Journal of Biotechnology 257 (2017) 222–232

Contents lists available at ScienceDirect

Journal of Biotechnology

journal homepage: www.elsevier.com/locate/jbiotec

Four distinct types of E.C. 1.2.1.30 enzymes can catalyze the reduction of carboxylic acids to aldehydes

Holly Stolterfoht[a,b], Daniel Schwendenwein[b], Christoph W. Sensen[c], Florian Rudroff[d], Margit Winkler[a,b,*]

[a] Institute of Molecular Biotechnology, Graz University of Technology, Petersgasse 14, 8010 Graz, Austria
[b] acib GmbH, Petersgasse 14, 8010 Graz, Austria
[c] Institute of Computational Biotechnology, Graz University of Technology, Petersgasse 14, 8010 Graz, Austria
[d] Institute of Applied Synthetic Chemistry, TU Wien, Getreidemarkt 9/OC-163, 1060 Vienna, Austria

ARTICLE INFO

Article history:
Received 13 December 2016
Received in revised form 8 February 2017
Accepted 14 February 2017
Available online 20 February 2017

Keywords:
Carboxylate reductase
Biocatalysis
Signature sequence
Bootstrap analysis
Aryl-aldehyde dehydrogenase (NADP+)
Flavor and fragrance

ABSTRACT

Increasing demand for chemicals from renewable resources calls for the development of new biotechnological methods for the reduction of oxidized bio-based compounds. Enzymatic carboxylate reduction is highly selective, both in terms of chemo- and product selectivity, but not many carboxylate reductase enzymes (CARs) have been identified on the sequence level to date. Thus far, their phylogeny is unexplored and very little is known about their structure-function-relationship. CARs minimally contain an adenylation domain, a phosphopantetheinylation domain and a reductase domain. We have recently identified new enzymes of fungal origin, using similarity searches against genomic sequences from organisms in which aldehydes were detected upon incubation with carboxylic acids. Analysis of sequences with known CAR functionality and CAR enzymes recently identified in our laboratory suggests that the three-domain architecture mentioned above is modular. The construction of a distance tree with a subsequent 1000-replicate bootstrap analysis showed that the CAR sequences included in our study fall into four distinct subgroups (one of bacterial origin and three of fungal origin, respectively), each with a bootstrap value of 100%. The multiple sequence alignment of all experimentally confirmed CAR protein sequences revealed fingerprint sequences of residues which are likely to be involved in substrate and co-substrate binding and one of the three catalytic substeps, respectively. The fingerprint sequences broaden our understanding of the amino acids that might be essential for the reduction of organic acids to the corresponding aldehydes in CAR proteins.

1. Introduction

The demand for non-petroleum based chemical products is steadily increasing and platform technologies are emerging in which organic feedstock or waste is converted to short-chain carboxylic acids (sugar, syngas and carboxylate platform) (Agler et al., 2011; Perez et al., 2013). In order to access many of the intermediates, which are used in current petrol based processes, efficient reductive processes, starting from bio-based carboxylates need to be developed. Although enzymes are already available for several of the required functional group transformations, such as the selective reduction of double bonds (Toogood and Scrutton, 2014) or the reduction of imines (Schrittwieser et al., 2015), the industrial conversion of organic acids to aldehydes is still in need of significant further development. Aldehydes are not only used in the flavor/fragrance sector (Carroll et al., 2016; Sell, 2006) and as agrochemicals but are also important reactive intermediates for a whole range of commercial products, such as pharmceutical intermediates (van Langen et al., 2003).

Over the last sixty years, a number of organisms have been identified to be able to catalyze the reduction of carboxylic acids. The majority of literature reports relate to identifying this activity in fungi, but bacteria, archea and plants, respectively, have also been shown to host carboxylate reductase activity

Abbreviations: CAR, carboxylate reductase; ATP, adenosine 5′-triphosphate; AMP, adenosine 5′-monophosphate; NADPH, β-nicotinamide adenine dinucleotide phosphate; A, adenylation; T, transthiolation; R, reductase.

* Corresponding author at: Institute of Molecular Biotechnology, Graz University of Technology, Petersgasse 14, 8010 Graz, Austria.
E-mail addresses: margit.winkler@tugraz.at, margit.winkler@acib.at (M. Winkler).

http://dx.doi.org/10.1016/j.jbiotec.2017.02.014
0168-1656/

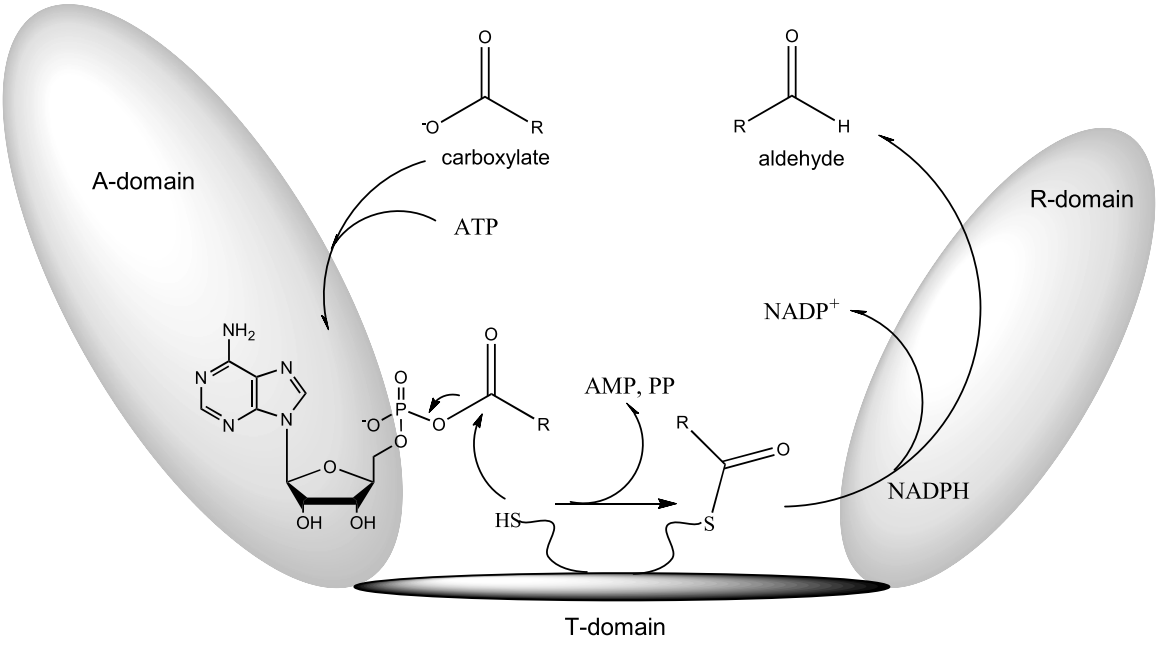

Fig. 1. Schematic representation of E.C.1.2.1.30 carboxylate reductase mediated aldehyde formation.

(da S. Amaral and Rodrigues-Filho, 2015; Brenna et al., 2015; Napora-Wijata et al., 2014). Archea and anaerobic bacteria possess metal-dependent enzymes, which are called aldehyde oxidoreductases (AORs) or carboxylate reductases (CARs), belonging to the E.C. 1.2.99.6 family, to reduce acids to aldehydes. Aerobic bacteria and fungi typically express ATP and NADPH dependent enzymes, which were initially named aryl-aldehyde dehydrogenases ($NADP^+$), but are now also mostly referred to as carboxylate reductases (CARs). These enzymes are classified as the E.C. 1.2.1.30 family. Both the E.C. 1.2.99.6 CAR family, as well as the E.C. 1.2.1.30 CAR family, respectively, exhibit a broad substrate tolerance for the conversion of organic acids to the respective aldehydes. While the E.C. 1.2.99.6 enzymes (which are not in the focus of this study) have predominantly been reported to act on short chain aliphatic substrates, the E.C. 1.2.1.30 enzymes have been mostly reported using aromatic substrates (Napora-Wijata et al., 2014). E.C. 1.2.1.30 enzymes are comprised of three domains: an adenylation domain (A-domain), a phosphopantetheinyl binding domain (also called transthiolation domain (T-domain), or peptidyl carrier protein (PCP domain)), and a reductase domain (R-domain). The phosphopantetheinyl-binding domain is recognized by a phosphopantetheinyl transferase enzyme, which attaches a phosphopantetheinyl residue to a conserved serine (Venkitasubramanian et al., 2007). Only upon this post-translational modification, the enzymes become active and are able to engage in the catalytic cycle. This cycle starts with the activation of the carboxylate substrate with ATP in the A-domain, yielding an AMP-ester intermediate under release of pyrophosphate as the co-product. The active thiol tether of the phosphopantetheinyl moiety then binds the carboxylate, releasing AMP as a leaving group. The resulting thioester is subsequently transferred to the R-domain, where it is reduced to the corresponding aldehyde product. Fig. 1 summarizes the current understanding of the reaction (Venkitasubramanian et al., 2007). The aldehyde is not amenable to enter a second catalytic cycle. Enzymatic reduction by E.C. 1.2.1.30 CARs is product-selective because the enzyme does not catalyze the overreduction of the aldehyde product to the respective alcohol. This product selectivity is an essential asset of an enzyme if it is used for the biocatalytic synthesis of organic molecules on the preparative level. Another equally important asset is the relaxed substrate specificity of E.C. 1.2.1.30 CARs, which allows their use for the reduction of a broad range of compounds (Napora-Wijata et al., 2014) such as vanillin (Venkitasubramanian et al., 2008) and other aromatic acid derivatives (Finnigan et al., 2016) as well as aliphatic acids (Akhtar et al., 2013; Fraatz et al., 2016).

Another enzyme family with the ability to reduce the carboxylate moiety to the respective aldehyde in a highly similar manner, are L-aminoadipate reductases (Guo and Bhattacharjee, 2004). As their name implies, these enzymes are highly specific for the reduction of L-aminoadipate to L-aminoadipate-semialdehyde. These reductases are classified as E.C. 1.2.1.95 enzymes. In their domain architecture, an additional *N*-terminal domain was reported to exert an activating function (Kalb et al., 2015). Recently, a homologue of E.C. 1.2.1.30 CARs was described, which was named tyrosine reductase. The name of this enzyme refers to the apparently strict specificity for the reduction of tyrosine (Forseth et al., 2013). This enzyme is also characterized by A-T-R domain architecture and believed to follow the same reaction mechanism as shown in Fig. 1.

2. Materials and methods

2.1. Alignment, Phylogenetic Tree Construction and Bootstrap Analysis

Alignment: 17 Sequences including the NRPS sequence from *Neosartorya fisheri*, which was used as an outgroup to be able to root the CAR sequences amongst each other (Supporting information Table S1) were aligned using T-Coffee (http://www.tcoffee.org). *Phylogenetic Analysis:* The phylogenetic analysis was performed using Phylip version 3.696 (http://evolution.genetics.washington.edu/phylip.html). *Bootstrap Analysis:* a delete-half Jacknife dataset of 1000 sub-alignments was created using seqboot, followed by protdist (with M option) and fitch (also with M option). The resulting trees were collated using the program consense. The bootstrapped tree was subsequently visualized and re-labeled using FigTree V 1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/).

2.2. Multiple sequence alignment with highlights of functional regions

Seven CAR sequences with confirmed function, representing members from all four distinct groups of E.C. 1.2.1.30 CARs, were aligned using T-Coffee. Motifs identified with the Motif Search Tool at GenomeNet Japan (http://www.genome.jp/tools/motif/) were highlighted using background colors in the alignment. In addition, A-domains of fourteen CAR sequences with confirmed function were aligned with A-domains of selected NRPS proteins (Fig. S1). Those positions of the 10-residue specific codes of NRPS, were identified manually in CAR sequences, respectively (Supporting information Table S3).

2.3. Chemicals and solvents

ATP was obtained from Roche Diagnostics. NADPH, MES and $MgCl_2$ hexahydrate were purchased from Roth. All other chemicals were obtained from Sigma–Aldrich/Fluka or Roth and used without further purification.

2.4. Site directed mutagenesis

The mutants of the CAR_{Nc} gene, *Nc*CAR, were amplified from the vector pETDUET1_*Ec*PPTase_HT*Nc*CAR by polymerase chain reaction (PCR) using the following primers: pETDuet1_fwd as forward primer and M1r to M17r as reverse primer were used for amplification of the 5′ part of the gene, pETDuet1_rev as forward and M1f to M17f as reverse primer for the 3′ part. For primer sequences see Supporting information (Table S2). The PCR was performed using the following thermal cycle parameters: initial denaturation at 98 °C for 30 s, followed by 25 cycles of denaturation at 98 °C for 10 s, annealing at 56 °C for 15 s, extension at 72 °C for 40 s, and a final extension at 72 °C for 2 min with the Phusion® High-Fidelity DNA Polymerase. The PCR products were purified via agarose gel electrophoresis, the desired fragments were extracted using the GeneJET Gel Extraction Kit (Thermo Scientific). The purified PCR products were subsequently joined via overlap extension PCR with the primers pETDuet1_fwd and pETDuet1_rev to obtain the complete *Nc*CAR gene carrying the desired mutation. The PCR reaction was performed using the following conditions: initial denaturation at 98 °C for 30 s, followed by 25 cycles of denaturation at 98 °C for 10 s, annealing at 56 °C for 15 s, extension at 72 °C for 1 min, and a final extension at 72 °C for 2 min.

The vector backbone was amplified by PCR using the following primers: pETDuet1:*Ec*PPTase rev and pETDuet1:*Ec*PPTase fwd and the pETDUET1_*Ec*PPTase_HT*Nc*CAR as a template. The PCR was performed using the following conditions: initial denaturation at 98 °C for 30 s, followed by 25 cycles of denaturation at 98 °C for 10 s, annealing at 56 °C for 15 s, extension at 72 °C for 2 min, and a final extension at 72 °C for 4 min. The PCR product was purified via agarose gel and extracted as described above.

The mutated PCR products and the vector backbone were subsequently used for Gibson assembly (Gibson et al., 2009) with a 1:3 vector: insert ratio at 50 °C for 60 min, followed by a desalting step and transformation into electrocompetent *E. coli* TOP10. The recombinant vectors, designated pETDUET1_*Ec*PPTase_HT*Nc*ACAR_M1-M17, were purified and the correct mutation was confirmed by automated sequencing (Microsynth).

2.5. Expression and purification of Nc*CAR mutants*

E. coli K12 MG1655 RARE was transformed with the plasmid pETDUET1:*Ec*PPTaseHT*Nc*CAR wild-type and mutants, and colonies were selected on LB/Amp. For protein expression, the autoinduction protocol as described by Studier was used (Studier, 2005). Cultivations were performed in an RS 306 shaker (Infors AG) and Multitron shakers (Infors AG), and the cells were harvested with an Avanti J-20 centrifuge (Beckman Coulter). After 24 h at 20 °C, the cells were harvested by centrifugation and stored at −20 °C. *Nc*CAR expression levels were visualized by SDS-PAGE® usingNovex® 4–12% Bis-Tris-gels (50 min at 200 V in MOPS-buffer) and SimplyBlue SafeStain (Invitrogen) after protein extraction with BugBuster® 10 x Protein Extraction Reagent (EMD Milipore corp.). Thawed cells were disrupted by sonication using a 102C converter with a Sonifier 250 (Branson), and the cell-free extract was obtained by centrifugation in an Ultracentrifuge Optima LE80 K (Beckman). Proteins were purified by nickel affinity chromatography on Ni Sepharose 6 Fast Flow (GE Healthcare) using the gravity flow protocol. The protein containing fractions were pooled and the buffer exchanged for 50 mM MES buffer, pH 7.5, containing 10 mM $MgCl_2$, 1 mM EDTA, and 1 mM DTT by size exclusion chromatography on PD-10 columns (GE-Healthcare) using the gravity flow protocol. Protein concentrations were determined with the BCA Kit (Thermo Scientific) using BSA as the standard, and a Nanodrop 2000c spectrophotometer (Thermo Scientific). Aliquots of the protein solutions were shock frozen in liquid nitrogen and stored at −80 °C.

2.6. Spectrophotometric assay

An NADPH depletion assay was used to compare the activity of mutant *Nc*CARs to the wild-type. The assay composition was as follows: 10 μL of NADPH (10 mM in water), 10 μL of ATP (20 mM in water), 10 μL of CAR enzyme preparation (0.3-2.9 mg mL^{-1}), 160 μL of MES buffer (50 mM, pH 6.0, containing 10 mM $MgCl_2$) and 10 μL of cinnamic acid (100 mM in 0.1 M KOH). The depletion of NADPH was followed on a Synergy Mx Platereader (BioTek) at 340 nm and 28 °C for 10 min. Appropriate blank reactions (same mixture but without ATP) were carried out in parallel and each reaction was carried out up to five times in four technical replicates, respectively.

3. Results and discussion

In contrast to widely studied enzyme classes, such as alcohol dehydrogenases, esterases or transaminases, where hundreds of protein sequences have been examined already, the number of sequences with the experimentally confirmed ability to reduce acids to aldehydes is rather limited thus far. Therefore our research aims were to identify new enzyme sequences with an A-T-R domain architecture with a relaxed substrate specificity, to study their structure-function-relationship and to investigate their potential as biocatalysts for organic synthesis, respectively.

Our strategy for the identification of new CAR sequences was to use four CAR sequences with confirmed activity (*Nocardia iowensis* CAR: Q6RKB1.1, *Mycobacterium marinum* CAR: B2HN69, *Segniliparius rotundus* CAR: WP_013138593.1 and *Aspergillus terreus* CAR: XP_001212808.1) as target sequences for BLASTP searches (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) against the GenBank non-redundant protein sequences of organisms showing a reported ability to catalyze acid reduction. Approximately 250 sequences were identified and manually analyzed. Sequences with truncated A- or R-domain, as well as sequences having additional *C*- or *N*-terminal domains were discarded. The remaining sequences were ranked, based on the BLAST p-values, the number of predicted transmembrane domains, and potential subcellular targeting. Finally, the most promising candidates were subjected to expression in *E. coli*, leading to the identification of the CAR function in the protein with the GenBank accession number XP_008043822.1 from the basidiomycete *Trametes versicolor* (*Tv*CAR) (Winkler and

Winkler, 2016) as well as the protein with GenBank accession number XP_955820.1 from the ascomycete *Neurospora crassa* (*Nc*CAR) (Schwendenwein et al., 2016).

3.1. Preliminary sequence analysis

In the course of the sequence analysis described above, we noticed that the A- and R-domains of bacterial CARs, respectively, were classified as FAA1 and Lys2B by the multi-domain model (conserved domain search result, using 'reverse position-specific BLAST') (Marchler-Bauer and Bryant, 2004), whereas the domains of *Aspergillus terreus* CAR (*At*CAR) (Wang et al., 2014; Wang and Zhao, 2014) were classified as CaiC (A-domain) and Thioesterredct (R-domain). The CAR from *Trametes versicolor* seemed to be a hybrid, because this sequence contains an A-domain with similarity to bacterial CARs (FAA1) and an R-domain that resembles the one of the fungal *At*CAR. *Nc*CAR, however, was again distinct, with an 'AMP-binding' A-domain, according to the multi-domain model. This observation indicates that a certain level of modularity is inherent to CARs, which means that e.g. different A-domains can be linked to one PCP/R-domain combination or different R-domains to one A-domain/PCP combination. Our results prompted us to extend a study on fungal adenylate-forming reductases (Kalb et al., 2014). A preliminary phylogenetic analysis of all confirmed CAR sequences (Table 1, entries 1–14) indicated a clustering of the 10 bacterial CAR sequences, whereas the four fungal CAR sequences included did not seem to be too closely related.

3.2. Bootstrap analysis

A bootstrap analysis with 1000 replicates revealed four distinct groups of enzymes (type I: bacterial CARs; type II: *Tv*CAR group, type III: *At*CAR group; type IV: *Nc*CAR group) in the phylogenetic tree, each of which is supported by bootstrap values of 100%. The criterion for selecting sequences for this analysis was CAR activity that has been confirmed in the wet lab. Since all CARs show broad substrate specificity, a pattern of substrate specificity can not be assigned to the groups. The type III of A-T-R reductases resembles class III fungal adenylate-forming reductases, as suggested by the Hoffmeister group (Kalb et al., 2014).

3.3. CAR domains

Adenylation domains are not only integral to CARs, but have been widely studied in the context of non-ribosomal peptide synthesis. A-Domains of non-ribosomal peptide synthetases (NRPS) recognize and activate a particular amino acid to be incorporated into a specific location in the peptide product (Marahiel et al., 1997) and significant efforts have been made to understand and predict substrate specificity in a given sequence. Key positions in these A-domains have been identified and selectivity conferring NRPS-codes defined (Stachelhaus et al., 1999 Bushley et al., 2008). In contrast, the A-domains of CARs show broad substrate selectivity and activate structurally diverse carboxylates, as summarized in Table 1. The most-widely studied and longest-known CAR (*i.e. Nocardia iowensis*) has been shown to reduce approximately one hundred different carboxylic acids (Table 1, entry 1). Other bacterial CARs (Table 1, entries 2,4,5,11-14) also have been shown to possess a broad substrate scope. Two recently published bacterial CARs (Table 1, entries 6 and 7) were only reported to be active for the reduction of benzoic acid and short chain aliphatic acids (Moura et al., 2016), but due to high homology to the afore-mentioned bacterial CARs, these CARs might likely be able to reduce a number of other substrates as well. CARs from the fungal kingdom are not different with respect to relaxed substrate specificity, as shown in Table 1, (see entries 3, 8 and 10). Entry No. 9 in Table 1, describes an enzyme with only one known substrate, however to our knowledge, its substrate scope has not been further explored yet (Li et al., 2016). Tyrosine reductases (Table 1, entries 15 and 16), by contrast, have been assayed with some potential substrates, however, tyrosine was found to be the only compound which was reduced up to now (Forseth et al., 2013). Although tyrosine reductases share the general domain architecture of CARs, their substrate specificity indicates a closer relationship to NRPS enzymes with highly specific A-domains than to CARs, as supported by a phylogenetic analysis (Brandenburger et al., 2016). Moreover, tyrosine reductases form a dimer of 2-amino-3-(4-hydroxyphenyl)propanal instead of the aldehyde monomer. L-Aminoadipate reductases (LAAR) are another enzyme class with the ability to convert an acid to the respective aldehyde. LAARs are also highly specific for their name-giving substrate (Moura et al., 2016) and a particular amino acid pattern for LAARs has been extracted (Kalb et al., 2014). Table 1, entry 17 shows that not only the strict substrate scope of LAARs diverges them from CARs: it was shown recently that an additional *N*-terminal X domain seems to be necessary for substrate activation (Kalb et al., 2015), therefore tyrosine reductases and LAARs were omitted in our analysis of CARs.

3.4. Identification of key residues for CAR activity

Until recently, only bacterial CAR sequences were known to exist (Table 1, entries 1,2 and 4–7). The alignment of experimentally confirmed bacterial CARs resulted in >300 conserved amino acids (approximately 160 in the A-domain (Finnigan et al., 2016), 25 in the T-domain and 170 in the R-domain), which made it impossible to narrow down the most relevant amino acid residues, which participate in the catalysis using solely the multiple alignment as the template. Crystal structures of CAR proteins are unfortunately not available yet. Crystal structures might provide a deeper insight into the molecular level of CAR catalysis and could support rational protein design. CAR-encoding genes are relatively long (approximately 3000–3500 bp) and there are no reports about random mutagenesis of CARs in the literature thus far. Most likely, this can be attributed to a lack of appropriate high-throughput screening methods. The knowledge of key residues for carboxylate reduction in CARs is not only essential to understand the mechanism on the molecular level, but also for rational engineering of CARs to generate enzymes with improved properties, or efficiency for industrial applications.

In 2014, an NRPS-like protein from *Aspergillus terreus* was reported to reduce aryl acids to aryl aldehydes (Table 1, entry 3) (Wang and Zhao, 2014), thus providing the first fungal CAR sequence. This was, to the best of our knowledge, followed shortly thereafter by three more fungal sequences with confirmed CAR activity (Li et al., 2016; Schwendenwein et al., 2016; Winkler and Winkler, 2016). When all experimentally confirmed CAR sequences with A-T-R architecture and broad substrate scope are aligned, a significantly more focused amino acid pattern becomes apparent (bold residues in Fig. 3 and Table 2), that seems to be inherent to the carboxylate reductase function in all E.C. 1.20.1.30 CARs. Four CAR fingerprint sequences are present in the A-domain (Table 2, entries 1–4). Entry 1 resembles core A3, entry 3 core A5 and entry 5 core A8, as defined by the group of Marahiel (Marahiel et al., 1997). When the A-domains of four specific NRPS proteins are aligned to the A-domains of CARs, a wide variety of residues occupy these positions. A 10-residue specificity code for broad substrate scope CARs can not be deduced by this alignment (Table S3).

We hypothesized that essential residues for the reduction of acids to aldehydes in CARs are among the residues P and H in motif P xx H (Table 2, entry 2), as well as the residues D, R, GE and E (Table 2, entry 4). Some of those residues may play a significant role in the proposed gate-keeper/substrate recognition function

Table 1
Overview of enzymes that catalyze the reduction of a carboxylic acid to the respective aldehyde.

Entry	Substrates reduced	Enzyme Abbrev.	Accession Nr.	Origin	Domain architecture	References
1	>100 compounds (benzoic acid and derivatives, phenylacetic acid and derivatives, phenylpropanoic and propenoic acids and derivatives, acids with heterocycles or polycycles, aliphatic acids etc)	*Ni*CAR	WP_012393886	Bacteria	A-T-R	He et al. (2004), references in Napora-Wijata et al. (2014), (Finnigan et al. (2016), Kunjapur et al. (2014) and Moura et al. (2016)
2	>30 compounds (aliphatic acids, keto acids etc)	*Mm*CAR	Q6RKB1.1	Bacteria	A-T-R	Akhtar et al. (2013), Fraatz et al. (2016) and France et al. (2016)
3	Benzoic acid and derivatives	*At*CAR	XP_001212808.1	Ascomycete	A-T-R	Wang et al. (2014) and Wang and Zhao (2014)
4	>26 compounds (benzoic acid and derivatives, phenylacetic acid and derivatives, phenylpropanoic and propenoic acids and derivatives, aliphatic acids	*Sr*CAR	WP_013138593.1	Bacteria	A-T-R	Duan et al. (2015a)
5	>22 compounds (benzoic acid and derivatives, phenylacetic acid and derivatives, phenylpropanoic and propenoic acids and derivatives, aliphatic acids)	*Mn*CAR	WP_019510583.1	Bacteria	A-T-R	Duan et al. (2015b)
6	Benzoic acid, butyric acid, 2-methylbutyric acid	*Nb*CAR	AFU02004.1	Bacteria	A-T-R	Moura et al. (2016)
7	Benzoic acid, butyric acid, 2-methylbutyric acid, 2-oxo-butyric acid, acetic acid	*Ms*CAR1	WP_015306631.1	Bacteria	A-T-R	Moura et al. (2016)
8	>9 compounds (benzoic acid and derivatives, phenylpropanoic anc propenoic acid, aliphatic acids)	*Tv*CAR	XP_008043822.1	Basisdiomycete	A-T-R	Winkler and Winkler (2016)
9	2,4-dihydroxy-3-((2*E*,6*E*)-3,7,11-trimethyldodeca-2,6,10-trien-1-yl)benzoic acid	*Sb*CAR	BAV19380.1	Ascomycete	A-T-R	Li et al. (2016)
10	>25 compounds (benzoic acid and derivatives, phenylacetic acid and derivatives, phenylpropanoic and propenoic acids and derivatives, pyridine-2-carboxylic acid, aliphatic acids)	*Nc*CAR	XP_955820.1	Ascomycete	A-T-R	Schwendenwein et al. (2016)
11	>11 compounds (benzoic acid and derivatives, phenylpropanoic and propenoic acids and derivatives, acids with heterocycles, aliphatic acids)	*Mp*CAR	WP_003889896.1	Bacteria	A-T-R	Finnigan et al. (2016)
12	>14 compounds (benzoic acid and derivatives, phenylpropanoic and propenoic acids and derivatives, acids with heterocycles, aliphatic acids)	*Ms*CAR2	AFP42026.1	Bacteria	A-T-R	Finnigan et al. (2016)
13	>14 compounds (benzoic acid and derivatives, phenylpropanoic and propenoic acids and derivatives, acids with heterocycles, aliphatic acids)	*No*CAR	WP_029928026.1	Bacteria	A-T-R	Finnigan et al. (2016)
14	>16 compounds (benzoic acid and derivatives, phenylpropanoic and propenoic acids and derivatives, acids with heterocycles, aliphatic acids)	*Tp*CAR	WP_013126039.1	Bacteria	A-T-R	Finnigan et al. (2016)
15	Tyrosine	LnaA		Ascomycete	A-T-R	Forseth et al. (2013)
16	Tyrosine	LnbA		Ascomycete	A-T-R	Forseth et al. (2013)
17	L-aminoadipate	NPS3		Basidiomycete	X-A-T-R	Kalb et al. (2015)

of the A-domain (Moura et al., 2016). In the reductase domain, a lone phenylalanine and the residues R, G and N (Table 2, entry 8) are most likely to be involved in substrate binding. The Y and K of the motif GY xx K (Table 2, entry 9) have been recognized earlier to be characteristic for tyrosine-dependent oxidoreductases of the short chain dehydrogenase family (Chhabra et al., 2012) and are expected to be essential for the thioester reduction step in CARs. In order to understand which residues contribute to the catalytic steps, *Nc*CAR was modeled using Phyre2 (http://www.sbg.bio.ic.ac.uk/phyre2). Eleven residues pointing into a cavity with the potential ATP binding site in the A-domain, as well as six residues pointing into a cavity with the potential NADPH binding site in the R-domain were selected for mutation against alanine. Wild-type and mutants were expressed as described (Schwendenwein et al., 2016) and soluble expression was analysed by SDS-PAGE. Mutation of G184 and R870 to alanine, respectively, was detrimental to soluble expression (Fig. 4). Wild-type *Nc*CAR and the fifteen soluble mutants were purified by Ni-affinity chromatography and their activity was assessed using a spectrophotometric assay (Schwendenwein et al., 2016). The exchange of two amino acid residues in the R-domain (Y844 and K848) abolished CAR activity for the reduction of cinnamic acid completely (Fig. 5). Similarly,

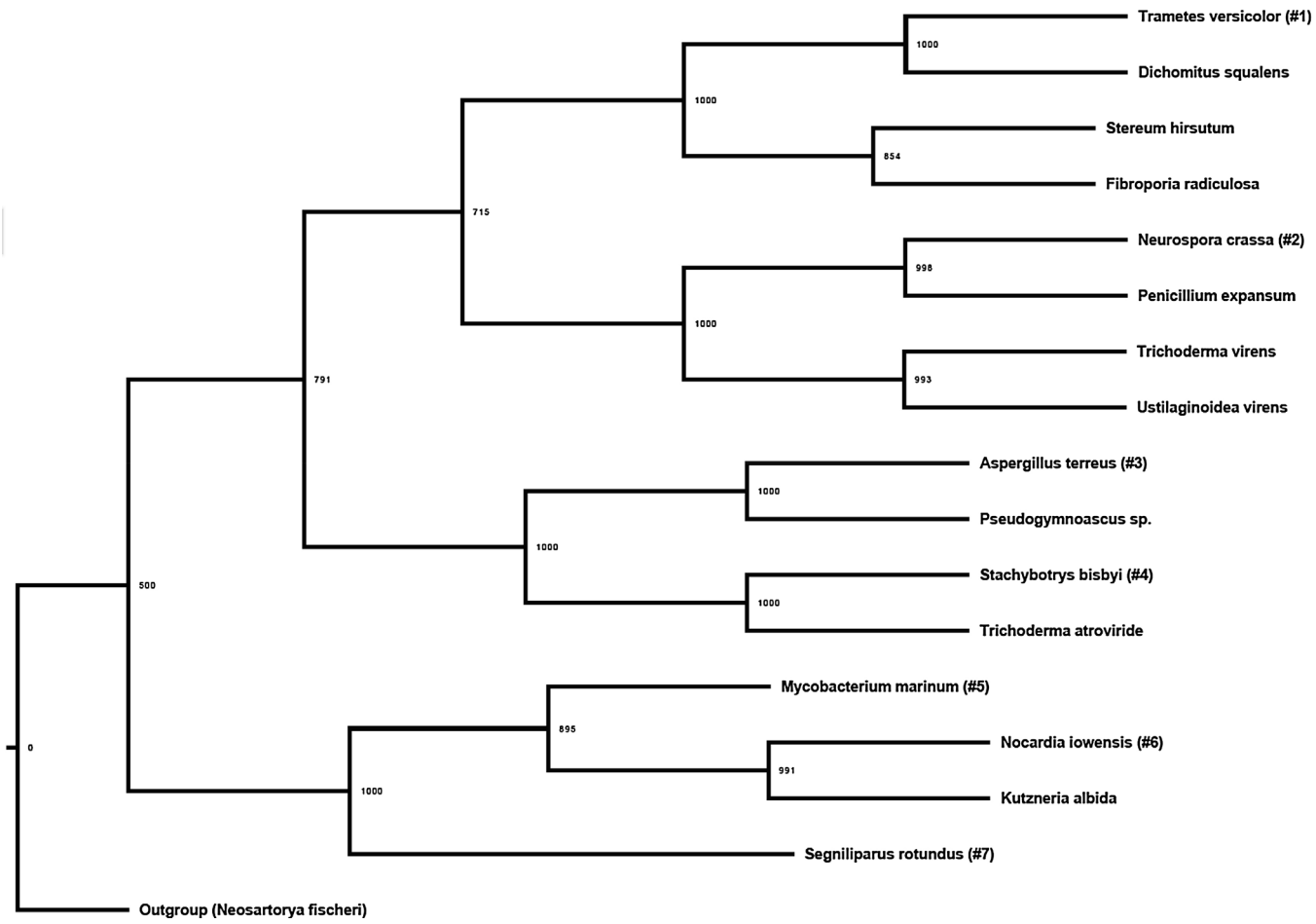

Fig. 2. Phylogenetic tree, showing distribution of carboxylate reductases based on a protein distance analysis with Phylip Version 3.696. The four groups are each supported by a bootstrap value of 100% (*i.e.* 1000 delete-half Jacknife replicates forming the same topology in each subtree). For GenBank accession numbers of the sequences used in the tree, see Table S1. Bacterial CAR: type I (represented by sequences #5-7), *At*CAR/*Sb*CAR: type II (represented by sequence #3 and 4), *Nc*CAR: type III (represented by sequence #2), and *Tv*CAR: type IV (represented by sequence #1).

Table 2
Highly conserved signature sequences in the catalytic domains of CARs (bold). Residues in italics: Exchange for alanine results in insoluble expression; Underlined residues: Exchange for alanine results in loss of carboxylate reductase activity.

entry	Domain	Consensus sequence	Predicted function
1	A	S/T **SG** S/T **T** G/S x **PK**	Phosphate binding
2	A	**P** xx **H**	Substrate binding
3	A	Y/F/I G/S x **TE** x G/A/T	AMP binding
4	A	**D** x_{14-17} **R** xx N/D/E x L/I x L/F A/S x **GE** K/F x_5 L/I **E**	AMP binding, substrate binding
5	T	**G** x D/N **S** L/M	Phosphopantetheine binding site
6	R	**V** L/V/I L/I/V **TG** xx **G** xx **G**	NADP(H) binding
7	R	**GY** xxx **K** xxx E/S	NADP(H) binding
8	R	***R*** x_{11} **G** xx **N**	Substrate binding

three residues in the A-domain (H237, E337 and E433) reduced activity to a non-significant level (Fig. 5). The exchange of T186, P189 and G432 for alanine resulted in wild-type level activities. Interestingly, mutations K190A and P234A resulted in increased activity for the reduction of cinnamic acid.

3.5. Motif analysis

The motif search revealed AMP-binding motifs in all CAR subgroups and also, consistently throughout all sequences, an AMP-binding domain signature (Fig. 3). Type I CARs are additionally characterized by an AMP-binding enzyme C-terminal domain (Pfam PF13193), which is not present in fungal CARs of type II–IV. Phosphopantetheine attachment sites are rather short stretches with low similarity among the four types. The *Nc*CAR sequence is the only one, for which a PROSITE pattern with the function "Phosphopantetheine attachment sites" was recognized. In alignments with other CARs, only the phosphopantetheine binding serine and a glycine are conserved within the 16 amino acid stretch of the PROSITE pattern. The NAD-binding-4 domain in the R-domains of all CARs begin with the typical G xx G xx G motif of Rossman fold proteins (Kavanagh et al., 2008). Again, bacterial CARs (type I) are characterized by a unique short-chain dehydrogenase/reductase family signature (PROSITE Pattern) which is not present in fungal CARs.

4. Conclusions and outlook

In conclusion, the detailed analysis of E.C.1.2.1.30 carboxylate reductase enzymes (CARs) allows us to propose a CAR amino acid signature and we also suggest a subgrouping into four CAR types,

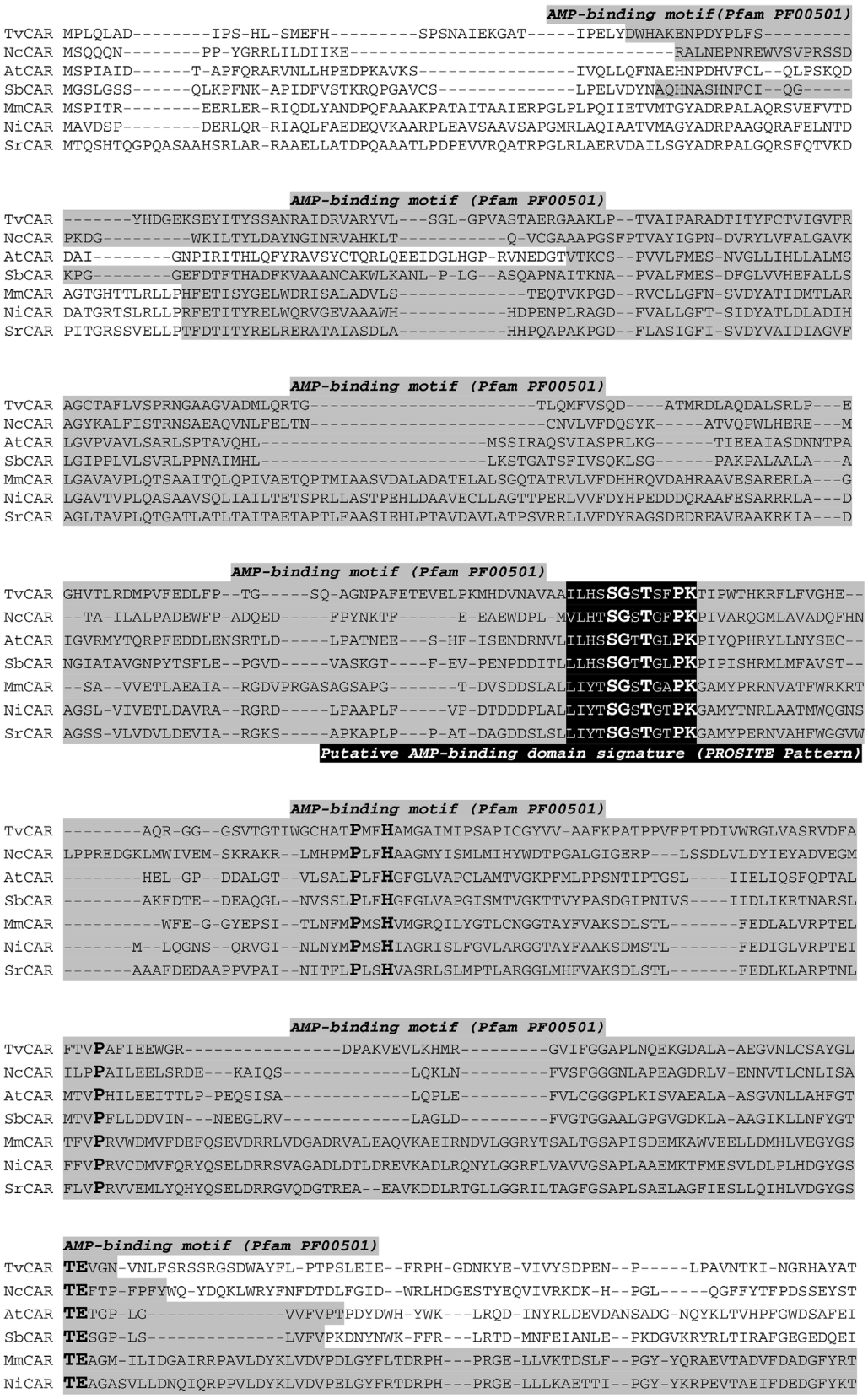

Fig. 3. Multiple sequence alignment of confirmed CARs of each CAR-type as shown in Fig. 2. Bold letters indicate residues conserved in all four CAR types. The phopspho-pantetheinyl attachment residue (serine) is shown in bold and italics.

Fig. 3. (*Continued*)

```
NcCAR RWH----DEDR-IVPEASLDDLS-LAAGGYGQSKLVSSLIFDKAA-EVS-GVPTEVVRVGQVAGPSSEKGYWNKQEWLPS
AtCAR QYPRT---HGTRLIPEVPSDKSSIIEDFGYGKAKYVCEEIMRAAA-DRYPEMQLGIVRVGQMSG-SSRTGYWNPKEHFPT
SbCAR RLSLS---NSGAMISETPVEPVDAACGIGYADGKLVCEKILEKAAVSHAGQLEIAYVRCGQMTG-SRATGAWNADEQIPM
MmCAR QIPPSAFTEDADIRVISATRAVDDSYANGYSNSKWAGEVLLREAH-DLC-GLPVAVFRCDMILADTTWAGQLNVPDMFTR
NiCAR QVDPAEYQEDSDVREMSAVRVVRESYANGYGNSKWAGEVLLREAH-DLC-GLPVAVFRSDMILAHSRYAGQLNVQDVFTR
SrCAR GVEPSALDEDGDIRTVSAERSVDEGYANGYGNSKWGGEVLLREAH-DRT-GLPVRVFRSDMILAHQKYTGQVNATDQFTR
Short-chain dehydrogenase/reductase family signature (PROSITE Pattern)

          NAD_binding_4 (Pfam PF07993)
TvCAR LVKSAQFQRCLPDIE----------GSVSWIPGYEAAKAFTEMRHS-PHP--------FLHLVHPKPVS--WHTLISAI
NcCAR IVASSAYLGVLPDSLGQ--------MTTIDWTPIEAIAKLLLEVSGV-IDNVPLDKINGYFHGVNPERTS--WSALAPAV
AtCAR LIKFASMVGQLPAIK----------QTLSWIAVDNAATVLSDILFA-PSLS------GIYHLENPIRQA--WQDVLDIF
SbCAR IFRTAKNLGVLPRIP----------GTLSWIPVDDAAQYIMDLSFFEGALP------IASHLENPVRQS--WADLMDGA
MmCAR MILSLAATGIAPGSFYELAADGARQRAHYDGLPVEFIAEAISTLGAQSQDGF------HTYHVMNPYDDGIGLDEFVDWL
NiCAR LILSLVATGIAPYSFYRTDADGNRQRAHYDGLPADFTAAAITALGIQATEGF------RTYDVLNPYDDGISLDEFVDWL
SrCAR LVQSLLATGLAPKSFYELDAQGNRQRAHYDGIPVDFTAESITTLGGDGLEGY------RSYNVFNPHRDGVGLDEFVDWL

TvCAR AK-E-FGNVPLVPYDEWLSALQASVSEGDAAEVELMRANPALRLLP-FFQAVNQHASLDR-EPLGLVYL-----STEKSA
NcCAR QEYYGDRIQKIVPLDEWLEALEKSQEKAE--D-----VTRNPG-I-KLIDTYRTW---SEGYKKGTKFVPLDMTRTK---
AtCAR AS-SLYINTVNVPFDQWLRNVQAAVQELGTED-----ERMEYDLLAEFLEKD--F---QR-MATGKVILDTSR-----SR
SbCAR GK-F-LGIQKSVSWPEWLELAGAAEDGP--QD-----KYPVKKLF-AFFKFS--F---GP-MASGAVILGTDV-----AR
MmCAR NE-SGCPIQRIADYGDWLQRFETALRALP--D-----RQRHSSLL-PLLHNYRQP---ER-PVRGSI-APTDRFRAAVQE
NiCAR VE-SGHPIQRITDYSDWFHRFETAIRALP--E-----KQRQASVL-PLLDAYRNP---CP-AVRGAI-LPAKEFQAAVQT
SrCAR IE-AGHPITRIDDYDQWLSRFETSLRGLP--E-----SKRQASVL-PLLHAFARP---GP-AVDGSP-FRNTVFRTDVQK

TvCAR AVSGALANLPQLDAERAKGWLAAWKSAGF-----L
NcCAR EYSKTMREMHAVTPELMKNWCRQW----N-----F
AtCAR AVSETLREVGEISEEVVWKYVREWRRAGTLRAPLE
SbCAR AHSATIKNMGALDASTIMKYFLHWQKINY-----L
MmCAR AKIGPDKDIPHVGAPIIVKYVSDLRLLGL-----L
NiCAR AKIGPEQDIPHLSAPLIDKYVSDLELLQL-----L
SrCAR AKIGAEHDIPHLGKALVLKYADDIKQLGL-----L
```

Fig. 3. (*Continued*)

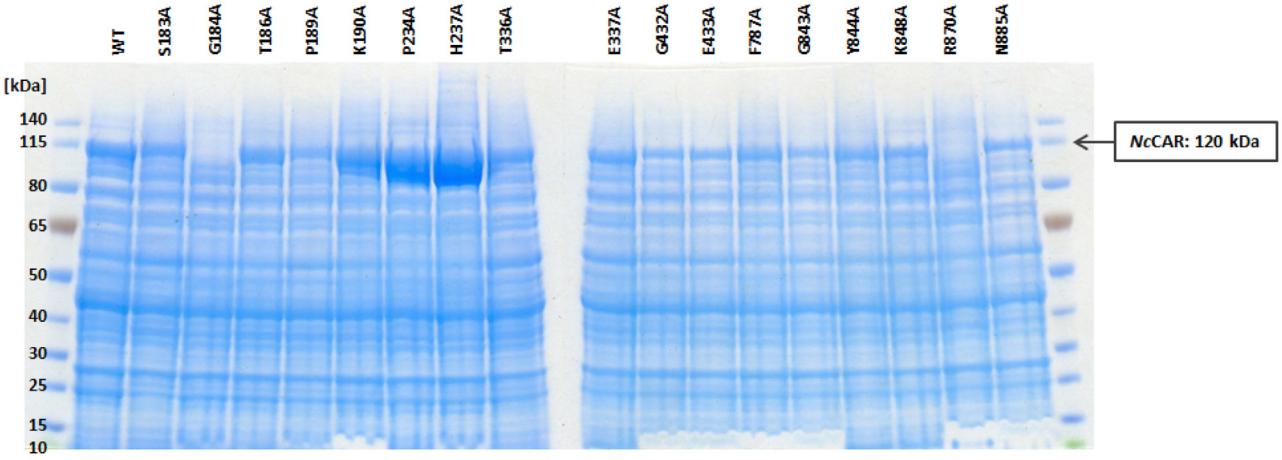

Fig. 4. SDS-PAGE of *Nc*CAR wild-type and mutant proteins. Ladder: PageRuler™ prestained Protein Ladder (Thermo Scientific).

based on a bootstrapped distance analysis with 100% bootstrap value for each of the four types. We are convinced that three fungal and one bacterial CAR types can be distinguished based on this topology, respectively. We are currently aiming to increase the understanding of this enzyme class in general and another important milestone in this direction will be the investigation of mutants of the CAR signature residues, which might reveal the real significance of each of the residues contained in the signature sequence in the future. The identification of the CAR active sites and a better understanding of the CAR signature sequence will assist in the identification of new carboxylate reductase enzymes in the future and their application in biotechnological processes.

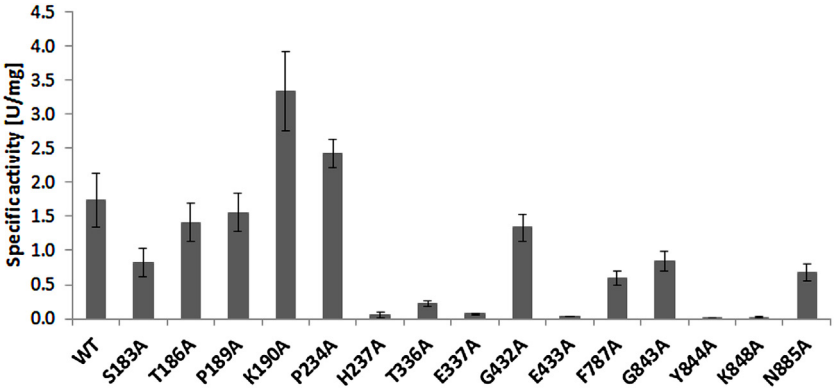

Fig. 5. Activity of *Nc*CAR wild-type and mutants for the reduction of cinnamic acid. The results represent 3–5 repetitions of each experiment in four technical replicates.

Acknowledgements

The skillful technical assistance of K. Rudnicka is gratefully acknowledged. The Austrian science fund FWF is kindly acknowledged for financial support (Elise-Richter fellowship V415-B21 and P28477-B21). This research has also been supported by the Austrian BMWFW, BMVIT, SFG, Standortagentur Tirol, Government of Lower Austria and ZIT through the Austrian FFG-COMET- Funding Program. We thank the COST action Systems Biocatalysis WG2 and TU Wien ABC-Top Anschubfinanzierung for financial support.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jbiotec.2017.02.014.

References

Agler, M.T., Wrenn, B.A., Zinder, S.H., Angenent, L.T., 2011. Waste to bioproduct conversion with undefined mixed cultures: the carboxylate platform. Trends Biotechnol. 29, 70–78.

Akhtar, M.K., Turner, N.J., Jones, P.R., 2013. Carboxylic acid reductase is a versatile enzyme for the conversion of fatty acids into fuels and chemical commodities. Proc. Natl. Acad. Sci. U. S. A. 110, 87–92.

Brandenburger, E., Braga, D., Kombrink, A., Lackner, G., Gressler, J., Künzler, M., Hoffmeister, D., 2016. Multi-genome analysis identifies functional and phylogenetic diversity of basidiomycete adenylate-forming reductases. Fungal Genet. Biol., http://dx.doi.org/10.1016/j.fgb.2016.07.008.

Brenna, E., Cannavale, F., Crotti, M., Parmeggiani, F., Romagnolo, A., Spina, F., Varese, G.C., 2015. Biocatalysed reduction of carboxylic acids to primary alcohols in aqueous medium: a novel synthetic capability of the zygomycete fungus *Syncephalastrum racemosum*. J. Mol. Catal. B Enzym. 116, 83–88.

Bushley, K.E., Ripoll, D.R., Turgeon, B.G., Helm Van der, D., Winkelmann, G., et al., 2008. Module evolution and substrate specificity of fungal nonribosomal peptide synthetases involved in siderophore biosynthesis. BMC Evol. Biol. 8, 328.

Carroll, A.L., Desai, S.H., Atsumi, S., 2016. Microbial production of scent and flavor compounds. Curr. Opin. Biotechnol. 37, 8–15.

Chhabra, A., Haque, A.S., Pal, R.K., Goyal, A., Rai, R., Joshi, S., Panjikar, S., Pasha, S., Sankaranarayanan, R., Gokhale, R.S., 2012. Nonprocessive [2+2]e- off-loading reductase domains from mycobacterial nonribosomal peptide synthetases. Proc. Natl. Acad. Sci. U. S. A. 109, 5681–5686.

da S. Amaral, L., Rodrigues-Filho, E., 2015. Aryl carboxylic acid reduction and further reactions with GABA and glucose promoted by whole cells of *Xylaria arbuscula*. J. Mol. Catal. B Enzym. 113, 90–94.

Duan, Y., Yao, P., Chen, X., Liu, X., Zhang, R., Feng, J., Wu, Q., Zhu, D., 2015a. Exploring the synthetic applicability of a new carboxylic acid reductase from *Segniliparus rotundus* DSM 44985. J. Mol. Catal. B Enzym. 115, 1–7.

Duan, Y., Yao, P., Du, Y., Feng, J., Wu, Q., Zhu, D., 2015b. Synthesis of α, β-unsaturated esters via a chemo-enzymatic chain elongation approach by combining carboxylic acid reduction and Wittig reaction. Beilstein J. Org. Chem. 11, 2245–2251.

Finnigan, W., Thomas, A., Cromar, H., Gough, B., Snajdrova, R., Adams, J.P., Littlechild, J.A., Harmer, N., 2016. Characterization of carboxylic acid reductases as enzymes in the toolbox for synthetic chemistry. ChemCatChem, http://dx.doi.org/10.1002/cctc.201601249.

Forseth, R.R., Amaike, S., Schwenk, D., Affeldt, K.J., Hoffmeister, D., Schroeder, F.C., Keller, N.P., 2013. Homologous NRPS-like gene clusters mediate redundant small-molecule biosynthesis in *Aspergillus flavus*. Angew. Chem. Int. Ed. Engl. 52, 1590–1594.

Fraatz, M.A., Goldmann, M., Geissler, T., Gross, E., Backes, M., Hilmer, J.-M., Ley, J., Rost, J., Francke, A., Zorn, H., 2016. Biotechnological production of methyl-branched aldehydes. J. Agric. Food Chem., http://dx.doi.org/10.1021/acs.jafc.6b04793.

France, S.P., Hussain, S., Hill, A.M., Hepworth, L.J., Howard, R.M., Mulholland, K.R., Flitsch, S.L., Turner, N.J., 2016. One-pot cascade synthesis of mono- and disubstituted piperidines and pyrrolidines using carboxylic acid reductase (CAR), ω-transaminase ω(-TA), and imine reductase (IRED) biocatalysts. ACS Catal. 6, 3753–3759.

Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., Smith, H.O., 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods 6, 343–345.

Guo, S., Bhattacharjee, J.K., 2004. Posttranslational activation, site-directed mutation and phylogenetic analyses of the lysine biosynthesis enzymes alpha-aminoadipate reductase Lys1p (AAR) and the phosphopantetheinyl transferase Lys7p (PPTase) from *Schizosaccharomyces pombe*. Yeast 21, 1279–1288.

He, A., Li, T., Daniels, L., Fotheringham, I., Rosazza, J.P.N., 2004. Nocardia sp. carboxylic acid reductase: cloning, expression, and characterization of a new aldehyde oxidoreductase family. Appl. Environ. Microbiol. 70, 1874–1881.

Kalb, D., Lackner, G., Hoffmeister, D., 2014. Functional and phylogenetic divergence of fungal adenylate-forming reductases. Appl. Environ. Microbiol. 80, 6175–6183.

Kalb, D., Lackner, G., Rappe, M., Hoffmeister, D., 2015. Activity of α-aminoadipate reductase depends on the *N*-terminally extending domain. ChemBioChem 16, 1426–1430.

Kavanagh, K.L., Jörnvall, H., Persson, B., Oppermann, U., 2008. Medium- and short-chain dehydrogenase/reductase gene and protein families. Cell. Mol. Life Sci. 65, 3895–3906.

Kunjapur, A.M., Tarasova, Y., Prather, K.L.J., 2014. Synthesis and accumulation of aromatic aldehydes in an engineered strain of Escherichia coli. J. Am. Chem. Soc. 136, 11644–11654.

Li, C., Matsuda, Y., Gao, H., Hu, D., Yao, X.S., Abe, I., 2016. Biosynthesis of LL-Z1272β: discovery of a new member of NRPS-like enzymes for aryl-aldehyde formation. ChemBioChem 17, 904–907.

Marahiel, M.A., Stachelhaus, T., Mootz, H.D., 1997. Modular peptide synthetases involved in nonribosomal peptide synthesis. Chem. Rev. 97, 2651–2673.

Marchler-Bauer, A., Bryant, S.H., 2004. CD-Search: protein domain annotations on the fly. Nucleic Acids Res. 32, W327–W331.

Moura, M., Pertusi, D., Lenzini, S., Bhan, N., Broadbelt, L.J., Tyo, K.E.J., 2016. Characterizing and predicting carboxylic acid reductase activity for diversifying bioaldehyde production. Biotechnol. Bioeng. 113, 944–952.

Napora-Wijata, K., Strohmeier, G.A., Winkler, M., 2014. Biocatalytic reduction of carboxylic acids. Biotechnol. J. 9, 822–843.

Perez, J.M., Richter, H., Loftus, S.E., Angenent, L.T., 2013. Biocatalytic reduction of short-chain carboxylic acids into their corresponding alcohols with syngas fermentation. Biotechnol. Bioeng. 110, 1066–1077.

Schrittwieser, J.H., Velikogne, S., Kroutil, W., 2015. Biocatalytic imine reduction and reductive amination of ketones. Adv. Synth. Catal. 357, 1655–1685.

Schwendenwein, D., Fiume, G., Weber, H., Rudroff, F., Winkler, M., 2016. Selective enzymatic transformation to aldehydes *in vivo* by fungal carboxylate reductase from *Neurospora crassa*. Adv. Synth. Catal. 358 (21), 3414–3421.

Sell, C., 2006. The Chemistry of Fragrances: From Perfumer to Consumer, 2nd ed. RSC Publishing.

Stachelhaus, T., Mootz, H.D., Marahiel, M.A., 1999. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. Chem. Biol. 6, 493–505.

Studier, F.W., 2005. Protein production by auto-induction in high-density shaking cultures. Protein Expr. Purif. 41, 207–234.

Toogood, H.S., Scrutton, N.S., 2014. New developments in ene-reductase catalysed biological hydrogenations. Curr. Opin. Chem. Biol. 19, 107–115.
van Langen, L.M., van Rantwijk, F., Sheldon, R.A., 2003. Enzymatic Hydrocyanation of a Sterically Hindered Aldehyde. Optimization of a Chemoenzymatic Procedure for (*R*)-2-Chloromandelic Acid. Org. Proc. Res. Devel. 7, 828–831.
Venkitasubramanian, P., Daniels, L., Rosazza, J.P.N., 2007. Reduction of carboxylic acids by *Nocardia* aldehyde oxidoreductase requires a phosphopantetheinylated enzyme. J. Biol. Chem. 282, 478–485.
Venkitasubramanian, P., Daniels, L., Das, S., Lamm, A.S., Rosazza, J.P.N., 2008. Aldehyde oxidoreductase as a biocatalyst: reductions of vanillic acid. Enzyme Microb. Technol. 42, 130–137.
Wang, M., Zhao, H., 2014. Characterization and engineering of the adenylation domain of a NRPS-like protein: a potential biocatalyst for aldehyde generation. ACS Catal. 4, 1219–1225.
Wang, M., Beissner, M., Zhao, H., 2014. Aryl-aldehyde formation in fungal polyketides: discovery and characterization of a distinct biosynthetic mechanism. Chem. Biol. 21, 257–263.
Winkler, M., Winkler, C.K., 2016. *Trametes versicolor* carboxylate reductase uncovered. Chem. Mon. 147, 575–578.