrate **1c** was omitted for the respective negative control reaction.

As the *car* gene from *Nocardia iowensis* sp. NRRL 5646 was initially expressed with a stop codon right before the His-tag sequence, the His-tag was not expressed in the protein sequence. Using the HisTrap™ FF 5 mL column could therefore not entirely separate and purify the enzyme alone, but more or less all proteins in the cell extract, depending on their histidine content on the protein surface. Nevertheless, a reasonable amount of *N. iowensis* CAR was overexpressed and "semi-purified" from the cell-free extract, according to SDS-PAGE (figure 4).

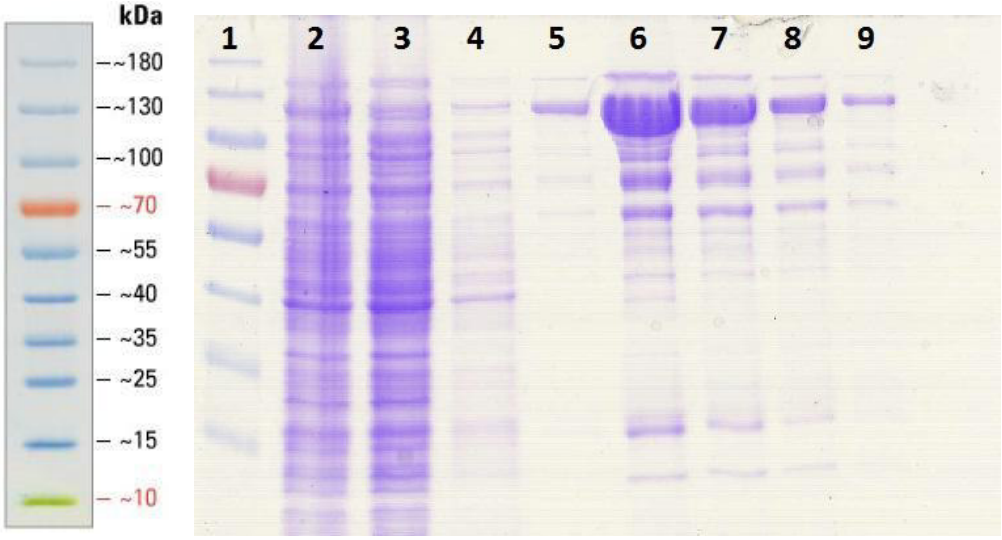


Figure 4: SDS-PAGE of *N. iowensis* CAR with coexpressed *B. subtilis* PPTase (129 kDa) after the first purification attempt; 1: standard, 2: lysate, 3: flowthrough, 4: wash fraction, 5 - 9: fraction 5 - 9

Protein concentration after dialysis was determined by measuring the absorption at 280 nm. The enzyme activity was measured *via* a photometric activity assay, using **1c** as ATP-dependent and **2a** as ATP-independent test substrates. The combined results for the different assays are shown in table 3.

Table 3: Results of activity assays of CAR stored under different storage, expression, and purification conditions. [a] expression of *N. iowensis* CAR, [b] coexpression of *N. iowensis* CAR and *B. subtilis* PPTase; n.a. no activity. Change in concentration of NADPH over the course of 10 min. 0.2 mM starting concentration for NADPH

Type of enzyme	Amount of CFE/enzyme	Substrate (5 mM)	Δc [µmol/L]
CFE[a] stored at 4 °C (1 day)	50 µL	**1c**	7.6
CFE[a] stored at - 20 °C (1 day)	50 µL	**1c**	4.8
CFE[a] stored at 4 °C (4 days)	50 µL	**1c**	6.5
CFE[a] stored at - 20 °C (4 days)	50 µL	**1c**	8.6
CFE[a] stored at 4 °C (4 days)	150 µL	**1c**	10.8
CFE[a] stored at - 20 °C (4 days)	150 µL	**1c**	9.3
CFE[a]	150 µL	**2a**	n.a.
CFE[b]	150 µL	**1c**	4.6
CFE[b]	150 µL	**2a**	n.a.
Semi-purified enzyme[b]	50 µg	**1c**	7.8
Semi-purified enzyme[b]	50 µg	**2a**	n.a.

The activity of the enzyme in the CFE was slightly lower in total change of concentration when *N. iowensis* CAR is coexpressed with *B. subtilis* PPTase, compared to CAR alone. No activity is detected with benzoyl methyl phosphate (**2a**) as the substrate.

2.2.2 Reaction engineering

Having the active enzyme available, end-point activity after 24 h was determined by gas chromatography. In the first assay, optimum reaction conditions were identified by variation of the addition of EDTA, $MgCl_2$, DTT, and glycerol, respectively.[38] **1c** was used as standard test substrate. The reaction was conducted as described in the experimental section, with the variations indicated in table 4. Additionally, enzyme that was refrozen and thawed once and twice, respectively, was tested in order to check the enzyme durability and enzyme concentration was varied to determine the optimum enzyme amount. No enzyme was added for a negative control reaction.

The results showed that excluding EDTA and glycerol, respectively, was not disadvantageous. Excluding $MgCl_2$ resulted in a complete loss of enzyme activity. Even excluding all the components in question except for $MgCl_2$ did not show negative results, so Mg^{2+} is clearly required for this ATP-dependent enzyme. Therefore, the optimum buffer had to contain $MgCl_2$, whereas EDTA and glycerol could be omitted. Since a higher enzyme loading led to higher conversion, 100 µg enzyme/sample were used for all of the following screenings. Repeated freezing and thawing decreased enzyme activity, so only unthawed enzyme was used.

Furthermore, the influence of pH and reaction temperature, as well as a different buffer system were investigated. It was tested at pH 7.5 for direct comparison with the previous screening and at pH 6, 7, 8, and 9 to find the optimum pH. Additionally, the same pH values were tested with K-phosphate buffer. For finding the optimum reaction temperature, samples in 50 mM Tris buffer pH 7.5 were incubated in a thermocycler at 300 rpm and 25 °C, 30 °C, and 35 °C, respectively. This buffer system was also used for negative control reactions and with DTT (1 mM) to clarify the influence of DTT.

Performing the reaction in phosphate buffer clearly led to higher conversions than in Tris buffer. Conversion also increased with increasing pH, regardless of the buffer system. Adding DTT here showed significantly improved conversion, so it was included in further screenings.

The different shaking conditions in a thermocycler had slightly positive effects, however, the varying reaction temperature did not significantly change the outcome of the reaction.

In the next experiment, a final comparison of phosphate buffer at pH 8 and pH 9 was performed. Additionally, the difference between addition of NADPH in excess and in catalytic amounts using an NADPH recycling system ($NADP^+$, glucose, and glucose dehydrogenase)[57] was examined, as well as again the influence of DTT, each in phosphate buffer at pH 8 and pH 9. The collected results for all reaction engineering reactions are displayed in table 4.

Table 4: Collected results for reaction engineering experiments. 50 mM buffer concentration for each buffer. [a] refrozen once, [b] refrozen twice, [c] incubation at 25 °C, [d] incubation at 30 °C, [e] incubation at 35 °C; n.c. no conversion; cofactor: NADPH (15 mM) or recycling system of $NADP^+$ (1 mM), glucose (30 mM), and glucose dehydrogenase (0.2 mg/mL)

Enzyme amount [µg]	Buffer	Cofactor	pH	EDTA	Glycerol	DTT	Mg^+	Conversion [%]
25	Tris	NADPH	7.5	+	+	+	+	1
50	Tris	NADPH	7.5	+	+	+	+	3
50[a]	Tris	NADPH	7.5	+	+	+	+	3
50[b]	Tris	NADPH	7.5	+	+	+	+	2
100	Tris	NADPH	7.5	+	+	+	+	7
50	Tris	NADPH	7.5	-	+	+	+	3
50	Tris	NADPH	7.5	+	+	-	+	2
50	Tris	NADPH	7.5	+	+	+	-	n.c.
50	Tris	NADPH	7.5	+	-	+	+	5
50	Tris	NADPH	7.5	-	-	-	+	3
100	Tris	Rec. Sys.	6	-	-	-	+	1
100	Tris	Rec. Sys.	7	-	-	-	+	10
100	Tris	Rec. Sys.	7.5	-	-	-	+	18
100	Tris	Rec. Sys.	8	-	-	-	+	19
100	Tris	Rec. Sys.	9	-	-	-	+	30
100	Tris	Rec. Sys.	7.5	-	-	+	+	55
100[c]	Tris	Rec. Sys.	7.5	-	-	-	+	24

Enzyme amount [µg]	Buffer	Cofactor	pH	EDTA	Glycerol	DTT	Mg^+	Conversion [%]
100[d]	Tris	Rec. Sys.	7.5	-	-	-	+	23
100[e]	Tris	Rec. Sys.	7.5	-	-	-	+	22
100	Phosphate	Rec. Sys.	6	-	-	-	+	17
100	Phosphate	Rec. Sys.	7	-	-	-	+	31
100	Phosphate	Rec. Sys.	8	-	-	-	+	43
100	Phosphate	Rec. Sys.	9	-	-	-	+	45
100	HEPES	Rec. Sys.	7	-	-	-	+	11
100	Phosphate	Rec. Sys.	8	-	-	-	+	72
100	Phosphate	Rec. Sys.	8	-	-	+	+	82
100	Phosphate	NADPH	8	-	-	-	+	40
100	Phosphate	NADPH	8	-	-	+	+	72
100	Phosphate	Rec. Sys.	9	-	-	-	+	69
100	Phosphate	Rec. Sys.	9	-	-	+	+	84
100	Phosphate	NADPH	9	-	-	-	+	44
100	Phosphate	NADPH	9	-	-	+	+	74

In regard to the optimum pH, pH 9 led to slightly higher conversions than pH 8 in phosphate buffer. DTT again proved to be required for optimum conversion. It was not necessary to include NADPH directly, using the NADPH recycling system even showed better results.

Summarizing the results of the different screenings, K-phosphate buffer at pH 9 including DTT and $MgCl_2$, excluding EDTA and glycerol, was shown to be the optimum reaction conditions. NADPH was supplied using a recycling system of $NADP^+$, glucose, and glucose dehydrogenase. The following reactions were conducted under these reaction conditions (see reaction buffer E in the experimental section).

2.2.3 Biotransformations

Two additional substrates, 4-hydroxyphenylacetic acid (**9a**) and phenylacetic acid (**10a**) (10 mM each), were tested in a HEPES-buffer system at pH 9. The general functionality of the enzyme in this buffer was already confirmed with the standard substrate benzoic acid (**1c**) at pH 7 (table 3). Substrates **9a** and **10a** were also tested in Tris buffer (50 mM, pH 9), results are shown in table 5.

Table 5: Reduction of two additional aromatic substrates, 50 mM buffer concentration and pH 9 for each buffer, cofactor: NADPH recycling system of $NADP^+$ (1 mM), glucose (30 mM), and glucose dehydrogenase (0.2 mg/mL)

Enzyme amount [µg]	Buffer	Substrate	EDTA	Glycerol	DTT	Mg^+	Conversion [%]
100	Tris	**9a**	-	-	+	+	29
100	HEPES	**9a**	-	-	+	+	49
100	Tris	**10a**	-	-	+	+	12
100	HEPES	**10a**	-	-	+	+	32

Enzyme activity was shown for both substrates in both buffers. HEPES buffer was the preferred buffer system for both substrates compared to Tris buffer. Higher conversion was achieved with substrate **9a** compared to substrate **10a** in both buffers.

In the following experiments, the main question of this work was targeted, whether it was possible to use chemically activated carboxylic acid derivatives as substrates in an ATP-independent reaction.

The reaction was performed under the identified optimum conditions with reaction buffer E as the standard buffer system. However, as in theory ATP would not be needed for activating the carboxylic acid, ATP and $MgCl_2$ were excluded from the buffer. Substrates **1a** to **7a** that were tested as well as their potential reduction products are shown in table 1 in the introduction section. The buffer for the positive control included ATP, $MgCl_2$, and the natural substrate benzoic acid (**1c**).

Unfortunately, while the enzyme was shown to be active according to the positive controls, no conversion of the non-natural substrates to any of the theoretically expected products could be

detected using GC or GC-MS. There are several possible explanations as to why the conversion did not occur, depending on the different steps in the catalytic cycle. Substrates **1a** and **2a** are artificially activated mimics of acyl-AMP and should theoretically enter the reaction at the stage after adenylation. However, due to the different interaction with the enzymatic active site compared to the bulkier acyl-AMP, the Ppant moiety might not recognize and bind these substrates. Substrates **3a** to **7a** are reported in literature to react with Ppant moieties,[48,49,50] yet steric hindrance might still prevent the availability of the substrates for the Ppant moiety if no adenylation is involved. Additionally, the reduction reaction itself might not occur for the artificial substrates.

2.2.4 Purified enzyme

The redundant stop codon in the gene sequence was removed by executing a PCR as described in the experimental section (3.2). After a 1% agarose gel separation and the subsequent gel extraction, the gene was ligated into a pET28 vector. The plasmid that was obtained after the transformation was then used for another transformation into the prepared chemically competent BL21 (DE3) cells with the *B. subtilis* PPTase plasmid. Overexpression and purification was then performed according to the experimental section, SDS-PAGE of the protein-containing fractions is displayed in figure 5.

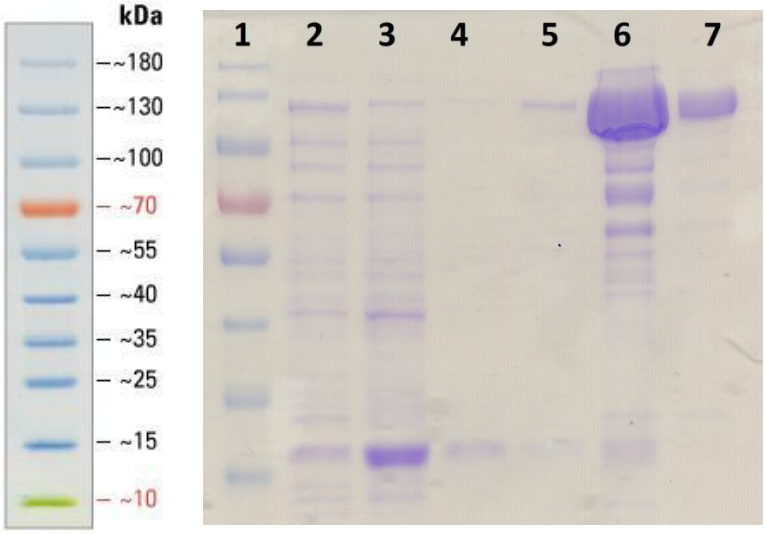


Figure 5: SDS-PAGE of purified *Nocardia iowensis* CAR (129 kDa) after coexpression with *B. subtilis* PPTase (26 kDa); 1: standard, 2: fraction 1, 3: fraction 2, 4: fraction 4, 5: fraction 5, 6: fraction 6, 7: fraction 9

SDS-PAGE in figure 5 shows the purification of the His-tagged enzyme. It can be clearly seen that the two overexpressed enzymes are separated into fractions 1 to 4 (PPTase) and 5 to 9 (CAR), so that fractions 5 to 9 were then concentrated. Desalting on the desalting column was performed with phosphate buffer (50 mM, pH 9) as this was going to be used for the experiments.

The last experiment with the "semi-purified" enzyme was then repeated analogously with the new enzyme for the non-natural substrates in table 1, including substrates **6a** and **7a**, but excluding self-synthesized BDMP (**1a**), which was not available anymore. The reactions with substrates **6a** and **7a** were additionally conducted including ATP.[58] Again, no conversion of non-natural substrates could be detected neither with nor without the addition of ATP. Impurities of other proteins or protein fragments were removed by FPLC, so the explanation as to why no conversion occurred has to be linked to the enzyme itself. The reasons discussed for the semi-purified enzyme are still applicable here, so the steps of the reaction with Ppant or the reduction may be impaired by non-natural enzyme-substrate interactions.

2.3 α-Aminoadipate reductase

Sequencing of the plasmid that was obtained after the Gibson Assembly® showed 3 point mutations which also translated into a mutated amino acid sequence. Repeating the Gibson Assembly® was not successful and by using site-directed mutagenesis, only one of the mutations could be removed. In regard to the difficulty and length of time repeated site-directed mutageneses would require,[59] work on α-aminoadipate reductase was discontinued, also considering its natural substrate limitation to α-aminoadipate,[35] making it unlikely to accept non-natural substrates.

2.4 GriC/D

GriC and GriD expressed separately:

After overexpression of *Streptomyces griseus* GriC and GriD in *E. coli* and cell disruption in Tris buffer (50 mM, pH 7.5), one half of each cell lysate was stored at 4 °C and the other half was shock-frozen in liquid nitrogen and stored at - 20 °C. Although GriC and GriD were very insoluble in the cell-free extract (SDS-PAGE in figure 3), they were still used in a photometric activity assay. Enzyme activity was qualitatively determined by the change in concentration of NADPH. The photometric activity assay was repeated after 4 days with 50

µL and 150 µL CFE, respectively, to test the enzyme stability after storage under different conditions. As for *Nocardia iowensis* sp. NRRL 5646 CAR, in the next assay the non-natural substrate BMP (**2a**) was tested for enzyme activity with 150 µL CFE of GriC and GriD, respectively, stored at – 20 °C.

The change in NADPH concentration here was clearly lower than for *Nocardia iowensis* sp. NRRL 5646 CAR (table 3) and activity of GriD alone would not be expected, so no unambiguous NADPH-dependent reducing activity could be assumed even for benzoic acid (**1c**) and no activity for the non-natural substrate (**2a**) was observed.

Both proteins were overexpressed in *E. coli* again, disrupted, purified, and dialysed (details in the experimental section). However, also the genes for GriC and GriD were designed and expressed with a stop codon right before the His-tag sequence. HisTrap™ purification could therefore not specifically separate the respective proteins alone, however, the purification for GriC worked well nevertheless (figure 6). For GriD the purification itself was successful as well, however, the concentration in the lysate was very low. Protein concentration after dialysis was determined by measuring the absorption at 280 nm.

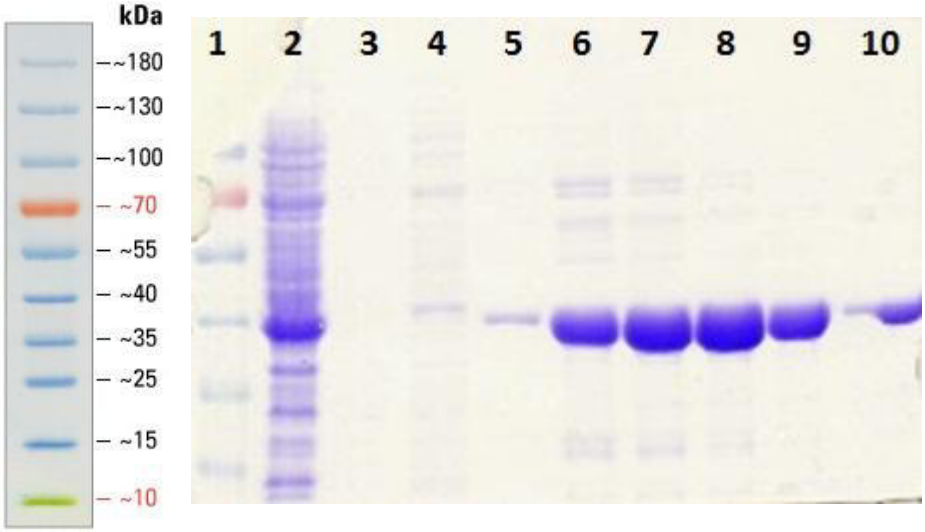


Figure 6: SDS-PAGE for purified GriC (39 kDa), 1: standard, 2: lysate, 3: flow through, 4: wash fraction, 5 to 10: fractions 3 to 8

Another photometric activity assay was performed with 50 µg of either enzyme and benzoic acid (**1c**) and BMP (**2a**) as substrates. The reaction was executed with and without the addition of ATP and $MgCl_2$, respectively. No activity was detected for either substrate. The combined results for the assays are shown in table 6.

Table 6: Results of activity assays of GriC and GriD after different storage and purification conditions; n.a. no activity. Change in concentration in respect to NADPH over the course of 10 min. 0.2 mM starting concentration for NADPH

Type of enzyme	Amount of CFE/enzyme	Substrate (5 mM)	Δc [µmol/L]
CFE[a] stored at 4 °C (1 day)	50 µL (GriD)	**1c**	3.3
CFE[a] stored at 4 °C (4 days)	50 µL (GriD)	**1c**	1.4
CFE[a] stored at - 20 °C (4 days)	50 µL (GriD)	**1c**	4.7
CFE[a] stored at 4 °C (4 days)	50 µL (GriC) + 50 µL (GriD)	**1c**	4.0
CFE[a] stored at - 20 °C (4 days)	50 µL (GriC) + 50 µL (GriD)	**1c**	2.3
CFE[a] stored at 4 °C (4 days)	150 µL (GriD)	**1c**	6.5
CFE[a] stored at - 20 °C (4 days)	150 µL (GriD)	**1c**	3.6
CFE[a] stored at 4 °C (4 days)	150 µL (GriC) + 150 µL (GriD)	**1c**	5.8
CFE[a] stored at - 20 °C (4 days)	150 µL (GriC) + 150 µL (GriD)	**1c**	3.7
CFE[a]	150 µL (GriC) + 150 µL (GriD)	**2a**	n.a.
Semi-purified enzyme[b]	50 µg (GriC) + 50 µg (GriD)	**1c**	n.a.
Semi-purified enzyme[b]	50 µg (GriC) + 50 µg (GriD)	**2a**	n.a.

Coexpression of GriC and GriD:

Due to the complete loss of activity after purification, it was then suspected that GriC and GriD have to be coexpressed in order to be more soluble and to show enzyme activity, so the gene coding for GriC was re-ligated into a different vector [pET 21a(+)] to be transformed into chemically competent BL21 (DE3) cells already containing the GriD plasmid. The steps up to cell disruption were performed according to the experimental section with two slightly different procedures after induction of protein expression: one culture was incubated at 20 °C, 120 rpm overnight after adding IPTG and culture was incubated for only 3.5 h at 37 °C, 120 rpm. SDS-PAGE of the coexpressed proteins for both methods is shown in figure 7.

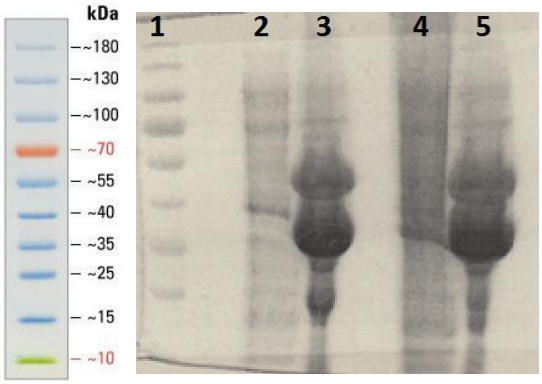


Figure 7: SDS-PAGE of coexpressed GriC (39 kDa) and GriD (48 kDa) at 20 °C and 37 °C, respectively. 1: standard, 2: 37 °C lysate, 3: 37 °C pellet, 4: 20 °C lysate, 5: 20 °C pellet

No significant difference for the induction procedures was observed and in both cases GriC, as well as GriD, were highly unsoluble. Nevertheless, an enzymatic screening in 800 µL reaction buffer E was performed using 100 µL CFE from both induction methods and additionally, a biotransformation in living/fermenting cells was conducted according to the experimental section. After work-up, the samples from both experiments were analyzed using GC-MS but no conversion to benzaldehyde (**1b**) was detected.

2.5 Conclusion and outlook

During the course of this work, 7 different non-natural, activated carboxylic acid derivatives have been tested as potential candidates for ATP-independent reduction to their corresponding aldehydes. The enzyme that was used for this task was carboxylic acid reductase from *Nocardia iowensis* sp. NRRL 5646. Unfortunately, for none of the substrates conversion could be detected. Whether the activation of the non-natural substrate was not adequate or whether the substrates were inactive due to steric reasons remains unclear. Carboxylic acid reductases are highly complex enzymes consisting of three domains, which have to work in sequence – carboxyl group activation of the substrate, attachment to the enzyme, and nicotinamide-dependent carboxyl reduction - in order to produce the desired products. Inefficiency in any of the three catalyzed steps can lead to an inactive enzyme. Future work with this enzyme can still follow this direction in finding other carboxylic acid derivatives which might be accepted substrates.

If none of those attempts show success, one could think of individual expression and characterization of the three enzyme domains. Substrates **1a** (BDMP)and **2a** (BMP) are chemically activated as phosphor-anhydrides, mimicking small variants of acyl-AMP, therefore, the acyl-AMP formation in the adenylating *N*-terminal domain of CAR would not be required. Expressing the enzyme without the adenylating domain might create easier access for the substrate to be picked up by the 4-phosphopantetheinyl moiety for transfer to the reducing domain. CAR has been described as a fusion protein of the enzyme system of GriC and GriD, which does not possess a 4-phosphopantetheinyl moiety,[40,47] so it might also appear possible to not only remove the adenylating, but also the phosphopantetheinylating domain of CAR to express the reducing domain alone. This additional removal of steric hindrance might further facilitate reduction of this kind of non-natural, activated carboxylic acid derivatives.

However, also with their natural substrates, this whole class of enzymes still provides many useful applications in biocatalytic synthesis. As shown here, chemoselective carboxylic acid reduction to aldehydes can be performed using a purified enzyme instead of whole cell catalysis. One of the disadvantages of using purified enzymes compared to whole cells is the necessity of providing cofactors like NAD(P)H which would be produced by the cells, but it is also shown here that this enzyme works with an NADPH recycling system as well.

Aldehydes are highly required in chemical industry, so the natural substrate scope can be broadened in finding new carboxylic acids as enzyme substrates. In general, aldehyde synthesis in an economically and ecologically convenient way as in biocatalysis, makes this enzyme a worthy subject for future investigation under many different aspects.

The enzyme α-aminoadipate reductase could not be prepared in an active form, so it remains questionable if this enzyme provides additional opportunities in this enzyme class. However, the limited substrate scope of this enzyme[35] makes non-natural activity even more unlikely than for *Nocardia iowensis* sp. NRRL 5646 CAR.

The two-enzyme system of GriC and GriD appeared to be a likely candidate for the proposed non-natural activity because of its different catalytic mechanism. GriC provides an activated acyl-AMP substrate, which is mimicked as a smaller variant by substrates **1a** and **2a**, so even GriD alone could have been thought to reduce these substrates. However, no non-natural activity has been achieved with these enzymes for any substrate. The exact reason remains unclear, as protein solubility was very low throughout the course of this work and also coexpression in living/fermenting cells did not lead to aldehyde formation even for a carboxylic acid substrate. *In vivo*, the genes coding for these enzyme overlap, as shown in the introduction section, so possibly the enzymes may therefore not be active when coded and expressed from different plasmids. Also, *in vitro* reactions with the proteins as well as with cell-free extracts of *E. coli* and *S. griseus* attempted by Suzuki *et al*. were not successful, with similar problems to those in this work. Consequently, with *in vitro* reactions being necessary to fully elucidate the proposed mechanism, this mechanism, although probable by enzyme homology, has still not been indisputably verified.[47] For usage with natural substrates and potential industrial applications, CAR from *Nocardia iowensis* sp. NRRL 5646 therefore appears to be a more feasible choice. Still, for mechanistic research, GriC and GriD represent an interesting alternative enzyme system.

3. Experimental

3.1 Materials

DNA-based:

T4 DNA ligase and 10x ligation buffer, New England Biolabs

Xho I, Nde I Fast Digest restriction enzymes, Thermo Scientific

Fast Digest Green Buffer, Fermentas

Gene Ruler™ DNA ladder mix, 6x orange loading dye, Fermentas

Phusion™ PCR Kit, New England Biolabs

10x ultrapure TAE buffer, SYBR® Safe DNA gel stain, UltraPure agarose, Invitrogen

BL21 (DE3) chemically competent *E. coli* cells, Neb 5α chemically competent *E. coli* cells, New England Biolabs

XL10-Gold ultracompetent cells, XL10-Gold β-mercaptoethanol mix, Agilent Technologies

pET 21a(+), pET 28a(+), Novagen

pJET1.2/blunt cloning vector, Life Technologies

Gibson Assembly® master mix, New England Biolabs

Dream Taq master mix, Thermo Scientific

T7 promoter primer, T7 terminator primer, Thermo Scientific

QuikChange® II XL site-directed mutagenesis kit, QuikSolutions, Pfu Ultra HF DNA polymerase, Agilent Technologies

dNTPs, H_2O (nuclease free), Dpn I, Thermo Scientific

Genes purchased from Gene Art AG

Primers purchased from Eurofins Genomics

QIAprep® Spin Miniprep Kit, QIAquick® PCR Purification Kit, QIAquick® Gel Extraction Kit, QIAGEN

Protein-based:

HisTrap™ FF 5 mL column, PD-10 desalting column, GE Healthcare

PageRuler Prestrained Protein Ladder, Thermo Scientific

Acrylamide (30%), ammonium persulfate (APS), N,N,N´,N´ tetramethylethylethylenediamine (TEMED), Biorad protein assay, Biorad

6x DNA Loading Dye & SDS Solution, Fermentas

Isopropyl β-D-thiogalactopyranoside (IPTG), PEQLAB

Laemmli 2x sample buffer, Sigma Aldrich

SDS, D(+)-glucose, glycin, Tris, NaCl, ampicillin, Roth

Trypton, yeast extract, agar bacteriological, Oxoid

Nicotinamide adenine dinucleotide phosphate (NADPH) tetrasodium salt, AppliChem

Nicotinamide adenine dinucleotide phosphate ($NADP^+$) disodium salt, adenosine triphosphate (ATP) disodium salt, Roche

Imidazole, kanamycin sulfate, $Na_2HPO_4x2(H_2O)$, NaH_2PO_4, KH_2PO_4, K_2HPO_4, Sigma Aldrich

Glucose dehydrogenase (GDH) from *Bacillus megaterium*, X-zyme

Chemicals:

Triethylamine, 1-pentanol, trimethyl phosphate, benzoic acid, benzoyl chloride, *N*-acetylcysteamine, sodium iodide, β-butyrolactone, 2-azetidinone, magnesium chloride, EDTA, magnesium sulfate, ammonium chloride, Sigma Aldrich

4-(Dimethylamino)-pyridine, glyceraldehyde, calcium chloride, Fluka

Sodium sulfate, VWR International

Tetrahydrofuran, dithiothreitol, rubidium chloride, manganese(II) chloride tetrahydrate, Roth

Instruments:

Gas chromatograph

Agilent Technologies 7890 A equipped with an Agilent Technologies 7693 Autosampler

GC-column

GC: J&W DB-1701 Capillary GC column, Agilent Technologies,
length: 30 m, diameter: 0.25 mm, film: 0.25 µm

GC-MS: HP-5 19091S-433, Agilent Technologies,
length: 30 m, diameter: 0.25 mm film: 0.25 µm

Mass spectrometer

Agilent Technologies 5975C inert XL MSD

NMR

Bruker Ultrashield (300 MHz) spectrometer

Centrifuge

Thermo Scientific Sorvall RC 5C PLUS

FPLC

BIORAD BioLogic DuoFlow

Ultrasound

BRANSON Digital Sonifier

Photometers

Eppendorf BioPhotometer plus

Genesys 10UV Scanning

Safe Imager™ 2.0 Blue-Light Transilluminator

Autoclave

VARIOKLAV

3.2 Microbiological work

PCR:

PCR for the amplification of the *car* coding gene in *Nocardia iowensis* sp. NRRL 5646 was prepared with the Phusion™ PCR Kit and the according manual. The reaction mix is given in table 7:[60]

Table 7: PCR reaction mix

Component	Volume [µL]	Final concentration
5x Phusion HF Buffer	10	1
dNTP (10 mM)	1	200 µM (for each ATGC)
Forward primer (10 µM)	2.5	0.5 µM
Reverse primer (10 µM)	2.5	0.5 µM
DNA template	0.5	1 µg/mL
Phusion™ DNA Polymerase	0.5	0.02 U/mL
H_2O (nuclease free)	33	
	50	

The reaction was executed according to the program in table 8:

Table 8: PCR time program

PCR phase	Temperature [°C]	Time [s]	Cycles
Initial denaturation	98	30	1
Denaturation	98	10	
Annealing	59	30	3
Extension	100	100	
Final extension	72	600	1
	4	Hold	

The PCR product was further separated and isolated by a 1% agarose gel, followed by a gel extraction using the QIAGEN Kit.

PCR for the site-directed mutagenesis (SDM) in the *aar* coding gene of *Saccharomyces cerevisiae* was prepared according to table 9:[59]

Table 9: PCR reaction mix for SDM

Component	Volume [µL]	Final concentration
5x Phusion HF Buffer	10	1
dNTP (10 mM)	1	200 µM (for each ATGC)
Forward/reverse mutagenesis primer (10 µM)	2.5	0.5 µM
DNA template	0.5	
Phusion™ DNA Polymerase	0.5	0.02 U/mL
H_2O (nuclease free)	35.5	
	50	

The reaction was conducted separately for forward and reverse primer, respectively, according to the program in table 10 after which the reaction mixtures were merged into one tube and the second reaction was performed according to the program in table 11:

Table 10: PCR time program for SDM for forward and reverse primers separately (linear PCR)

PCR phase	Temperature [°C]	Time [s]	Cycles
Initial denaturation	98	30	1
Denaturation	98	15	
Annealing	55	30	3
Extension	72	1140[a]	
	4	Hold	

[a] 0.5 kb/min

Table 11: PCR time program for SDM with both primers combined (exponential PCR)

PCR phase	Temperature [°C]	Time [s]	Cycles
Initial denaturation	98	30	1
Denaturation	98	15	
Annealing	55	30	3
Extension	72	300[a]	
Final extension	72	600	1
	4	Hold	

[a] 2 kb/min

After digesting the parent DNA with Dpn I at 37 °C, 2 µL of the PCR product were used for transformation without further isolation.

Agarose gels:

1% Agarose gels were prepared by dissolving agarose (0.5 g) in TAE buffer (50 mL) under microwave heating. After cooling, the gel was poured into the form, SYBR® Safe DNA gel stain (5 µL) was added and an 8-slot-comb was fixed in the form. The samples were mixed accordingly with a 6x loading dye and Gene Ruler™ DNA ladder mix (5 µL; figure 8) was used as a standard.[61]

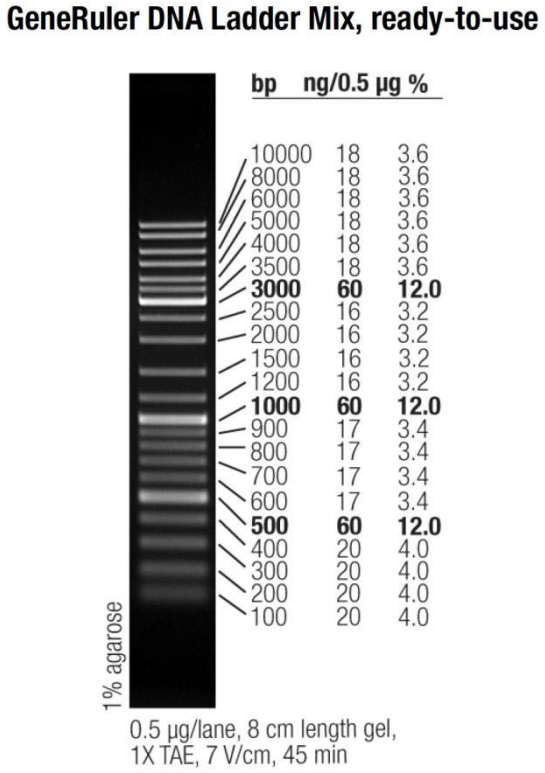


Figure 8: Reference bands and their indicative sizes of Gene Ruler™ DNA ladder mix

The gel was run in TAE buffer for 60 min at 100 mV and detected *via* a Safe Imager™ 2.0 Blue-Light Transilluminator at 470 nm.

Determination of DNA concentration:

DNA concentration was determined based on the absorption at 260 nm using an Eppendorf BioPhotometer plus.

DNA restriction:

The standard composition of the DNA restriction, shown in table 12, was varied according to DNA availability and concentration. The volumes of DNA sample, buffer, and H_2O were adapted in order to reach the smallest possible total volume.

Table 12: Example for a DNA restriction mix

Component	Volume [µL]
DNA	10
Xho I	2
Nde I	2
10x FD Green Buffer	2
H_2O (nuclease free)	4
	20

After restriction took place at 37 °C for 1 h, purification was achieved with an agarose gel followed by a gel extraction using the QIAGEN Kit. When low DNA concentration was used in the restriction, the gel extraction was sipped and only the PCR purification using the QIAGEN Kit was performed directly after restriction.

DNA ligation:

For ligation of vector DNA with the desired insert, vector DNA (70 ng) and a 1:5 and 1:3 vector:insert ratio were ideally used in the lowest possible volume. The composition of the DNA ligation mix varied depending on DNA concentration and availability. An exemplary ligation mix for a ligation of pET28a(+) vector (5.4 kb) with GriC insert (1.1 kb) is given in table 13:[62]

Table 13: Example for a DNA ligation mix

Component	Volume [µL] (1:5 ratio)	Volume [µL] (1:3 ratio)
DNA vector	6.5	6.5
DNA insert	17.0	10.2
10x T4 DNA ligase buffer	3.0	3.0
T4 DNA ligase	1.7	1.7
H_2O (nuclease free)	1.8	8.6
	30	30

The ligation was carried out at room temperature overnight, followed by heat inactivation at 65 °C for 10 min for dissociation of the ligase-DNA complex to ensure a more effective transformation.

pJET DNA ligation for amplification of insert DNA:

After several unsuccessful ligation attempts of the *griC* and *griD* genes, the remaining original blunt end genes were amplified to allow further ligations. For this, ligation of the *griC/D* genes into pJET1.2/blunt cloning vector was performed.[63] A gene:vector ratio of 3:1 or 0.15:0.05 pmol for blunt ends but a molecular ratio of 0.075:0.025 pmol (two blunt ends per molecule) was required, which was calculated using formula (1) based on the Gibson Assembly® protocol.

$$\text{m [ng]} = \frac{n\,[pmol] * gene\ length\,[bp] * 650\,[\frac{g}{mol}]}{1000} \qquad (1)$$

650 g/mol is therein assumed as the average molecular weight for a base pair. Using 0.075 pmol and 1080 bp and 1392 bp for GriC and GriD, respectively, and 0.025 pmol and 2974 bp for the pJET vector, gives the required amount of DNA:

m (GriC) = 52.65 ng

m (GriD) = 67,86 ng

m (pJET) = 48.33 ng

With c = 20 ng/µL for GriC and GriD, and c = 50 ng/µL for the pJET vector, the actual volumes to be used in the reaction were calculated, as seen in table 14:

Table 14: DNA ligation mix for GriC/D with pJET1.2/blunt cloning vector

Component	Volume [µL] (GriC)	Volume [µL] (GriD)
2x reaction buffer	10	10
Blunt end DNA fragment	2.6	3.4
pJET1.2/blunt cloning vector	1	1
H_2O (nuclease free)	5.4	4.6
T4 DNA ligase	1	1
	20	20

After mixing and spinning down for 5 s, the reaction took place for 5 min at 22 °C, followed by transformation.

Gibson Assembly:

For genes like the *car* coding gene from *N. iowensis* and the *aar* coding gene from *S. cerevisiae*, that are too big to be ordered in one piece, their fragments were ordered and assembled into the complete gene and into the plasmid vector *via* the Gibson Assembly® procedure.[64] Both genes were ordered in two parts, making it a 3 fragment assembly, which requires vector (50-100 ng), a 2-3 fold excess of insert and 0.02-0.5 pmol of fragments overall. By using formula (1) with vector (100 ng), the molar amounts of vector and fragments were calculated. The ordered gene fragments were concentrated in vacuum and then dissolved in an amount of sterile water that yielded the correct concentration in order to use 1 - 2 µL of the solution. The calculation for the *N. iowensis car* gene fragment 1 according to formula (1) is outlined below [100 ng of the pET28a(+) vector of 5369 bp length]:

$$\text{n (pET28)} = \frac{100\ ng * 1000}{5369\ bp * 650} = \underline{0.0287\ \text{pmol}}$$

Using a 3 fold excess of *N. iowensis car* gene fragment 1 (0.0860 pmol) and its size (1794 bp), gives

$$\text{m (}\textit{N. iowensis car}\text{ gene fragment 1)} = \frac{0.0860\ pmol * 1794\ bp * 650 \frac{g}{mol}}{1000} = 100.24\ \text{ng}$$

For the calculation of the necessary dilution, the total amount of the ordered gene fragments (1050 ng in this example) and the desired volume to be used in the reaction (1 µL for all fragments, 2 µL for the vector) are used in formula (2):

$$\text{V (dilution volume)} = \frac{1050\ ng * 1\ \mu L}{100.24\ ng} = 10.5\ \mu\text{L} \qquad (2)$$

The reaction mix for all fragments is shown in table 15:

Table 15: Gibson Assembly® reaction mix

Component	Volume [µL]
DNA fragment 1	1
DNA fragment 2	1
DNA vector	2
	4
H_2O (nuclease free)	6
Gibson Assembly® Master Mix	10
	16
Total volume	20

Due to the difficulty of pipetting such small amounts, the 3 DNA fragments as well as the Gibson Assembly® Master Mix with water were prepared separately, then the diluted Gibson Assembly® Master Mix was added to the DNA fragments. The combined reaction mix was then incubated at 50 °C for 1 h, put on ice briefly and used as a whole for the following transformation.

Transformations:

Neb 5α chemically competent *E. coli* cells were thawed on ice, then plasmid DNA (4 µL) was added to the cells (50 µL) and left on ice for 30 min. A heat shock for 30 s at 42 °C was applied and the cells were placed back on ice for 2 min. After the addition of LB-medium (250 µL), the tubes were incubated for 1 h at 37 °C at 300 rpm in a thermocycler that was turned onto its side. 50 µL and all the remaining volume, respectively, were spread on agar plates containing the appropriate antibiotic, which were then incubated overnight at 37 °C or for 48 to 72 h at room temperature.

The same protocol was applied when the transformation was conducted with BL21 (DE3) cells for protein expression and with self-made chemically competent cells already containing one target plasmid.

For transformations after Gibson Assembly® or site-directed mutagenesis, XL10-Gold ultracompetent cells were thawed on ice. Cells (45 µL) were pipetted into ice-chilled tubes, XL10-Gold β-mercaptoethanol mix (2 µL) was added to the cells and incubated on ice for 10 min. DNA sample (2 µL) was added to the cells and further incubated on ice for 30 min. After a heat shock at 42 °C for 30 s, LB-medium (900 µL) was added and the tubes were incubated for 1 h at 37 °C at 300 rpm in a thermocycler that was turned onto its side. 250 µL were spread on a agar plate containing the appropriate antibiotic and the remaining cells were pelleted by centrifugation at 14860 rpm for 30 s. The resuspended pellet was then spread on an agar plate containing the appropriate antibiotic and incubated overnight at 37 °C or for 48 to 72 h at room temperature.

Plasmid isolation:

Overnight cultures (ONC) were prepared from one colony after transformation in 7 mL LB-medium including the appropriate antibiotic [Ampicillin (100 µg/mL) and Kanamycin (30 µg/mL), respectively], and then used for plasmid isolation following the QIAGEN Miniprep Kit. For the sequencing of the whole length of the *N. iowensis car* gene and the *S. cerevisiae aar* gene, it was necessary to use 4 sequencing primers, each requiring 10 µL of isolated plasmid DNA.

Self-made chemically competent cells:

For simultaneous expression of two genes (*Nocardia iowensis car* + *Bacillus subtilis* PPTase, *griC* + *griD*), chemically competent cells already containing one plasmid, were required. For this, 0.5 mL of an ONC of BL21 (DE3) cells with one plasmid were added to LB-medium (50 mL), together with sterile $MgSO_4$ (0.5 mL; 1 M in water) and sterile $MgCl_2$ (0.5 mL; 1 M in water). The cells were grown at 37 °C and 120 rpm until an OD_{600} of 0.6-0.8 was reached. The cells were harvested by centrifugation at 5000 rpm for 15 min. For the following steps, the cells were kept on ice whenever possible. The supernatant was discarded and the pellet was resuspended in ice-cold, sterile TMF buffer (4 mL) and glycerol (1 mL). The aliquoted cells (100 µL) were shock-frozen in liquid nitrogen and stored at –80 °C.

Glycerol stocks:

In order to store cells ready for protein expression, glycerol stocks were prepared from ONCs with *Nocardia iowensis* CAR + *Bacillus subtilis* PPTase plasmids and Ampicillin (100 µg/mL) and Kanamycin (30 µg/mL) antibiotics after transformation of the plasmids into BL21 (DE3) cells. The ONCs (5.5 mL) were treated with glycerol (66%) for a final glycerol concentration of 15% and aliquots (1 mL) were shock-frozen in liquid nitrogen and stored at – 80 °C.

Media:

Lysogeny broth medium:

- 5 g/L NaCl
- 5 g/L Yeast extract
- 10 g/L Tryptone

Terrific broth medium:

- Autoclaved separately (900 mL):
    - 24 g/L Yeast extract
    - 12 g/L Tryptone
    - 4 mL/L Glycerol
- Autoclaved separately (100 mL):
    - 23.1 g/L KH_2PO_4
    - 125.4 g/L K_2HPO_4

The indicated concentrations refer to the final volume. The two parts of 900 mL and 100 mL, respectively, were prepared and autoclaved separately and then mixed for 1 L TB-medium.

For antibiotics and induction of protein expression, the following (final) concentrations were used:

- 100 µg/mL Ampicillin
- 30 µg/mL Kanamycin
- 1 mM IPTG

3.3 Protein expression

Induction and harvesting:

After transformation of the plasmids into BL21 (DE3) cells, an ONC (7 mL) of an isolated colony in LB-medium with the appropriate antibiotics [Ampicillin (100 µg/mL) and/or Kanamycin (30 µg/mL), respectively] was prepared and the ONC was used for inoculation of TB-medium (1.5 mL ONC for 300 mL medium) containing the same antibiotics [Ampicillin (100 µg/mL) and/or Kanamycin (30 µg/mL), respectively]. The culture was grown at 37 °C, 120 rpm until an OD of 0.6 - 0.8 was reached, then protein expression was induced by adding IPTG (1 mM). The culture was further incubated at 20 °C, 120 rpm overnight. To harvest the cells, the culture was centrifuged at 5000 rpm, 20 min, 4 °C, the cell pellet was resuspended in wash buffer (1 - 2.5 g cells/10 mL buffer), and then centrifuged again under the same conditions.

Protein expression for the aldehyde dehydrogenases in the preliminary test was performed by inoculating 200 mL LB-medium with 2 mL of the ONCs after transformation of the plasmids into BL21 cells. After incubation at 30 °C, 120 rpm, protein expression was induced by adding 1 mM IPTG (0.5 mM for Gox 5 and 6) at an OD_{600} of 0.7-1.0.

Cell disruption:

The resulting pellet was again resuspended in wash buffer (1 - 2.5 g cells/10 mL buffer) and the cells were disrupted using a BRANSON digital sonifier® with 1 s pulse at 30% amplitude, 4 s pause, and 3.5 min combined pulse duration. The supernatant (cell-free extract) was obtained after centrifugation at 15000 rpm, 20 min, 4 °C and careful separation from the pellet.

For the aldehyde dehydrogenases, 100 mg of the harvested cell pellet were resuspended in 700 µL 50 mM Na/K-phosphate buffer (pH 7.5) and disrupted using ultrasound with 0.1 s pulse at 40% amplitude, 0.4 s pause, and 15 s combined pulse duration. The procedure was executed twice under ice cooling. After centrifugation at 14650 rpm for 5 min, the cell-free extract was used in the photometric activity assay.

Protein purification:

The His-tagged proteins were purified using a HisTrap™ FF 5 mL column according to the manual:[65]

- Column washed with 5 column volumes (CV) of distilled water
- Equilibration of the column with 5 CVs of wash buffer
- Application of the cell-free extract
- Washing with wash buffer until absorbance reached steady baseline
- Elution with elution buffer using a linear gradient over 20 min
- Washing with 5 CVs of distilled water
- Washing with 5 CVs of 20% ethanol

A flow rate of 5 mL/min was used for all steps.

The protein-containing fractions according to the FPLC were analyzed *via* SDS-PAGE and the appropriate fractions were then combined and concentrated to a volume of approximately 2.5 mL. Before using the enzyme, it had to be desalted and the buffer system was changed to the respective reaction buffer which was going to be used in the reaction as well as for storage. Desalting was performed with a PD-10 desalting column according to the manual.[66] After the column was equilibrated with potassium phosphate buffer (50 mM, pH 9), the concentrated protein was applied to the column and eluted with the same buffer (3.5 mL). Protein concentration was calculated by measuring the absorption at 280 nm and using the Lambert-Beer law as depicted in formula (3)

$$c = \frac{A}{\varepsilon * d} \qquad (3)$$

with A = absorption at 280 nm, ε = extinction coefficient [L/mol*cm], and d = path length [= 1 cm]

Changing of the buffer and desalting was also carried out by dialysis instead of a desalting column by collecting the protein-containing fractions in a dialysis tube, which was then fixed in 5 L Tris buffer (50 mM, pH 7.5). Buffer exchange occurred at 4 °C overnight, followed by concentrating and determination of the concentration as above.

SDS-PAGE:

4 10% Gels were prepared according to the following recipe.

Separating gel:

- 5 mL 30% acrylamide
- 5.625 mL separating gel buffer (1 M Tris-HCl pH 8.8)
- 4.1 mL H_2O
- 150 µL 10% SDS solution
- 120 µL 10% APS
- 12 µL TEMED

The gel was filled into 4 Biorad gel cassettes and overlaid with isopropanol. After polymerization, the isopropanol was carefully removed and the stacking gel was added. The second polymerization then took place with a 10-slot-comb fixed in the separating gel.

Stacking gel:

- 0.833 mL 30% acrylamide
- 0.625 mL collecting gel buffer (1 M Tris-HCl pH 6.8)
- 3.45 mL H_2O
- 50 µL 10% SDS solution
- 25 µL 10% APS
- 5 µL TEMED

Before the application to the gel, the protein samples were treated with 2x sample buffer in the same amount as the sample and then heated for 10 min at 95 °C.

2x sample buffer:

- 1 mL 1 M Tris-HCl pH 6.8
- 400 mg SDS
- 300 mg DTT
- 20 mg bromphenol blue
- 2 mL glycerol
- 10 mL H_2O

The gel was run at 160 V for about 1 h and 5 µL of PageRuler Prestrained Protein Ladder was used as a standard (fig 9).[67]

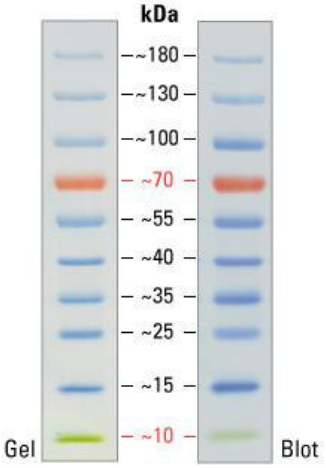


Figure 9: Reference bands and their indicative sizes of the PageRuler Prestrained Protein Ladder

For staining and destaining of the gels, the following solutions were used:

Commassie-Blue staining solution:

- 1 g Commassie Blue R250
- 450 mL H_2O
- 450 mL Methanol
- 100 mL glacial acetic acid

Destaining solution:

- 300 mL EtOH
- 100 mL acetic acid
- 600 mL H_2O

Buffers:

50 mM Na/K-Phosphate buffer, pH 7.5:

- 7.58 g/L $Na_2HPO_4x2H_2O$
- 1.01g/L KH_2PO_4

TMF buffer:

- 100 mM $CaCl_2$ (11.1 g/L)
- 50 mM RbCl (6.05 g/L)
- 40 mM $MnCl_2$ (7.92 $MnCl_2x4H_2O$ g/L)

Wash Buffer, pH 7.5, degassed after pH-adjustment:

- 2.40 g/L NaH_2PO_4 (20 mM)
- 29.2 g/L NaCl (500 mM)
- 1.36 g/L Imidazole (20 mM)

Elution Buffer, pH 7.5; degassed after pH-adjustment:

- 2.40 g/L NaH_2PO_4 (20 mM)
- 29.2 g/L NaCl (500 mM)
- 34.0 g/L Imidazole (500 mM)

Reaction buffer A, 100 mM K-Phosphate buffer, pH 7:

- 17.4 g/L K_2HPO_4
- 13.6 g/L KH_2PO_4
- 1.35 g/L NAD^+ (98% pure) (2 mM)
- 1.06 g/L benzaldehyde (**1b**; 10 mM) / 3.60 g/L glyceraldehyde (**8b**; 10 mM)

Reaction buffer B, 50 mM Tris buffer, pH 7.5:

- 6.06 g/L Tris
- 0.372 g/L EDTA disodium salt dihydrate (1 mM)
- 0.154 g/L DTT (1 mM)
- 10% v/v glycerol
- 0.167 g/L NADPH tetrasodium salt (0.2 mM)
- [0.952 g/L $MgCl_2$ (10 mM)]
- [0.605 g/L ATP disodium salt trihydrate (1 mM)]
- [0.611 g/L benzoic acid (**1c**; 5 mM)] ($MgCl_2$, ATP, and benzoic acid were only used in ATP-dependent reactions and were omitted in the experiments with the non-natural substrates)

Reaction buffer C, 50 mM Tris buffer, pH 7.5:

- 50 mM Tris
- 1 mM EDTA disodium salt dihydrate
- 1 mM DTT
- 10% v/v glycerol
- 15 mM NADPH tetrasodium salt
- [10 mM $MgCl_2$]
- [15 mM ATP disodium salt trihydrate]
- [10 mM benzoic acid, **1c**] ($MgCl_2$, ATP, and benzoic acid were only used in ATP-dependent reactions and were omitted in the experiments with the non-natural substrates)

Reaction buffer D, 50 mM Tris buffer:

- 50 mM Tris
- 1 mM $NADP^+$ disodium salt
- 30 mM glucose
- 0.2 mg/mL glucose dehydrogenase
- [10 mM $MgCl_2$]
- [15 mM ATP disodium salt trihydrate]
- [10 mM benzoic acid, **1c**] ($MgCl_2$, ATP, and benzoic acid were only used in ATP-dependent reactions and were omitted in the experiments with the non-natural substrates)

Reaction buffer D was also prepared with HEPES buffer and with K-phosphate at various pH instead of Tris. For the preparation with K-phosphate, 100 mM stocks were made, mixed in the appropriate amounts to adjust the pH and diluted to 20 mL with H_2O:

- 17.4 g/L K_2HPO_4 (100 mM stock)
- 13.6 g/L KH_2PO_4 (100 mM stock)

pH 6: 1.31 mL K_2HPO_4 (100 mM) + 8.68 mL KH_2PO_4 (100 mM)
pH 7: 6.15 mL K_2HPO_4 (100 mM) + 3.85 mL KH_2PO_4 (100 mM)
pH 8: 9.40 mL K_2HPO_4 (100 mM) + 0.60 mL KH_2PO_4 (100 mM)
pH 9: 9.93 mL K_2HPO_4 (100 mM) + 0.07 mL KH_2PO_4 (100 mM)

Reaction buffer E, 50 mM K-phosphate, pH 9:

- 1 mM DTT
- 1 mM $NADP^+$ disodium salt
- 30 mM glucose
- 0.2 mg/mL glucose dehydrogenase
- [10 mM $MgCl_2$]
- [15 mM ATP disodium salt trihydrate]
- [10 mM benzoic acid, **1c**] ($MgCl_2$, ATP, and benzoic acid were only used in ATP-dependent reactions and were omitted in the experiments with the non-natural substrates)

3.4 Substrate synthesis

Synthesis of benzoyl dimethyl phosphate (1a):

Step 1: Trimethyl phosphate (11.7 mL, 0.1 mol) and sodium iodide (15 g, 0.1 mol) were stirred for 4 days at room temperature until sodium dimethyl phosphate (11.7 g, 79%), precipitated as white crystals. The product was washed with ice-cooled acetone, and dried at 60 °C.

Step 2: Sodium dimethyl phosphate (1.0 g, 6.8 mmol) and 4-(dimethylamino)-pyridine (50 mg, 0.4 mmol) were suspended in dry tetrahydrofuran in a flask under argon atmosphere. Benzoyl chloride (0.8 mL, 6.9 mmol) was added and the reaction took place under stirring for 3 h at room temperature. After filtering the reaction mixture to remove sodium chloride, water (20 mL) was added. The aqueous phase was extracted with $CHCl_3$ (3x30 mL). The combined organic layers were washed with brine and dried with Na_2SO_4 overnight. After filtration and removal of the solvent in vacuum the crude product was purified *via* column chromatography (silica gel, petroleum ether:EtOAc 1.5:1; TLC: R_f = 0.07) which yielded the product (**1a**; 140 mg) as a colorless oil in ca. 25% purity.[68]

^{1}H NMR (300 MHz, $CDCl_3$) δ (ppm) = 7.3 (m, 5H) and 4.0 (d, 6H, *J* = 11.7 Hz).

^{31}P NMR (121 MHz, $CDCl_3$) δ (ppm) = - 5.1

Synthesis of benzoyl methyl phosphate (2a):

Benzoyl dimethyl phosphate (**1a**; 140 mg, 0.49 mmol) in ca. 25% purity and sodium iodide (92.5 mg, 0.62 mmol) were suspended in acetone (3 mL) and stirred for 4 days at room temperature, during which the product precipitated. Filtering and washing with ice-cooled acetone yielded the product, benzoyl methyl phosphate (**2a**; 72.3 g, 62%), as a white solid.[68]

^{1}H NMR (300 MHz, $CDCl_3$) δ (ppm) = 8.0 (d, 2H, *J* = 7.8 Hz), 7.6 (t, 1H, *J* = 7.5 Hz), 7.4 (t, 2H, *J* = 7.5 Hz), and 3.6 (d, 3H, *J* = 11.4 Hz).

^{13}C NMR (75 MHz, D_2O) δ (ppm) = 164.3, 134.5, 130.2, 128.8, 128.6, and 53.8.

^{31}P NMR (121 MHz, $CDCl_3$) δ (ppm) = - 5.3

Synthesis of benzoyl-*N*-acetylcysteamine (3a):

Benzoyl chloride (0.5 mL, 4.5 mmol) was added slowly to a solution of *N*-acetylcysteamine (288 µL, 3 mmol) and triethylamine (835 µL, 6 mmol) in 20 mL diethyl ether at 0 °C, which led to a white precipitate immediately. The reaction was quenched by adding saturated aqueous NH_4Cl (20 mL) and H_2O (40 mL), and extracted twice with ethyl acetate (80 mL each). The combined organic phases were concentrated *in vacuo*, which gave the solid product. In order to purify the product, it was filtered and washed with 5 mL ice-cooled ethyl acetate. After drying at 60 °C, the crude product was purified by column chromatography (silica gel, petroleum ether:EtOAc 1:10, TLC: R_f = 0.29) which yielded benzoyl-*N*-acetylcysteamine (**3a**; 0.54 g, 80%), as a white solid.[69]

^{1}H NMR (300 MHz, DMSO-d_6) δ (ppm) = 7.9 (d, 2H, *J* = 1.2 Hz), 7.7 (t, 1H, *J* = 7.5 Hz), 7.6 (t, 2H, *J* = 7.5 Hz), 3.3 (t, 2H, *J* = 6.3 Hz), 3.1 (t, 2H, *J* = 6.6 Hz), and 2.5 (s, 3H).

^{13}C NMR (75 MHz, DMSO-d_6) δ (ppm) = 191.5, 172.2, 136.8, 133.4, 128.5, 126.8, 38.8, 27.9, and 21.1.

3.5 Enzymatic reactions

Photometric activity assay:

Reaction buffer A or reaction buffer B (990 µL) with the respective substrate (10 mM **1b**, 10 mM **8b**, or 5 mM **1c**, respectively) were aliquoted into 1.5 mL plastic cuvettes. Immediately after adding 10 µL cell extract, the reaction mixture was stirred briefly and placed into the photometer. The change in absorption at 340 nm was measured over the course of 10 min.

Enzymatic screenings:

The respective reaction buffer (800 µL) was placed in 1.5 mL Eppendorf reaction tubes and the reaction was started by adding (20 µL; 100 µg) purified protein. The samples were incubated at 30 °C, 120 rpm for approximately 24 h before work-up.

Activity screenings in living/fermenting cells:

Biotransformations in living/fermenting cells were conducted by inoculating 10 mL LB-medium with a BL21 (DE3) colony containing the desired plasmids. After an OD_{600} of 0.6-0.8 was reached, protein expression was induced by IPTG (1 mM) and benzoic acid (**1c**, 10 mM) was added as a substrate. The reaction was performed overnight at 20 °C, 120 rpm for approximately 24 h before work-up.

Sample preparation for GC:

Aqueous HCl (100 µL; 2 M) was added to the sample and the reaction products were extracted with ethyl acetate (500 µL). 300 µL of the organic phase were diluted with additional ethyl acetate (100 µL) and subjected to GC analysis. In case of a following basic extraction, NaOH (ca. 200 µL, 10%) was added to the remaining sample until a pH of 10 was reached and the aqueous phase was extracted with ethyl acetate (300 µL) which was used fo GC analysis.

3.6 GC / GC-MS methods

For GC the following program was used (carrier gas: H_2, constant pressure: 0.3 bar):

Heating rate [°C/min]	Hold [°C]	Hold [min]
	100	1
20	170	4
30	250	1

Retention times: benzoic acid (**1c**; 7.9 min), benzaldehyde (**1b**; 5.7 min), hydroxyphenylacetic acid (**9a**; 11.2 min), hydroxyphenylacetaldehyde (**9b**; 9.4 min), phenylacetic acid (**10a**; 8.5 min), phenylacetaldehyde (**10b**; 6.6 min)

The response factor of benzaldehyde and benzoic acid was determined but insignificantly low and therefore not further considered.

For GC-MS the following program was used (carrier gas: He, flow: 0.7 mL/min):

Heating rate [°C/min]	Hold [°C]	Hold [min]
	40	2
20	180	1
20	300	4

4. References

[1] Kohlpaintner C, Schulte M, Falbe J, Lappe P, Weber J, Frey GD; *Ullmann's Encycl. Ind. Chem.* 2013, *2*, 285-316
[2] Kohlpaintner C, Schulte M, Falbe J, Lappe P, Weber J, Frey GD; *Ullmann's Encycl. Ind. Chem.* 2013, *2*, 317-324
[3] Hughes V, Kirshenbaum I; *Ind. Eng. Chem.* 1957, *49*, 1999-2003
[4] Harcken C, Brückner R; *Sci. Synth.* 2007, *25*, 65-136
[5] Corey EJ, Suggs JW; *Tetrahedron Lett.* 1975, *16*, 2647-2650
[6] Omura K, Swern D; *Tetrahedron*. 1978, *34*, 1651-1660
[7] Mancuso AJ, Huang S, Swern D; *J. Org. Chem.* 1978, *43*, 2480-2482
[8] Dess D, Martin J; *J. Org. Chem.* 1983, *48*, 4155-4156
[9] Semmelhack M, Chou CS, Cortes DA; *J. Am. Chem. Soc.* 1983, *105*, 4492-4494
[10] Addis D, Das S, Junge K, Beller M; *Angew. Chem., Int. Ed*. 2011, *50*, 6004-6011
[11] Cha JS, Lee KD, Kwon OO, Kim JM, Lee HS; *Bull. Korean Chem. Soc*. 1995, *16*, 561-565
[12] Burgstahler AW, Worden LR, Lewis TB; *J. Org. Chem*. 1963, *28*, 2918-2919
[13] Yokoyama T, Yamagata N; *Appl. Catal., A*. 2001, *221*, 227-239
[14] Farmer V, Henderson ME, Russell J; *Biochim. Biophys. Acta*. 1959, *35*, 202-211
[15] Gross G, Bolkart K, Zenk M; *Biochem. Biophys. Res. Commun.* 1968, *32*, 173-178
[16] Gross G, Zenk M; *Eur. J. Biochem.* 1969, *8*, 413-419
[17] Gross GG; *Eur. J. Biochem.* 1972, *31*, 585-592
[18] Kato N, Konishi H, Uda K, Shimao M, Sakazawa C; *Agric. Biol. Chem.* 1988, *52*, 1885-1886
[19] Kato N, Joung E, Yang H, Masuda M, Shimao M, Yanase H; *Agric. Biol. Chem.* 1991, *55*, 757-762
[20] Chen Y, Rosazza JP; *Appl. Environ. Microbiol.* 1994, *60*, 1292-1296
[21] Lamm AS, Khare A, Conville P, Lau PC, Bergeron H, Rosazza JP; *Int. J. Syst. Evol. Microbiol.* 2009, *59*, 2408-2414
[22] Bhattacharjee J; *CRC Crit. Rev. Microbiol.* 1985, *12*, 131-151
[23] Xu H, Andi B, Qian J, West AH, Cook PF; *Cell Biochem. Biophys.* 2006, *46*, 43-64
[24] Sinha AK, Bhattacharjee J; *Biochem. J.* 1971, *125*, 743-439
[25] Ohnishi Y, Furusho Y, Higashi T, Chun HK, Furihata K, Sakuda S, Horinouchu S; *J Antibiot.* 2004, *57*, 218-223
[26] Suzuki H, Ohnishi Y, Furusho Y, Sakuda S, Horinouchi S; *J. Biol. Chem.* 2006, *281*, 36944-36951
[27] Ohnishi Y, Horinouchi S; *Biofilms.* 2004, *1*, 319-328
[28] Suzuki H, Furusho Y, Higashi T, Ohnishi Y, Horinouchi S; *J. Biol. Chem.* 2006, *281*, 824-833
[29] Koketsu K, Watanabe K, Suda H, Oguri H, Oikawa H; *Nat. Chem. Biol.* 2010, *6*, 408-410
[30] Hastings JW, Nealson KH; *Annu. Rev. Microbiol*. 1977, *31*, 549-595
[31] Ferri SR, Meighen EA; *J. Biol. Chem.* 1991, *266*, 12852-12857
[32] Soly RR, Meighen EA; *J. Mol. Biol.* 1991, *219*, 69-77
[33] Meighen EA; *FASEB J.* 1993, *7*, 1016-1022
[34] Liu R, Zhu F, Lu L, Fu A, Lu J, Deng Z, Liu T; *Metab. Eng.* 2014, 22, 10-21
[35] Napora-Wijata K, Strohmeier GA, Winkler M; *Biotechnol. J.* 2014, *9*, 822-843
[36] Andexer JN, Richter M; *ChemBioChem.* 2015, *16*, 380-386
[37] Venkitasubramanian P, Daniels L, Das S, Lamm AS, Rosazza JP; *Enzyme Microb. Technol.* 2008, *42*, 130-137
[38] Li T, Rosazza JP; *J. Bacteriol.* 1997, *179*, 3482-3487
[39] Li T, Rosazza JP; *Appl. Environ. Microbiol.* 2000, *66*, 684-687
[40] Venkitasubramanian P, Daniels L, Rosazza JP; *J. Biol. Chem.* 2007, *282*, 478-485
[41] Schlumbohm W, Stein T, Ullrich C, Vater J, Krause M, Marahiel MA, Kruft V, Wittmann-Liebold B, *J. Biol. Chem.* 1991, *266*, 23135-23141
[42] Lambalot RH, Gehring AM, Flugel RS, Zuber P, LaCelle M, Marahiel MA, Reid R, Khosla C, Walsh CT; *Chem. Biol.* 1996, *3*, 923-936
[43] Sinha AK, Bhattacharjee JK; *Biochem. Biophys. Res. Commun.* 1970, *39*, 1205-1210
[44] Suvarna K, Seah L, Bhattacherjee V, Bhattacharjee JK, *Curr. Genet.* 1998, *33*, 268-275
[45] Pospiech A, Bietenhader J, Schupp T; *Microbiology.* 1996, *142*, 741-746
[46] Ehmann DE, Gehring AM; Walsh CT; *Biochemistry.* 1999, *38*, 6171-6177
[47] Suzuki H, Ohnishi Y, Horinouchi S; *J. Antibiot.* 2007, *60*, 380-387
[48] Prasad G, Borketey LS, Lin T, Schnarr NA; *Org. Biomol. Chem.* 2012, *10*, 6717-6723
[49] Amoroso JW, Borketey LS, Prasad G, Schnarr NA; *Org. Lett.* 2010, *12*, 2330-2333
[50] Prasad G, Amoroso JW, Borketey LS, Schnarr NA; *Org. Biomol. Chem.* 2012, *10*, 1992-2002
[51] Strope PK, Skelly DA, Kozmin SG, Mahadevan G, Stone EA, Maqwene P, Dietrich FS, McCusker JH; *Genome Res.* 2015, *25*, 762-774

[52] Barbe V, Vallenet D, Fonknechten N, Kreimeyer A, Oztas S, Labarre L, Cruveiller S, Robert C, Duprat S, Wincker P, Ornston LN, Weissenbach J, Marlière P, Cohen GN, Médique C; *Nucleic Acids Res.* 2004, *32*, 5766-5779
[53] Frese SA, Benson AK, Tannock GW, Loach DM, Kim J, Zhang M, Oh PL, Heng NC, Patil PB, Juge N, Mackenzie DA, Pearson BM, Lapidus A, Dalin E, Tice H, Goltsman E, Land M, Hauser L, Ivanova N, Kyrpides NC, Walter J; *PLoS Genet.* 2011, *7*, e1001314
[54] Jo J, Raj SM, Rathnasingh C, Selvakumar E, Jung W, Park S; *Appl. Microbiol. Biotechnol.* 2008, *81*, 51-60
[55] Schweiger P, Volland S, Deppenmeier U; *J. Mol. Microbiol. Biotechnol.* 2007, *13*, 147-155
[56] Hanke T, Richhardt J, Polen T, Sahm H, Bringer S, Bott M; *J. Biotechnol.* 2012, *157*, 359-372
[57] Faber K, *Biotransformations in organic chemistry: A textbook.* Springer-Verlag, Berlin, 2011
[58] Znidar D; *Laboratory project: Bioreduction of activated carboxylic acid derivatives with carboxylic acid reductase.* Karl-Franzens-University Graz, 2015
[59] Sabel J; *Ultramer™ oligonucleotides mutagenesis application guide. Integrated DNA Technologies*. 2011
[60] New England Biolabs; *PCR with phusion® polymerase (M0530).* 30.1.2015
[61] Thermo Fisher Scientific Inc; *PRODUCT INFORMATION thermo scientific GeneRuler DNA ladder mix, ready-to-use #SM0333.* 2012
[62] New England Biolabs; *Ligation protocol with T4 DNA ligase (M0202).* 25.2.2015
[63] Thermo Fisher Scientific Inc; *PRODUCT INFORMATION: Thermo ScientificCloneJET PCR CloningKit #K1231.* 2015
[64] New England Biolabs; *Gibson assembly® master mix – assembly (E2611).* 26.1.2015
[65] General Electric Company; *Instructions: HisTrap™ FF.* 2004
[66] General Electric Company; *Instructions: PD-10 desalting columns.* 2007
[67] Thermo Fisher Scientific Inc. *INSTRUCTIONS: PageRuler™ prestained protein ladder.* 2014
[68] Hikawa H, Hamada M, Yokoyama Y, Azumaya I; *RSC Adv.* 2014, *4*, 23131-23136
[69] Xie X, Watanabe K, Wojcicki WA, Wang CC, Tang Y; *Chem. Biol.* 2006, *12*, 1161-1169

5. Appendix

5.1 Gene sequences

Nocardia iowensis sp. NRRL 5646 CAR:
ATGGCAGTTGATAGTCCGGATGAACGTCTGCAGCGTCGTATTGCACAGCTGTTTGCAGAA
GATGAACAGGTTAAAGCAGCACGTCCGCTGGAAGCAGTTAGCGCAGCCGTTAGCGCACC
GGGTATGCGTCTGGCACAGATTGCAGCAACCGTTATGGCAGGTTATGCAGATCGTCCGGC
AGCAGGTCAGCGTGCATTTGAACTGAATACCGATGATGCAACCGGTCGTACCAGCCTGCG
TCTGCTGCCTCGTTTTGAAACCATTACCTATCGTGAACTGTGGCAGCGTGTTGGTGAAGTT
GCAGCAGCATGGCATCATGATCCGGAAAATCCGCTGCGTGCCGGTGATTTTGTTGCACTG
CTGGGTTTTACCAGCATTGATTATGCAACCCTGGATCTGGCAGATATTCATCTGGGTGCAG
TTACCGTTCCGCTGCAGGCAAGCGCAGCAGTGAGCCAGCTGATTGCAATTCTGACCGAAA
CCTCACCTCGTCTGCTGGCAAGCACACCGGAACATCTGGATGCAGCAGTTGAATGCCTGC
TGGCAGGTACAACTCCGGAACGTCTGGTTGTTTTTGATTATCATCCGGAAGATGATGATC
AGCGTGCCGCATTTGAAAGCGCACGTCGTCGTCTGGCCGATGCAGGTAGCCTGGTTATTG
TTGAAACACTGGATGCCGTTCGTGCACGTGGTCGTGATCTGCCTGCCGCACCGCTGTTTGT
TCCGGATACCGACGATGATCCGCTGGCACTGCTGATTTATACCAGTGGTAGCACCGGCAC
CCCGAAAGGTGCAATGTATACCAATCGTCTGGCAGCCACCATGTGGCAGGGTAATAGCAT
GCTGCAGGGCAATAGCCAGCGTGTGGGTATTAATCTGAATTATATGCCGATGAGCCATAT
TGCCGGTCGTATTAGCCTGTTTGGTGTTCTGGCACGTGGTGGCACCGCATATTTTGCAGCA
AAAAGCGATATGAGCACCCTGTTTGAAGATATTGGTCTGGTTCGTCCGACCGAAATCTTT
TTTGTTCCGCGTGTTTGTGATATGGTGTTTCAGCGTTATCAGAGCGAACTGGATCGTCGTA
GCGTTGCCGGTGCAGATCTGGATACGCTGGATCGTGAAGTTAAAGCCGATCTGCGTCAGA
ACTATCTGGGTGGTCGTTTTCTGGTTGCAGTTGTTGGTAGTGCACCGCTGGCAGCAGAAAT
GAAAACCTTTATGGAAAGCGTTCTGGATCTGCCGCTGCATGATGGTTATGGTAGTACCGA
AGCCGGTGCCAGCGTTCTGCTGGATAATCAGATTCAGCGTCCGCCTGTTCTGGATTATAA
ACTGGTTGATGTGCCGGAACTGGGTTATTTTCGTACCGATCGTCCGCATCCGCGTGGTGA
ACTGCTGCTGAAAGCAGAAACCACCATTCCGGGTTATTACAAACGTCCGGAAGTTACCGC
AGAAATCTTTGATGAAGATGGCTTCTATAAAACCGGTGATATTGTGGCAGAACTGGAACA
TGATCGTCTGGTTTATGTTGATCGTCGCAATAATGTTCTGAAACTGAGCCAGGGTGAATTT
GTTACCGTTGCACATCTGGAAGCCGTTTTTGCAAGCAGTCCGCTGATTCGTCAGATTTTTA
TCTATGGTAGCAGCGAACGTAGCTATCTGCTGGCCGTTATTGTTCCGACAGATGATGCAC
TGCGTGGTCGCGATACCGCAACCCTGAAAAGCGCACTGGCAGAAAGCATTCAGCGTATTG
CAAAAGATGCAAATCTGCAGCCGTATGAAATTCCGCGTGATTTTCTGATTGAAACCGAAC
CGTTTACCATTGCAAATGGTCTGCTGAGCGGTATTGCCAAACTGCTGCGTCCGAATCTGA
AAGAACGTTATGGTGCCCAGCTGGAACAAATGTATACAGATCTGGCCACCGGTCAGGCA
GATGAGCTGCTGGCCCTGCGTCGTGAAGCAGCGGATCTGCCAGTTCTGGAAACCGTTAGC
CGTGCAGCCAAAGCAATGCTGGGTGTTGCCAGCGCAGATATGCGTCCGGATGCACATTTT
ACCGATCTGGGAGGTGATAGCCTGAGCGCACTGAGCTTTAGCAATCTGCTGCATGAAATT
TTTGGTGTTGAAGTTCCGGTTGGTGTTGTTGTTAGTCCGGCAAATGAACTGCGTGATCTGG
CAAACTATATTGAAGCAGAACGTAATAGCGGTGCAAAACGTCCGACCTTTACCAGCGTTC
ATGGTGGTGGTAGCGAAATTCGTGCCGCAGATCTGACCCTGGATAAATTCATTGATGCAC
GTACCCTGGCAGCCGCAGATAGCATTCCGCATGCACCGGTTCCAGCACAGACCGTGCTGC
TGACAGGTGCCAATGGTTATCTGGGTCGCTTTCTGTGTCTGGAATGGCTGGAACGCCTGG
ACAAAACAGGTGGCACCCTGATTTGTGTTGTTCGTGGTAGTGATGCCGCAGCAGCCCGTA
AACGTCTGGATAGCGCCTTTGATAGCGGTGATCCGGGTCTGCTGGAACATTATCAGCAGC
TGGCAGCGCGTACACTGGAAGTTCTGGCAGGCGATATTGGTGATCCGAACCTGGGTCTGG
ATGATGCCACCTGGCAGCGTCTGGCGGAAACCGTGGATCTGATTGTTCATCCTGCAGCAC
TGGTTAATCATGTTCTGCCGTATACCCAGCTGTTCGGTCCGAATGTTGTGGGTACAGCCGA
AATTGTTCGTCTGGCGATTACCGCACGTCGCAAACCGGTTACCTATCTGAGCACCGTTGGT
GTGGCAGATCAGGTTGATCCGGCAGAATATCAAGAGGATAGTGATGTGCGTGAAATGAG
CGCAGTTCGTGTGGTTCGTGAAAGCTATGCAAATGGCTATGGTAATAGTAAATGGGCAGG
CGAAGTTCTGCTGCGTGAAGCACATGATCTGTGTGGTCTGCCGGTTGCCGTTTTTCGTAGT
GATATGATTCTGGCACATAGCCGTTATGCCGGTCAGCTGAATGTTCAGGATGTTTTTACCC

GTCTGATTCTGAGCCTGGTTGCCACCGGTATTGCACCGTATAGCTTTTATCGCACCGATGC
AGATGGTAATCGTCAACGTGCACATTATGATGGTCTGCCTGCGGATTTTACCGCAGCCGC
AATTACAGCACTGGGTATTCAGGCAACCGAAGGTTTTCGCACCTATGATGTGCTGAATCC
GTATGATGATGGTATCAGCCTGGATGAATTTGTGGATTGGCTGGTTGAAAGCGGTCATCC
GATTCAGCGCATTACCGATTATAGCGATTGGTTTCACCGCTTTGAAACCGCAATTCGTGCA
CTGCCTGAAAAACAGCGTCAGGCCAGCGTGCTGCCTCTGCTGGACGCCTATCGTAATCCG
TGTCCGGCAGTGCGTGGTGCCATTCTGCCTGCAAAAGAATTCCAGGCAGCCGTTCAGACC
GCAAAAATTGGTCCGGAACAGGATATTCCGCATCTGTCAGCACCGCTGATTGATAAATAT
GTTAGCGATCTGGAACTGCTGCAGCTGCTGTAA

Saccharomyces cerevisiae AAR (Lys2):
ATGACCAACGAAAAAGTGTGGATCGAGAAACTGGATAATCCGACCCTGAGCGTTCTGCC
GCATGATTTTCTGCGTCCGCAGCAAGAACCGTATACCAAACAGGCAACCTATAGCCTGCA
GCTGCCTCAGCTGGATGTTCCGCATGATAGCTTTAGCAACAAATATGCAGTTGCACTGAG
CGTTTGGGCAGCACTGATTTATCGTGTTACCGGTGATGATGATATCGTGCTGTATATTGCC
AACAACAAAATCCTGCGCTTTAACATTCAGCCGACCTGGTCATTTAATGAACTGTATAGC
ACCATCAACAACGAACTGAATAAACTGAACAGCATCGAAGCCAACTTTAGCTTTGATGAA
CTGGCCGAAAAAATTCAGAGCTGTCAGGATCTGGAACGTACACCGCAGCTGTTTCGTCTG
GCATTTCTGGAAAATCAGGATTTCAAACTGGACGAGTTTAAACATCACCTGGTTGATTTT
GCACTGAATCTGGATACCAGCAATAATGCCCATGTTCTGAACCTGATTTATAACAGCCTG
CTGTATAGTAATGAACGCGTTACCATTGTTGCAGATCAGTTTACCCAGTATCTGACCGCA
GCCCTGAGCGATCCGAGCAATTGTATTACCAAAATTAGCCTGATTACCGCCAGCAGCAAA
GATAGCCTGCCGGATCCGACCAAAAATCTGGGTTGGTGTGATTTTGTGGGTTGCATCCAT
GATATCTTTCAGGATAATGCAGAAGCATTTCCGGAACGTACCTGTGTTGTTGAAACCCCG
ACACTGAATAGCGATAAAAGCCGTAGCTTTACCTATCGTGATATTAATCGTACCAGCAAC
ATTGTGGCCCACTATCTGATTAAAACCGGTATTAAACGTGGTGACGTGGTGATGATTTAT
AGCAGTCGTGGTGTTGATCTGATGGTTTGTGTTATGGGTGTTCTGAAAGCCGGTGCAACCT
TTAGCGTTATTGATCCGGCATATCCGCCTGCACGTCAGACCATTTATCTGGGTGTTGCAAA
ACCGCGTGGTCTGATTGTTATTCGTGCAGCAGGCCAGCTGGATCAACTGGTTGAAGATTA
TATCAATGACGAGCTGGAAATTGTGTCCCGCATTAATAGCATTGCCATTCAAGAAAATGG
CACCATTGAAGGTGGTAAACTGGACAATGGTGAAGATGTTCTGGCACCGTATGATCACTA
TAAAGATACCCGTACCGGTGTGGTTGTTGGTCCGGATAGCAATCCGACACTGTCATTTAC
CAGCGGTAGCGAAGGTATTCCGAAAGGTGTTCTGGGTCGTCATTTTAGCCTGGCCTATTA
TTTCAATTGGATGAGCAAACGCTTTAATCTGACCGAGAATGACAAATTTACCATGCTGAG
CGGTATTGCACATGATCCGATTCAGCGTGATATGTTTACACCGCTGTTTCTGGGTGCACAG
CTGTATGTGCCGACCCAGGATGATATTGGTACACCGGGTCGTCTGGCCGAATGGATGAGT
AAATATGGTTGTACCGTTACCCATCTGACACCGGCAATGGGCCAGCTGCTGACCGCACAG
GCCACCACCCCGTTTCCGAAACTGCATCATGCATTTTTTGTTGGTGATATCCTGACCAAAC
GTGATTGCCTGCGCCTGCAGACCCTGGCAGAAAATTGTCGTATTGTTAATATGTATGGCA
CCACCGAAACCCAGCGTGCAGTTAGCTATTTTGAAGTGAAAAGCAAAAACGATGACCCG
AACTTCCTGAAAAAACTGAAAGATGTTATGCCTGCCGGTAAAGGTATGCTGAATGTGCAG
CTGCTGGTTGTTAATCGTAATGATCGTACACAGATTTGCGGCATTGGTGAAATTGGCGAA
ATTTATGTTCGTGCCGGTGGTCTGGCGGAAGGTTATCGCGGTCTGCCGGAACTGAACAAA
GAAAAATTTGTGAACAACTGGTTTGTGGAAAAAGATCACTGGAACTATCTGGATAAAGAT
AATGGTGAACCGTGGCGTCAGTTTTGGCTGGGTCCGCGTGATCGTCTGTATCGCACCGGT
GATCTGGGACGTTATCTGCCGAATGGTGATTGTGAATGTTGTGGTCGTGCAGATGATCAG
GTTAAAATTCGTGGTTTTCGTATTGAACTGGGCGAGATTGATACCCATATTTCACAGCATC
CGCTGGTTCGTGAAAACATTACCCTGGTGCGTAAAAATGCAGATAATGAACCGACGCTGA
TCACCTTTATGGTTCCGCGTTTTGATAAACCGGATGATCTGAGCAAATTTCAGAGTGATGT
TCCGAAAGAAGTTGAAACCGATCCGATTGTTAAAGGCCTGATTGGTTATCATCTGCTGAG
TAAAGATATTCGCACCTTTCTGAAAAAACGTCTGGCAAGCTATGCAATGCCGAGCCTGAT
TGTTGTTATGGATAAACTGCCGCTGAATCCGAATGGTAAAGTGGATAAACCTAAACTGCA
ATTTCCGACCCCTAAACAGCTGAACCTGGTTGCAGAAAATACCGTTAGCGAAACCGATGA
TAGCCAGTTTACAAATGTTGAACGTGAAGTTCGTGATCTGTGGCTGAGTATTCTGCCGAC
CAAACCGGCAAGCGTTAGTCCGGATGATAGTTTTTTTGATCTGGGTGGTCATAGCATTCTG
GCCACCAAAATGATTTTTACACTGAAAAAAAAACTGCAGGTCGATCTGCCGCTGGGCACC

ATTTTCAAATATCCGACCATCAAAGCATTTGCAGCCGAAATTGATCGCATTAAAAGCAGC
GGTGGTAGCAGCCAGGGTGAAGTGGTTGAAAATGTTACCGCAAATTATGCCGAGGATGC
CAAAAAACTGGTGGAAACCCTGCCGAGCAGCTATCCGAGCCGTGAATATTTTGTTGAACC
GAATAGCGCAGAAGGCAAAACCACCATTAATGTTTTTGTGACCGGTGTGACAGGTTTTCT
GGGTAGCTATATCCTGGCCGATCTGCTGGGTCGTAGCCCGAAAAACTATAGCTTTAAAGT
TTTTGCCCATGTGCGTGCCAAAGATGAAGAAGCAGCATTTGCCCGTCTGCAGAAAGCAGG
TATTACCTATGGCACCTGGAATGAAAAATTCGCGAGCAACATTAAAGTTGTGCTGGGTGA
TCTGAGTAAAAGCCAGTTTGGTCTGAGTGATGAAAAATGGATGGATCTGGCCAATACCGT
GGATATTATCATTCATAATGGTGCGCTGGTTCATTGGGTTTATCCGTATGCCAAACTGCGT
GATCCGAACGTTATTAGCACAATTAATGTTATGAGCCTGGCAGCAGTTGGTAAACCGAAA
TTTTTCGATTTTGTTAGCAGCACCAGCACCCTGGATACCGAATATTACTTTAACCTGTCAG
ATAAACTGGTGAGCGAAGGCAAACCGGGTATTCTGGAAAGCGACGATCTGATGAATAGC
GCCAGCGGTCTGACCGGTGGTTATGGTCAGAGCAAATGGGCAGCAGAATACATTATTCGT
CGTGCGGGTGAACGTGGTCTGCGTGGTTGTATTGTTCGTCCGGGTTATGTTACAGGTGCA
AGCGCAAATGGTAGCAGTAATACCGATGATTTCCTGCTGCGTTTTCTGAAAGGTAGCGTT
CAGCTGGGTAAAATTCCGGATATCGAAAATAGCGTTAATATGGTGCCGGTTGATCATGTT
GCACGTGTTGTTGTTGCAACCAGCCTGAATCCGCCTAAAGAAAACGAACTGGCAGTTGCC
CAGGTTACCGGTCATCCGCGTATTCTGTTTAAAGATTATCTGTATACCCTGCACGATTATG
GCTATGATGTTGAAATCGAGAGCTATAGCAAATGGAAAAAAAGCCTGGAAGCAAGCGTG
ATTGATCGTAACGAAGAAAATGCACTGTATCCGCTGCTGCACATGGTTCTGGATAACCTG
CCTGAAAGCACCAAAGCACCGGAACTGGATGATCGTAATGCCGTTGCAAGCCTGAAAAA
AGATACCGCATGGACCGGTGTTGATTGGAGCAATGGTATTGGTGTTACACCGGAAGAAGT
TGGCATTTATATCGCGTTTCTGAATAAAGTGGGTTTTCTGCCTCCTCCGACCCATAATGAC
AAACTGCCTCTGCCGTCAATTGAACTGACCCAGGCACAGATTAGCCTGGTTGCCAGCGGT
GCCGGTGCACGTGGTAGCTCAGCAGCAGCATAA

Streptomyces griseus GriC:
ATGAGCCTGATTGATGATCCGGAATTTGATAGCCATGTTGTTAAACTGATGGCCTGGCAT
TTTGGTGCAGAAACCGGTAGCCCGTTTTGGCTGGGTAAACTGGCAGGTCTGGGTTTTGAT
CCGGTTACCGATGTTCGTGGTCTGGCAGATCTGACCCGTTTTCCGGATGTTAGCGCAGAA
CTGCGTAGCGTTCCGGCACAGGATCTGATTCCGCGTGGTCTGAGCGATCGTCCGTTTCGTG
TTTATGATAGCGGTGGCACCACCGGTAGTCCGAAACGTATTGTGGATAGCGGTTATCGTA
GCCTGGTTACCGAATGGGCACGTGAACGTCTGATTGCAAATGGTGTTCCGGCAGATGGTA
ATTGGCTGCATCTGGGTCCGGGTGGTCCGCATGTTATTGGTTTTGATACCGCACGTTATGC
AGCACTGGGTGGTGGTGTGTTTTATACCGTTGATCTGGATCCGCGTTGGGTTAAACGTCTG
CTGAGCCAGGGTCGTGGTGAAGAAAGCGAAGCCTATGTTCAGCATCTGCTGGATCAGGC
AGAAACCATTCTGGAAAGCCAAGAAATTGCAGTTCTGAATACCACCCCTCCGCTGCTGGA
AGCAATTTGTGCACGTCCGCGTCTGTATGAACTGGTTCGTAGCGGTGTTCGTGCAATTATT
TGGGCAGGCACCAGCATTAGCCGTGAAACCCTGACCCAGCTGGATGATGTTTTTTTTCCG
GCAGCAAGCGTTGTTGGTGTTTATGGTAATAGCCTGATGGGTGTTGCACCGCAGCGTAGC
GGTCGTCCGGGTGATGCACATCGTTGTGTTTTTGAACCGTATCCGGAAACCACCCGTCTGC
ACCTGGTTGATGAAGATGGTACACCGGTTGCCTATGGTGAACGTGGTCGTGTTCGTCTGC
ATCTGGTGAGCGAAGAAATGTTTCTGCCGAATGTTCTGGAACGTGATACCGCCATTCGTG
TTGAACCGGGTGAAGGTAGCACCGTTGATGGTCTGGCCGATGTTCAGACCTATCGTAGTC
TGGATGACGTTGCCATTATTGAAGGTGTGTATTAA

Streptomyces griseus GriD:
ATGAGCACCGTTGCACCGCCTGCAGTTGTTCTGGATCCGCGTGTTCGTGGTCGTCTGGTTG
CAAGCAGCGATCGTACCCGTCTGCCGAGCGTTCTGGGTGGTGATCTGGCAGATGTTGGCA
CCGCACCGCGTCTGCTGGCAGTTGCAGCACTGAATGAAATTCGTGCACATGCAGATGGTC
GTCCGCCTGCAAGCGAAGTTTTTACCGAAGCAGCACGTCTGTTTGCAACCGCAACCCTGG
ATGGTGAAACACCGCGTGAATATGCCCGTCGTGTTTGTCATGCAACCGGTCTGACCGCAA
CCGCAGTTGAACAGGCAGTTACCGATCTGACCGGTGAACTGCGTGCACTGCCTGCAACCA
CCGCAGCAGAACTGCCTGCCACCGGTTTTGGTGATGGTTTTGATACCCGTTGGGTTCCGCG
TGGTCGTACCTTTGCAGCAGTTATGGCAAGCAATCATCCGGTTCCGAATATTAGCTGGGC
ACAGGCACTGTTTCATGGTTATAGCGTGCTGGTTAAACCGGGTAGCCGTGATCCGTTTAC

ACCGGCACGCCTGCTGGCTGCACTGACCGCTGCAGGTCTGCCTCCGGATCGTGCAGCATT
TCTGCCGTGCAGCCGTGAAGTTGGTGCATATCTGCTGCGTGAAGCAGATCGTGGTATTGT
TTATGGTGGTGATAGCGCAGTTGCAACCTGGCATCAGCGTGAAAGCGTTGCAGTGCGTGG
TCCGGGTCGTACCAAAGCCTTTCTGGATCGTGCTCCGGATGATGCAGTTATTGATCATCTG
GCACTGAGCGCAAGCTTTGATGGTGGCACCCGTTGTACCAATCTGAGCGCAGTTCTGACC
ACCGGTCCGGTTGAAGAAGTTGCAGATCGTCTGGCCGAACGTCTGGCACGTCTGCCGTCA
CTGCCAGCGACCGATGAAGCAGCAACCCTGCTGGTTGTTAATCGTGCCCGTGCCGATGGT
CTGCGTGAACAGGTTGCCGCACTGCGTGCCAGCCTGACCGATCATAGCGCACGTGCAGGC
GATCCGGAAACCGTTGTTGAACTGCCGGATGGTAGCTTTCTGCCTCGTCCGGTTGTTCTGA
GTGCCGATCGTGCGGATCATCCTGCCGTTGGTACGGAACTGCCGTTTCCGTTTGTTGTTGT
TGCACCGTGGGCTGAAGCCGATGGTGTTGAACCGCTGCGTGATAGCCTGGTTCTGAATCT
GCTGACAGATCGTGAAGAACTGGTTGAAGCCGCAGTTCGTGAACCGAGCGTGCGTAAAG
TTACCCGTGGTCCTGTGCTGCCGTGGACCGCAACACCGGGTATTCCGCATGATGATAACT
ATACCCAGTTTCTGCTGGAACCGAAAGGTGTTGTTGCGCGTCGTTAA

Bacillus subtilis PPTase:
ATGAAAATCTACGGCATCTATATGGATCGTCCGCTGAGCCAAGAAGAAAACGAACGTTTT
ATGAGCTTTATCAGTCCGGAAAAACGTGAAAAATGCCGTCGCTTTTATCACAAAGAAGAT
GCACATCGTACCCTGCTGGGTGATGTTCTGGTTCGTAGCGTTATTAGCCGTCAGTATCAGC
TGGATAAATCCGATATTCGTTTTAGCACCCAAGAATATGGTAAACCGTGTATTCCGGATC
TGCCGGATGCACATTTTAACATTAGCCATAGCGGTCGTTGGGTTATTTGTGCATTTGATAG
CCAGCCGATTGGCATTGATATCGAAAAAACCAAACCGATCAGCCTGGAAATTGCCAAAC
GTTTTTTTAGCAAAACCGAGTATAGCGATCTGCTGGCCAAAGATAAAGATGAACAGACCG
ATTATTTCTATCATCTGTGGTCCATGAAAGAGAGCTTCATTAAACAAGAAGGTAAAGGTC
TGAGCCTGCCGCTGGATAGCTTTAGCGTTCGTCTGCATCAGGATGGTCAGGTTAGCATTG
AACTGCCGGATAGCCATAGTCCGTGTTATATCAAAACCTATGAAGTTGACCCTGGCTACA
AAATGGCAGTTTGTGCAGCACATCCGGATTTTCCGGAAGATATTACCATGGTTAGCTATG
AAGAACTGCTGTAA

Saccharomyces cerevisiae PPTase (Lys5):
ATGGTGAAAACCACCGAAGTTGTTAGCGAAGTTAGCAAAGTTGCCGGTGTTCGTCCGTGG
GCAGGTATTTTTGTTGTTGAAATTCAAGAGGATATTCTGGCCGATGAATTTACCTTTGAAG
CACTGATGCGTACCCTGCCGCTGGCAAGCCAGGCACGTATTCTGAACAAAAAAAGCTTTC
ATGATCGCTGCAGCAATCTGTGTAGCCAGCTGCTGCAACTGTTTGGTTGTAGCATTGTTAC
CGGTCTGAATTTCCAAGAGCTGAAATTTGATAAAGGCAGCTTTGGTAAACCGTTCCTGGA
TAATAATCGTTTTCTGCCGTTTAGCATGACCATTGGTGAACAGTATGTTGCCATGTTTCTG
GTTAAATGTGTTAGCACCGATGAGTATCAGGATGTGGGTATTGATATTGCAAGCCCGTGT
AATTATGGTGGTCGTGAAGAACTGGAACTGTTTAAAGAAGTGTTTAGCGAACGCGAATTT
AACGGTCTGCTGAAAGCAAGCGATCCGTGTACCATTTTTACCTATCTGTGGTCACTGAAA
GAGAGCTATACCAAATTTACCGGTACAGGTCTGAATACCGATCTGAGCCTGATTGATTTT
GGTGCCATTAGCTTTTTTCCAGCCGAAGGTGCAAGCATGTGTATTACCCTGGATGAAGTTC
CGCTGATTTTTCATAGCCAGTGGTTTAACAATGAGATCGTGACCATTTGTATGCCGAAAA
GCATTAGCGATAAAATCAATACCAATCGTCCGAAACTGTACAACATTAGCCTGAGCACCC
TGATCGATTATTTCATTGAAAATGATGGCCTGTAA

Acinetobacter sp. BALD:
ATGACCCTGCTGGATGCGAGCATTTGGAACAAAAAACTGTTTAACGGCGGCTGGTTTGAA
AGCGGCCAGCCGTATGGCGTGGTGGAAGTGGCGACCGGCGAACAGCTGGGCCAGACCGG
CAGCGCGAGCCCGACCGATGTGGCGCAGGCGGCGAAAGAAGCGCAGACCGCGCAGCGCC
AGTGGTGGGCGCTGGATTATCTGGAACGCCAGGCGGTGTTTGAAAAAGCGGTGCAGATT
GCGACCGAACATCAGGATGAACTGGTGAACTGGCTGATTCGCGAAAGCGGCAGCCTGGC
GCTGAAAGCGCAGTTTGAAGTGAAAGTGAGCATTCAGATGCTGAAAAACTGCATTGCGTT
TCCGCAGCTGGATCAGGGCAACATTCTGCCGAGCCGCAACGGCAAACTGAGCCTGGCGA
AACGCCTGCCGCTGGGCGTGGTGGGCGTGATTAGCCCGTTTAACTTTCCGCTGTATCTGGC
GCTGCGCGCGGTGGGCCCGGCGCTGGCGTTTGGCAACGCGGTGGTGCTGAAACCGGATG
AACGCACCGCGGTGTGCAGCGGCTATGCGATTGCGCGCATTTTTGAACTGGCGGGCCTGC

CGAAAGGCCTGCTGCATGTGCTGCCGGGCGGCGCGGATACCGGCGAAGCGCTGACCCTG
GATCCGCATATTGCGAGCATTCAGTTTACCGGCAGCACCCAGGTGGGCCGCATTATTGGC
GCGAACGCGGGCAAAACCCTGAAAAAAGTGAGCCTGGAACTGGGCGGCAAAAACAGCCT
GATTATTCTGGATGATGCGGATCTGGATCTGGCGGCGGAAAACATTGCGTGGGGCGCGTT
TCTGCATAGCGGCCAGATTTGCATGACCAGCGGCAAAATTCTGATTCATGAAAAAATTTA
TGATGCGCTGAAAGCGCGCGTGATTGAAAAAGTGAAACGCTTTGTGGTGGGCAACCCGCT
GGAACAGAACGTGACCATTGGCCCGCTGATTAACGAAAAACAGAGCAAACGCGTGGAAG
CGCTGGTGCAGGCGGCGGTGGAACATGGCGCGACCCTGGAAATTGGCGGCAAAGCGAAC
GGCCCGTTTTTTGAACCGAGCGTGCTGAGCCAGGTGCAGGCGGATAACCCGATTTTTAGC
GAAGAAATTTTTGGCCCGGTGGCGGTGCTGATTCCGTTTGCGAGCGATGAACAGGCGATT
GAACTGGCGAACATGGGCGATTATGGCCTGAGCGCGGGCATTATTAGCAGCAACGTGGG
CCGCGCGATGCAGCTGGGCGCGCAGCTGAACGTGGGCCTGCTGCATATTAACGATCAGAC
CGTGAACGATGAAACCATTAACCCGTTTGGCGGCTTTGGCGCGAGCGGCAACGCGACCCG
CATTGGCGGCCCGGCGAACCCGGATGAATTTACCCAGTGGCAGTGGATGACCATTCAGGC
GGAAGCGCCGCATTATCCGTTTAGCGCGTGGAGCCATCCGCAGTTTGAAAAA

Saccharomyces cerevisiae ALDH4:
ATGTTTAGCCGCAGCACCCTGTGCCTGAAAACCAGCGCGAGCAGCATTGGCCGCCTGCAG
CTGCGCTATTTTAGCCATCTGCCGATGACCGTGCCGATTAAACTGCCGAACGGCCTGGAA
TATGAACAGCCGACCGGCCTGTTTATTAACAACAAATTTGTGCCGAGCAAACAGAACAAA
ACCTTTGAAGTGATTAACCCGAGCACCGAAGAAGAAATTTGCCATATTTATGAAGGCCGC
GAAGATGATGTGGAAGAAGCGGTGCAGGCGGCGGATCGCGCGTTTAGCAACGGCAGCTG
GAACGGCATTGATCCGATTGATCGCGGCAAAGCGCTGTATCGCCTGGCGGAACTGATTGA
ACAGGATAAAGATGTGATTGCGAGCATTGAAACCCTGGATAACGGCAAAGCGATTAGCA
GCAGCCGCGGCGATGTGGATCTGGTGATTAACTATCTGAAAAGCAGCGCGGGCTTTGCGG
ATAAAATTGATGGCCGCATGATTGATACCGGCCGCACCCATTTTAGCTATACCAAACGCC
AGCCGCTGGGCGTGTGCGGCCAGATTATTCCGTGGAACTTTCCGCTGCTGATGTGGGCGT
GGAAAATTGCGCCGGCGCTGGTGACCGGCAACACCGTGGTGCTGAAAACCGCGGAAAGC
ACCCCGCTGAGCGCGCTGTATGTGAGCAAATATATTCCGCAGGCGGGCATTCCGCCGGGC
GTGATTAACATTGTGAGCGGCTTTGGCAAAATTGTGGGCGAAGCGATTACCAACCATCCG
AAAATTAAAAAAGTGGCGTTTACCGGCAGCACCGCGACCGGCCGCCATATTTATCAGAGC
GCGGCGGCGGGCCTGAAAAAAGTGACCCTGGAACTGGGCGGCAAAAGCCCGAACATTGT
GTTTGCGGATGCGGAACTGAAAAAAGCGGTGCAGAACATTATTCTGGGCATTTATTATAA
CAGCGGCGAAGTGTGCTGCGCGGGCAGCCGCGTGTATGTGGAAGAAAGCATTTATGATA
AATTTATTGAAGAATTTAAAGCGGCGAGCGAAAGCATTAAAGTGGGCGATCCGTTTGATG
AAAGCACCTTTCAGGGCGCGCAGACCAGCCAGATGCAGCTGAACAAAATTCTGAAATAT
GTGGATATTGGCAAAAACGAAGGCGCGACCCTGATTACCGGCGGCGAACGCCTGGGCAG
CAAAGGCTATTTTATTAAACCGACCGTGTTTGGCGATGTGAAAGAAGATATGCGCATTGT
GAAAGAAGAAATTTTTGGCCCGGTGGTGACCGTGACCAAATTTAAAAGCGCGGATGAAG
TGATTAACATGGCGAACGATAGCGAATATGGCCTGGCGGCGGGCATTCATACCAGCAAC
ATTAACACCGCGCTGAAAGTGGCGGATCGCGTGAACGCGGGCACCGTGTGGATTAACAC
CTATAACGATTTTCATCATGCGGTGCCGTTTGGCGGCTTTAACGCGAGCGGCCTGGGCCG
CGAAATGAGCGTGGATGCGCTGCAGAACTATCTGCAGGTGAAAGCGGTGCGCGCGAAAC
TGGATGAAAGCGCGTGGAGCCATCCGCAGTTTGAAAAA

Escherichia coli AldH:
ATGAACTTTCATCATCTGGCGTATTGGCAGGATAAAGCGCTGAGCCTGGCGATTGAAAAC
CGCCTGTTTATTAACGGCGAATATACCGCGGCGGCGGAAAACGAAACCTTTGAAACCGTG
GATCCGGTGACCCAGGCGCCGCTGGCGAAAATTGCGCGCGGCAAAAGCGTGGATATTGA
TCGCGCGATGAGCGCGGCGCGCGGCGTGTTTGAACGCGGCGATTGGAGCCTGAGCAGCC
CGGCGAAACGCAAAGCGGTGCTGAACAAACTGGCGGATCTGATGGAAGCGCATGCGGAA
GAACTGGCGCTGCTGGAAACCCTGGATACCGGCAAACCGATTCGCCATAGCCTGCGCGAT
GATATTCCGGGCGCGGCGCGCGCGATTCGCTGGTATGCGGAAGCGATTGATAAAGTGTAT
GGCGAAGTGGCGACCACCAGCAGCCATGAACTGGCGATGATTGTGCGCGAACCGGTGGG
CGTGATTGCGGCGATTGTGCCGTGGAACTTTCCGCTGCTGCTGACCTGCTGGAAACTGGG
CCCGGCGCTGGCGGCGGGCAACAGCGTGATTCTGAAACCGAGCGAAAAAAGCCCGCTGA

GCGCGATTCGCCTGGCGGGCCTGGCGAAAGAAGCGGGCCTGCCGGATGGCGTGCTGAAC
GTGGTGACCGGCTTTGGCCATGAAGCGGGCCAGGCGCTGAGCCGCCATAACGATATTGAT
GCGATTGCGTTTACCGGCAGCACCCGCACCGGCAAACAGCTGCTGAAAGATGCGGGCGA
TAGCAACATGAAACGCGTGTGGCTGGAAGCGGGCGGCAAAAGCGCGAACATTGTGTTTG
CGGATTGCCCGGATCTGCAGCAGGCGGCGAGCGCGACCGCGGCGGGCATTTTTTATAACC
AGGGCCAGGTGTGCATTGCGGGCACCCGCCTGCTGCTGGAAGAAAGCATTGCGGATGAA
TTTCTGGCGCTGCTGAAACAGCAGGCGCAGAACTGGCAGCCGGGCCATCCGCTGGATCCG
GCGACCACCATGGGCACCCTGATTGATTGCGCGCATGCGGATAGCGTGCATAGCTTTATT
CGCGAAGGCGAAAGCAAAGGCCAGCTGCTGCTGGATGGCCGCAACGCGGGCCTGGCGGC
GGCGATTGGCCCGACCATTTTTGTGGATGTGGATCCGAACGCGAGCCTGAGCCGCGAAGA
AATTTTTGGCCCGGTGCTGGTGGTGACCCGCTTTACCAGCGAAGAACAGGCGCTGCAGCT
GGCGAACGATAGCCAGTATGGCCTGGGCGCGGCGGTGTGGACCCGCGATCTGAGCCGCG
CGCATCGCATGAGCCGCCGCCTGAAAGCGGGCAGCGTGTTTGTGAACAACTATAACGATG
GCGATATGACCGTGCCGTTTGGCGGCTATAAACAGAGCGGCAACGGCCGCGATAAAAGC
CTGCATGCGCTGGAAAAATTTACCGAACTGAAAACCATTTGGATTAGCCTGGAAGCG

Lactobacillus reuteri Lreu0034:
ATGGCGTATCAGAGCATTAACCCGTTTACCAACCAGGTGGAAAAAACCTTTGAAAACACC
ACCGATGAAGAACTGGAACAGACCCTGACCACCGCGCATCAGCTGTATCTGGATTGGCGC
AAATATAACGATCTGGAAGAACGCAAACGCCAGATTCTGAAACTGGGCCAGATTCTGCG
CGAACGCCGCGTGGAATATGCGACCGTGATGAGCAAAGAAATGGGCAAACTGATTAGCG
AAGCGGAAGGCGAAGTGGATCTGTGCGCGAGCTTTTGCGATTATTATGCGGCGCATGCGG
ATGAATTTCTGCAGCCGAAAATTATTGCGACCACCAGCGGCCGCGCGAAAGTGCTGAAAC
AGAGCCTGGGCATTCTGGTGGCGGTGGAACCGTGGAACTTTCCGTTTTATCAGATTGCGC
GCGTGTTTATTCCGAACTTTATTGCGGGCAACCCGATGATTCTGAAAGATGCGAGCAACT
GCCCGGCGAGCGCGCAGGCGTTTAACGATGCGGTGAAAGAAGCGGGCGCGCCGGCGGGC
AGCCTGACCAACCTGTTTCTGAGCTATGATCAGGTGAACAAAGCGATTGCGGATAAACGC
GTGGCGGGCGTGTGCCTGACCGGCAGCGAACGCGGCGGCGCGACCGTGGCGAAAGAAGC
GGGCGCGAACCTGAAAAAAAGCACCCTGGAACTGGGCGGCAACGATGCGTTTATTATTCT
GGATGATGCGGATTGGGATCTGGTGGAAAAAGTGGCGCCGGCGGCGCGCCTGTATAACG
CGGGCCAGGTGTGCACCAGCAGCAAACGCTTTATTGTGCTGGAAAAAGATTATGATCGCT
TTCTGAAAATGATGAAAGATGCGTTTAGCAAAGTGAAAATGGGCGATCCGCTGGATCCGC
TGACCACCCTGGCGCCGCTGAGCAGCAAAAAAGCGAAAGAAAAACTGCAGCAGCAGGTG
GCGACCGCGGTGGAAAACGGCGCGAAAGTGTATTATGGCAACAAACCGGTGGATATGGA
AGGCCAGTTTTTTATGCCGACCATTCTGACCGATATTACCCCGGATAACCCGATTTTTGAT
ACCGAAATGTTTGGCCCGGTGGCGAGCGTGTATAAAGTGAGCAGCGAAGAAGAAGCGAT
TGAACTGGCGAACAACAGCAGCTATGGCCTGGGCAACACCATTTTTAGCAACGATAGCG
AACATGCGGAACGCGTGGCGGCGAAAATTGAAACCGGCATGAGCTGGATTAACGCGGGC
TGGGCGAGCCTGCCGGAACTGCCGTTTGGCGGCGTGAAAAACAGCGGCTATGGCCGCGA
ACTGAGCAGCTATGGCATTGATGAATTTACCAACAAACATCTGATTTATGAAGCGCGCCA
G

Gluconobacter oxydans Gox 0499:
ATGATTGTGTATCAGACCCTGAACCCGACCACCGAAACCGTGGAACGCAGCTTTGATCTG
CATACCCCGGCGCAGATGAAAGATATTACCGATCGCGCGGAACATGTGTGGAAAACCGA
TTGGAAACTGCGCAGCATTGCGCAGCGCAAAGAAATTGTGAGCCGCGCGGCGGATCTGC
TGCGCCGCGATCGCCAGCATCATGCGAGCCTGATTGCGACCGAAATGGGCAAAGCGCTG
CCGGATGCGCTGGAAGAAATTGATGTGACCGCGGATATTCTGAGCTTTTATGCGAACGGC
GCGGAAGAATTTCTGGCGCCGACCCCGCTGAAAGTGAAAACCGGCCAGGCGAAAATTAT
TAACCAGCCGCTGGGCATTATTTATTGCATTGAACCGTGGAACTTTCCGTATTATCAGCTG
GCGCGCGTGGCGGGCCCGAACCTGATGGCGGGCAACGTGGTGATTGCGAAACATGCGCC
GAACGTGCCGCAGTGCGCGCTGGCGTTTGAAAAACTGTTTCATGATGCGGGCGCGCCGGT
GGGCGCGTATGCGAACATTTTTCTGGATAACGATCAGAGCGCGGAACTGATTAAAGATGA
ACGCATTCGCGGCGTGGCGCTGACCGGCAGCGAACGCGCGGGCCAGGCGGTGGCGGCGC
AGGCGGGCGCGGCGCTGAAAAAAGATACCATGGAACTGGGCGGCAGCGATGCGTTTATT
GTGCTGGATGATGCGGATCTGGATCTGGCGGTGAAATGGGCGGTGTGGGGCCGCTTTGCG

AACAACGGCCAGGTGTGCACCGCGGCGAAACGCATGATTGTGCATGAAAAAGTGTATGA
TGCGTTTCTGGATGGCCTGAAAACCGCGATTACCCGCTTTCGCATTGGCAACCCGCTGGA
TCGCGATACCACCCATGGCCCGATGAGCAGCCTGCGCGCGATGGAACTGGCGCTGGATCA
GACCGCGGAAGCGGTGAAAGGCGGCGCGACCCTGGTGGCGGGCGGCAAACGCATGGATC
GCAAAGGCTTTTTTATGGAACCGACCATTCTGACCGATGTGAGCAAAGATAACCCGGTGT
TTTATCAGGAAATTTTTGGCCCGGTGGCGGTGGTGCATAAAGTGGCGAGCGAACAGGCGG
CGATTGATCTGGCGAACGATAGCCCGTATGGCCTGGGCGGCGCGGTGTTTAGCCGCGATA
TTGCGCGCGCGGAAAAAGTGGCGGAACAGGTGGAAACCGGCATGGTGTTTATTAACACC
GCGACCGCGGCGGCGCCGGAACTGCCGTTTGGCGGCATTAAAAACAGCGGCTTTGGCCG
CGAACTGAGCTTTCTGGGCATTGAAGAATTTATTAACCGCAAACTGGTGCGCATTGGC

Gluconobacter oxydans Gox 1712:
ATGACCGAAAAAAACAACCTGTTTATTAACGGCAGCTGGGTGGCGCCGAAAGGCGGCGA
ATGGATTAAAGTGGAAAACCCGGCGACCAAAGCGGTGGTGGCGGAAGTGGCGAAAGGC
GGCCAGGCGGATGTGGATGCGGCGGTGAGCGCGGCGAAAAGCGCGTTTATTGGCTGGAG
CCGCCGCACCGCGACCGAACGCGCGGATTATATTCATGCGCTGAAAGATCTGGTGAAACG
CGATAAAGAAAAACTGGCGGCGATTATTACCAGCGAAATGGGCAAACCGCTGAAAGAAG
CGCGCATTGAAGTGGATTTTGCGATTGGCCTGCTGCGCTTTAGCGCGGAAAACGTGCTGC
GCCTGCAGGGCGAAATTATTCCGGGCAGCAGCCCGGAAGAAAAAATTCTGATTGATCGC
GTGCCGCTGGGCGTGATTGGCGCGATTACCGCGTGGAACTTTCCGCTGGCGCTGTGCGCG
CGCAAAATTGGCCCGGCGGTGGCGGCGGGCAACACCATTGTGGTGAAACCGCATGAACT
GACCCCGCTGGCGTGCCTGCATCTGGCGAAACTGGTGGAAGAAGCGAAAATTCCGCATG
GCGTGATTAACGTGGTGACCGGCGATGGCAAAGATGTGGGCGTGCCGCTGGTGGCGCAT
AAAGATATTAAACTGATTACCATGACCGGCAGCACCCCGGCGGGCAAAAAAATTATGGC
GGCGGCGAGCGAAACCCTGAAAGAAGTGCGCCTGGAACTGGGCGGCAAAGCGCCGTTTA
TGGTGATGGAAGATGCGGATATTGATCGCGCGGCGGATGCGGCGGTGACCGCGCGCTTTA
ACAACGCGGGCCAGGTGTGCACCTGCAACGAACGCACCTATATTCATGAAGCGGTGTATG
ATCGCTTTGTGCAGAAAGTGCGCGAAAAAATTGAAGCGCTGAAAGTGGGCCTGCCGACC
GATCCGAGCACCGATATGGGCCCGAAAGTGAGCGAAGATGAACTGAACAAAGTGCATGA
AATGGTGGAACATGCGGTGCGCCAGGGCGCGCGCCTGGCGATTGGCGGCAAACGCCTGA
CCGGCGGCGTGTATGATAAAGGCTATTTTTATGCGCCGACCCTGCTGACCGATGTGACCC
AGGATATGGATATTGTGCATAACGAAGTGTTTGGCCCGGTGATGAGCCTGATTCGCGTGA
AAGATTTTGATCAGGCGATTGCGTGGGCGAACGATTGCCGCTATGGCCTGAGCGCGTATC
TGTTTACCAACGATCTGAGCCGCATTCTGCGCATGACCCGCGATCTGGAATTTGGCGAAG
TGTATGTGAACCGCCCGGGCGGCGAAGCGCCGCAGGGCTTTCATCATGGCTATAAAGAA
AGCGGCCTGGGCGGCGAAGATGGCCAGCATGGCATGGAAGCGTATGTGCAGACCAAAAC
CATTTATCTGAACGCG

5.2 Protein sequences

Nocardia iowensis sp. NRRL 5646 CAR:
MAVDSPDERLQRRIAQLFAEDEQVKAARPLEAVSAAVSAPGMRLAQIAATVMAGYADRPAA
GQRAFELNTDDATGRTSLRLLPRFETITYRELWQRVGEVAAAWHHDPENPLRAGDFVALLGF
TSIDYATLDLADIHLGAVTVPLQASAAVSQLIAILTETSPRLLASTPEHLDAAVECLLAGTTPER
LVVFDYHPEDDDQRAAFESARRRLADAGSLVIVETLDAVRARGRDLPAAPLFVPDTDDDPLA
LLIYTSGSTGTPKGAMYTNRLAATMWQGNSMLQGNSQRVGINLNYMPMSHIAGRISLFGVL
ARGGTAYFAAKSDMSTLFEDIGLVRPTEIFFVPRVCDMVFQRYQSELDRRSVAGADLDTLDR
EVKADLRQNYLGGRFLVAVVGSAPLAAEMKTFMESVLDLPLHDG1YGSTEAGASVLLDNQI
QRPPVLDYKLVDVPELGYFRTDRPHPRGELLLKAETTIPGYYKRPEVTAEIFDEDGFYKTGDIV
AELEHDRLVYVDRRNNVLKLSQGEFVTVAHLEAVFASSPLIRQIFIYGSSERSYLLAVIVPTDD
ALRGRDTATLKSALAESIQRIAKDANLQPYEIPRDFLIETEPFTIANGLLSGIAKLLRPNLKERY
GAQLEQMYTDLATGQADELLALRREAADLPVLETVSRAAKAMLGVASADMRPDAHFTDLG
GDSLSALSFSNLLHEIFGVEVPVGVVVSPANELRDLANYIEAERNSGAKRPTFTSVHGGGSEIR
AADLTLDKFIDARTLAAADSIPHAPVPAQTVLLTGANGYLGRFLCLEWLERLDKTGGTLICVV
RGSDAAAARKRLDSAFDSGDPGLLEHYQQLAARTLEVLAGDIGDPNLGLDDATWQRLAETV
DLIVHPAALVNHVLPYTQLFGPNVVGTAEIVRLAITARRKPVTYLSTVGVADQVDPAEYQED
SDVREMSAVRVVRESYANGYGNSKWAGEVLLREAHDLCGLPVAVFRSDMILAHSRYAGQL
NVQDVFTRLILSLVATGIAPYSFYRTDADGNRQRAHYDGLPADFTAAAITALGIQATEGFRTY
DVLNPYDDGISLDEFVDWLVESGHPIQRITDYSDWFHRFETAIRALPEKQRQASVLPLLDAYR
NPCPAVRGAILPAKEFQAAVQTAKIGPEQDIPHLSAPLIDKYVSDLELLQLL

Saccharomyces cerevisiae AAR (Lys2):
MTNEKVWIEKLDNPTLSVLPHDFLRPQQEPYTKQATYSLQLPQLDVPHDSFSNKYAVALSVW
AALIYRVTGDDDIVLYIANNKILRFNIQPTWSFNELYSTINNELNKLNSIEANFSFDELAEKIQS
CQDLERTPQLFRLAFLENQDFKLDEFKHHLVDFALNLDTSNNAHVLNLIYNSLLYSNERVTIV
ADQFTQYLTAALSDPSNCITKISLITASSKDSLPDPTKNLGWCDFVGCIHDIFQDNAEAFPERTC
VVETPTLNSDKSRSFTYRDINRTSNIVAHYLIKTGIKRGDVVMIYSSRGVDLMVCVMGVLKA
GATFSVIDPAYPPARQTIYLGVAKPRGLIVIRAAGQLDQLVEDYINDELEIVSRINSIAIQENGTI
EGGKLDNGEDVLAPYDHYKDTRTGVVVGPDSNPTLSFTSGSEGIPKGVLGRHFSLAYYFNW
MSKRFNLTENDKFTMLSGIAHDPIQRDMFTPLFLGAQLYVPTQDDIGTPGRLAEWMSKYGCT
VTHLTPAMGQLLTAQATTPFPKLHHAFFVGDILTKRDCLRLQTLAENCRIVNMYGTTETQRA
VSYFEVKSKNDDPNFLKKLKDVMPAGKGMLNVQLLVVNRNDRTQICGIGEIGEIYVRAGGLA
EGYRGLPELNKEKFVNNWFVEKDHWNYLDKDNGEPWRQFWLGPRDRLYRTGDLGRYLPN
GDCECCGRADDQVKIRGFRIELGEIDTHISQHPLVRENITLVRKNADNEPTLITFMVPRFDKPD
DLSKFQSDVPKEVETDPIVKGLIGYHLLSKDIRTFLKKRLASYAMPSLIVVMDKLPLNPNGKV
DKPKLQFPTPKQLNLVAENTVSETDDSQFTNVEREVRDLWLSILPTKPASVSPDDSFFDLGGH
SILATKMIFTLKKKLQVDLPLGTIFKYPTIKAFAAEIDRIKSSGGSSQGEVVENVTANYAEDAK
KLVETLPSSYPSREYFVEPNSAEGKTTINVFVTGVTGFLGSYILADLLGRSPKNYSFKVFAHVR
AKDEEAAFARLQKAGITYGTWNEKFASNIKVVLGDLSKSQFGLSDEKWMDLANTVDIIIHNG
ALVHWVYPYAKLRDPNVISTINVMSLAAVGKPKFFDFVSSTSTLDTEYYFNLSDKLVSEGKPG
ILESDDLMNSASGLTGGYGQSKWAAEYIIRRAGERGLRGCIVRPGYVTGASANGSSNTDDFLL
RFLKGSVQLGKIPDIENSVNMVPVDHVARVVVATSLNPPKENELAVAQVTGHPRILFKDYLY
TLHDYGYDVEIESYSKWKKSLEASVIDRNEENALYPLLHMVLDNLPESTKAPELDDRNAVAS
LKKDTAWTGVDWSNGIGVTPEEVGIYIAFLNKVGFLPPPTHNDKLPLPSIELTQAQISLVASGA
GARGSSAAA

Streptomyces griseus GriC:
MSLIDDPEFDSHVVKLMAWHFGAETGSPFWLGKLAGLGFDPVTDVRGLADLTRFPDVSAEL
RSVPAQDLIPRGLSDRPFRVYDSGGTTGSPKRIVDSGYRSLVTEWARERLIANGVPADGNWLH
LGPGGPHVIGFDTARYAALGGGVFYTVDLDPRWVKRLLSQGRGEESEAYVQHLLDQAETILE
SQEIAVLNTTPPLLEAICARPRLYELVRSGVRAIIWAGTSISRETLTQLDDVFFPAASVVGVYG
NSLMGVAPQRSGRPGDAHRCVFEPYPETTRLHLVDEDGTPVAYGERGRVRLHLVSEEMFLPN
VLERDTAIRVEPGEGSTVDGLADVQTYRSLDDVAIIEGVY

Streptomyces griseus GriD:
MSTVAPPAVVLDPRVRGRLVASSDRTRLPSVLGGDLADVGTAPRLLAVAALNEIRAHADGRP
PASEVFTEAARLFATATLDGETPREYARRVCHATGLTATAVEQAVTDLTGELRALPATTAAE
LPATGFGDGFDTRWVPRGRTFAAVMASNHPVPNISWAQALFHGYSVLVKPGSRDPFTPARLL
AALTAAGLPPDRAAFLPCSREVGAYLLREADRGIVYGGDSAVATWHQRESVAVRGPGRTKA
FLDRAPDDAVIDHLALSASFDGGTRCTNLSAVLTTGPVEEVADRLAERLARLPSLPATDEAAT
LLVVNRARADGLREQVAALRASLTDHSARAGDPETVVELPDGSFLPRPVVLSADRADHPAV
GTELPFPFVVVAPWAEADGVEPLRDSLVLNLLTDREELVEAAVREPSVRKVTRGPVLPWTAT
PGIPHDDNYTQFLLEPKGVVARR

Bacillus subtilis PPTase:
MKIYGIYMDRPLSQEENERFMSFISPEKREKCRRFYHKEDAHRTLLGDVLVRSVISRQYQLDK
SDIRFSTQEYGKPCIPDLPDAHFNISHSGRWVICAFDSQPIGIDIEKTKPISLEIAKRFFSKTEYSD
LLAKDKDEQTDYFYHLWSMKESFIKQEGKGLSLPLDSFSVRLHQDGQVSIELPDSHSPCYIKT
YEVDPGYKMAVCAAHPDFPEDITMVSYEELL

Saccharomyces cerevisiae PPTase (Lys5):
MVKTTEVVSEVSKVAGVRPWAGIFVVEIQEDILADEFTFEALMRTLPLASQARILNKKSFHDR
CSNLCSQLLQLFGCSIVTGLNFQELKFDKGSFGKPFLDNNRFLPFSMTIGEQYVAMFLVKCVS
TDEYQDVGIDIASPCNYGGREELELFKEVFSEREFNGLLKASDPCTIFTYLWSLKESYTKFTGT
GLNTDLSLIDFGAISFFPAEGASMCITLDEVPLIFHSQWFNNEIVTICMPKSISDKINTNRPKLYN
ISLSTLIDYFIENDGL

Acinetobacter sp. BALD:
MTLLDASIWNKKLFNGGWFESGQPYGVVEVATGEQLGQTGSASPTDVAQAAKEAQTAQRQ
WWALDYLERQAVFEKAVQIATEHQDELVNWLIRESGSLALKAQFEVKVSIQMLKNCIAFPQL
DQGNILPSRNGKLSLAKRLPLGVVGVISPFNFPLYLALRAVGPALAFGNAVVLKPDERTAVCS
GYAIARIFELAGLPKGLLHVLPGGADTGEALTLDPHIASIQFTGSTQVGRIIGANAGKTLKKVS
LELGGKNSLIILDDADLDLAAENIAWGAFLHSGQICMTSGKILIHEKIYDALKARVIEKVKRFV
VGNPLEQNVTIGPLINEKQSKRVEALVQAAVEHGATLEIGGKANGPFFEPSVLSQVQADNPIFS
EEIFGPVAVLIPFASDEQAIELANMGDYGLSAGIISSNVGRAMQLGAQLNVGLLHINDQTVND
ETINPFGGFGASGNATRIGGPANPDEFTQWQWMTIQAEAPHYPFSAWSHPQFEK

Saccharomyces cerevisiae ALDH4:
MFSRSTLCLKTSASSIGRLQLRYFSHLPMTVPIKLPNGLEYEQPTGLFINNKFVPSKQNKTFEVI
NPSTEEEICHIYEGREDDVEEAVQAADRAFSNGSWNGIDPIDRGKALYRLAELIEQDKDVIASI
ETLDNGKAISSSRGDVDLVINYLKSSAGFADKIDGRMIDTGRTHFSYTKRQPLGVCGQIIPWNF
PLLMWAWKIAPALVTGNTVVLKTAESTPLSALYVSKYIPQAGIPPGVINIVSGFGKIVGEAITN
HPKIKKVAFTGSTATGRHIYQSAAAGLKKVTLELGGKSPNIVFADAELKKAVQNIILGIYYNS
GEVCCAGSRVYVEESIYDKFIEEFKAASESIKVGDPFDESTFQGAQTSQMQLNKILKYVDIGKN
EGATLITGGERLGSKGYFIKPTVFGDVKEDMRIVKEEIFGPVVTVTKFKSADEVINMANDSEY
GLAAGIHTSNINTALKVADRVNAGTVWINTYNDFHHAVPFGGFNASGLGREMSVDALQNYL
QVKAVRAKLDESAWSHPQFEK

Escherichia coli AldH:
MNFHHLAYWQDKALSLAIENRLFINGEYTAAAENETFETVDPVTQAPLAKIARGKSVDIDRA
MSAARGVFERGDWSLSSPAKRKAVLNKLADLMEAHAEELALLETLDTGKPIRHSLRDDIPGA
ARAIRWYAEAIDKVYGEVATTSSHELAMIVREPVGVIAAIVPWNFPLLLTCWKLGPALAAGN
SVILKPSEKSPLSAIRLAGLAKEAGLPDGVLNVVTGFGHEAGQALSRHNDIDAIAFTGSTRTGK
QLLKDAGDSNMKRVWLEAGGKSANIVFADCPDLQQAASATAAGIFYNQGQVCIAGTRLLLE
ESIADEFLALLKQQAQNWQPGHPLDPATTMGTLIDCAHADSVHSFIREGESKGQLLLDGRNA
GLAAAIGPTIFVDVDPNASLSREEIFGPVLVVTRFTSEEQALQLANDSQYGLGAAVWTRDLSR
AHRMSRRLKAGSVFVNNYNDGDMTVPFGGYKQSGNGRDKSLHALEKFTELKTIWISLEA

Lactobacillus reuteri Lreu0034:
MAYQSINPFTNQVEKTFENTTDEELEQTLTTAHQLYLDWRKYNDLEERKRQILKLGQILRERR
VEYATVMSKEMGKLISEAEGEVDLCASFCDYYAAHADEFLQPKIIATTSGRAKVLKQSLGILV
AVEPWNFPFYQIARVFIPNFIAGNPMILKDASNCPASAQAFNDAVKEAGAPAGSLTNLFLSYD
QVNKAIADKRVAGVCLTGSERGGATVAKEAGANLKKSTLELGGNDAFIILDDADWDLVEKV
APAARLYNAGQVCTSSKRFIVLEKDYDRFLKMMKDAFSKVKMGDPLDPLTTLAPLSSKKAK
EKLQQQVATAVENGAKVYYGNKPVDMEGQFFMPTILTDITPDNPIFDTEMFGPVASVYKVSS
EEEAIELANNSSYGLGNTIFSNDSEHAERVAAKIETGMSWINAGWASLPELPFGGVKNSGYGR
ELSSYGIDEFTNKHLIYEARQ

Gluconobacter oxydans Gox 0499:
MIVYQTLNPTTETVERSFDLHTPAQMKDITDRAEHVWKTDWKLRSIAQRKEIVSRAADLLRR
DRQHHASLIATEMGKALPDALEEIDVTADILSFYANGAEEFLAPTPLKVKTGQAKIINQPLGIIY
CIEPWNFPYYQLARVAGPNLMAGNVVIAKHAPNVPQCALAFEKLFHDAGAPVGAYANIFLD
NDQSAELIKDERIRGVALTGSERAGQAVAAQAGAALKKDTMELGGSDAFIVLDDADLDLAV
KWAVWGRFANNGQVCTAAKRMIVHEKVYDAFLDGLKTAITRFRIGNPLDRDTTHGPMSSLR
AMELALDQTAEAVKGGATLVAGGKRMDRKGFFMEPTILTDVSKDNPVFYQEIFGPVAVVHK
VASEQAAIDLANDSPYGLGGAVFSRDIARAEKVAEQVETGMVFINTATAAAPELPFGGIKNSG
FGRELSFLGIEEFINRKLVRIG

Gluconobacter oxydans Gox 1712:
MTEKNNLFINGSWVAPKGGEWIKVENPATKAVVAEVAKGGQADVDAAVSAAKSAFIGWSR
RTATERADYIHALKDLVKRDKEKLAAIITSEMGKPLKEARIEVDFAIGLLRFSAENVLRLQGEII
PGSSPEEKILIDRVPLGVIGAITAWNFPLALCARKIGPAVAAGNTIVVKPHELTPLACLHLAKL
VEEAKIPHGVINVVTGDGKDVGVPLVAHKDIKLITMTGSTPAGKKIMAAASETLKEVRLELG
GKAPFMVMEDADIDRAADAAVTARFNNAGQVCTCNERTYIHEAVYDRFVQKVREKIEALKV
GLPTDPSTDMGPKVSEDELNKVHEMVEHAVRQGARLAIGGKRLTGGVYDKGYFYAPTLLTD
VTQDMDIVHNEVFGPVMSLIRVKDFDQAIAWANDCRYGLSAYLFTNDLSRILRMTRDLEFGE
VYVNRPGGEAPQGFHHGYKESGLGGEDGQHGMEAYVQTKTIYLNA