



Identification of Key Residues for Enzymatic Carboxylate Reduction

Holly Stolterfoht^{1,2}, Georg Steinkellner^{2,3}, Daniel Schwendenwein², Tea Pavkov-Keller^{2,3}, Karl Gruber^{2,3} and Margit Winkler^{1,2*}

¹ Institute of Molecular Biotechnology, Graz University of Technology, NAWI Graz, Graz, Austria, ² Austrian Centre of Industrial Biotechnology, Graz, Austria, ³ Structural Biology, Institute of Molecular Biosciences, University of Graz, Graz, Austria

OPEN ACCESS

Edited by:

Caroline Emilie Paul,
Wageningen University & Research,
Netherlands

Reviewed by:

Ligia O. Martins,
Instituto de Tecnologia Química e
Biológica, Universidade Nova de
Lisboa, Portugal
Peter-Leon Hagedoorn,
Delft University of Technology,
Netherlands

*Correspondence:

Margit Winkler
margit.winkler@tugraz.at

Specialty section:

This article was submitted to
Microbiotechnology, Ecotoxicology
and Bioremediation,
a section of the journal
Frontiers in Microbiology

Received: 23 November 2017

Accepted: 31 January 2018

Published: 19 February 2018

Citation:

Stolterfoht H, Steinkellner G,
Schwendenwein D, Pavkov-Keller T,
Gruber K and Winkler M (2018)
Identification of Key Residues for
Enzymatic Carboxylate Reduction.
Front. Microbiol. 9:250.
doi: 10.3389/fmicb.2018.00250

Carboxylate reductases (CARs, E.C. 1.2.1.30) generate aldehydes from their corresponding carboxylic acid with high selectivity. Little is known about the structure of CARs and their catalytically important amino acid residues. The identification of key residues for carboxylate reduction provides a starting point to gain deeper understanding of enzymatic carboxylate reduction. A multiple sequence alignment of CARs with confirmed activity recently identified in our lab and from the literature revealed a fingerprint of conserved amino acids. We studied the function of conserved residues by multiple sequence alignments and mutational replacements of these residues. In this study, single-site alanine variants of *Neurospora crassa* CAR were investigated to determine the contribution of conserved residues to the function, expressability or stability of the enzyme. The effect of amino acid replacements was investigated by analyzing enzymatic activity of the variants *in vivo* and *in vitro*. Supported by molecular modeling, we interpreted that five of these residues are essential for catalytic activity, or substrate and co-substrate binding. We identified amino acid residues having significant impact on CAR activity. Replacement of His 237, Glu 433, Ser 595, Tyr 844, and Lys 848 by Ala abolish CAR activity, indicating their key role in acid reduction. These results may assist in the functional annotation of CAR coding genes in genomic databases. While some other conserved residues decreased activity or had no significant impact, four residues increased the specific activity of *NcCAR* variants when replaced by alanine. Finally, we showed that *NcCAR* wild-type and mutants efficiently reduce aliphatic acids.

Keywords: carboxylate reductase, biocatalysis, signature sequence, aldehyde, flavor and fragrance, pharmaceutical intermediate

INTRODUCTION

Aldehydes are widely used as flavors and fragrances (Hagedorn and Kaphammer, 1994; Lesage-Meessen et al., 1996; Thibault et al., 1998; Kunjapur et al., 2014), chemicals and precursors for pharmaceuticals (Tripathi et al., 1997). Furthermore, they are precursors for fatty alcohols, which, in turn, can be used as biofuels (Akhtar et al., 2013), polymer constituents, surfactants, detergents, and cosmetics (Noweck and Grafahrend, 2006). Enzymatic carboxylate reduction is catalyzed either by ATP/NADPH dependent proteins of the family E.C. 1.2.1.30 [originally termed aryl-aldehyde dehydrogenase (NADP⁺); now known as carboxylate reductase (CAR)] or by molybdenum or tungsten dependent proteins of the family E.C. 1.2.99.6 (termed aldehyde:ferredoxin oxidoreductases or also carboxylate reductase) (White et al., 1989, 1993; Mukund and Adams, 1991; Huber et al., 1994; Ni et al., 2012). Carboxylate reductases (E.C. 1.2.1.30) are able to

generate various aldehydes with high selectivity from their corresponding carboxylic acid, which can be available from natural and renewable resources (Yang et al., 2007). Generally, enzymatic carboxylate reduction provides several advantages compared to chemical reductions: The reactions are carried out in one step under mild conditions in aqueous solutions. CARs accept a broad range of substrates, while being product-selective (no overreduction of the aldehydes to the respective alcohols), chemo-selective (no reduction of other reducible moieties), and to some degree enantioselective (Napora-Wijata et al., 2014; Qu et al., 2018; Winkler, 2018).

Due to these characteristics and the growing demand for bio-based products from renewable resources, CARs are gaining attention as biocatalysts for chemical synthesis.

However, only a few enzyme sequences have been identified and characterized to date (Table 1). The limited number of known enzymes is surprising in light of the plethora of fungi and bacteria that can perform carboxylate reduction (Napora-Wijata et al., 2014).

CARs show significant similarities to non-ribosomal peptide synthases (NRPS), members of the ANL (Acyl-CoA synthetase/NRPS adenylation domain/Luciferase) superfamily of adenylation enzymes (Gulick, 2009), as they are comprised of an N-terminal adenylation domain (A-domain), a phosphopantetheine attachment site, also called transthiolation domain (T-domain), or peptidyl carrier protein (PCP) domain, and a C-terminal reductase domain (R-domain). On the molecular level, a cascade reaction proceeds: First, deprotonated carboxylic acid is activated as adenosyl monophosphate at the expense of ATP at the A-domain. The AMP-intermediate is nucleophilically attacked by sulfhydryl of a phosphopantetheinyl moiety and, with the release of AMP, a covalently bound thioester is formed. Finally, the thioester is reduced by NADPH to release the aldehyde product at the R-domain (Gross, 1972; Li and Rosazza, 1997).

To obtain active enzyme, CARs require a post-translational phosphopantetheinylation catalyzed by a phosphopantetheine transferase (PPTase), that transfers a phosphopantetheine moiety from coenzyme A to a conserved serine residue of apo-CAR (Venkitasubramanian et al., 2007), which provides the sulfhydryl group mentioned above. In addition, like other ATP dependent enzymes, CARs are dependent on Mg^{2+} (Gross and Zenk, 1969).

Recently, first crystal structures of the individual (di)domains of *Sru*CAR, the A-domain of *Ni*CAR, and R-domain of *Mm*CAR shed some light on the complex mechanism underlying catalysis and revealed domain movements during the different steps of the reaction (Gahlth et al., 2017).

The N-terminus of CARs shows high homology to known AMP-binding proteins with a similar reaction mechanism to that of the ANL superfamily of adenylation enzymes, whereas the C-terminal domain strongly resembles NADPH-binding proteins, like terminal reductase domains of NRPS or short-chain dehydrogenases/reductases (SDRs). A-domains of NRPS enzymes have been studied in detail (Marahiel et al., 1997), key residues identified, and selectivity conferring codes of these highly substrate specific A-domains have been defined (Stachelhaus et al., 1999; Bushley et al., 2008). For SDRs,

conserved sequence motifs, and functionally important residues have been extensively characterized as well (Persson et al., 2003; Kavanagh et al., 2008; Chhabra et al., 2012), but CARs are much less thoroughly studied in this regard.

In this study, we used a CAR from the ascomycete *Neurospora crassa* (*Nc*CAR). *Nc*CAR was first isolated from the fungus (strain SY7A) in 1968 (Gross et al., 1968). Notably, the overall domain organization of *Nc*CAR resembles that of type I bacterial CARs, however, sequence identities are very low (~20%) and bacterial CARs are characterized by an A-domain stretch that is not found in type II-IV CARs (Stolterfoht et al., 2017). Consequently, detailed structural information is not easily transferable from one subtype to the other.

The substrate scope of *Nc*CAR determined in the 1970s led to the historic classification of this enzyme as an aryl-aldehyde:NADP oxidoreductase. Recently, the genome of *N. crassa* OR74A was sequenced (Galagan et al., 2003) and we identified the hypothetical protein with the NCBI accession no. XP_955820.1 from this strain to have CAR activity. The heterologously expressed enzyme was further characterized and used for the preparative scale synthesis of the aromatic aldehyde piperonal (Schwendenwein et al., 2016). In contrast to the observations concerning the substrate scope of *Nc*CAR (Gross and Zenk, 1969; Gross, 1972), we show herein that recombinant *Nc*CAR is not restricted to aromatic acids: *Nc*CAR can also reduce aliphatic acids.

MATERIALS AND METHODS

Chemicals and Solvents

ATP was obtained from Sigma-Aldrich. 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.

Multiple Sequence Alignments

Twenty-six confirmed CAR sequences that have been characterized to date (Table 1) were aligned using T-Coffee (<http://www.tcoffee.org>) as well as ClustalW (<http://www.genome.jp/tools-bin/clustalw>) with a gap opening penalty of 25 and gap extension penalty of 0.2. The alignment programs were run on the command line and visualized using Jalview (<http://www.jalview.org>). Additionally, alignments were generated using the program CLC Main Workbench Alignment with default settings. Only slight differences between the alignment tools were encountered, whereas highly conserved fractions of sequences were identical, independent of the alignment program. Conserved sequences presented in Table 2 are based on alignments using T-Coffee.

Modeling of *Nc*CAR-FL

The program YASARA structure (v. 16.6.24, Krieger et al., 2002) was used to build the models of the A-T-domain and the full length *Nc*CAR (*Nc*CAR-FL). The target sequence was used to perform a BLAST (Basic Local Alignment Search Tool) search against the UniProt database (PS-BLAST iterations 1, E-value cutoff 0.01) to build position-specific scoring matrices (PSSM)

TABLE 1 | Known CAR enzymes that have been characterized to date.

Entry	Enzyme Abbrev.	Organism	Accession No.	Phylogenetic group	References
1	NiCAR	<i>Nocardia iowensis</i>	Q6RKB1.1	type I	Li and Rosazza, 1997; He et al., 2004
2	MsCAR2	<i>Mycobacterium smegmatis</i> str. MC2 155	YP_889972.1	type I	Hu, 2010; Khusnutdinova et al., 2017
3	MtCAR	<i>Mycobacterium tuberculosis</i>	NP_217106.1	type I	Hu, 2010
4	MspCAR1	<i>Mycobacterium</i> sp. JLS	WP_011855500.1	type I	Behrouzian et al., 2010
5	MmCAR	<i>Mycobacterium marinum</i>	WP_012393886.1	type I	Akhtar et al., 2013
6	MsCAR1	<i>Mycobacterium smegmatis</i> str. MC2 155	YP_887275.1	type I	Schaffer et al., 2014; Khusnutdinova et al., 2017
7	SroCAR	<i>Segniliparus rotundus</i>	WP_013138593.1	type I	Duan et al., 2015a
8	MnCAR	<i>Mycobacterium</i> sp. (neoaurum)	WP_019510583.1	type I	Duan et al., 2015b
9	MspCAR2	<i>Mycobacterium</i> sp. JS623	WP_015306631.1	type I	Moura et al., 2016
10	NbCAR1	<i>Nocardia brasiliensis</i>	AFU02004.1	type I	Moura et al., 2016
11	MpCAR	<i>Mycobacterium phlei</i>	WP_003889896.1	type I	Finnigan et al., 2017
12	MsCAR3	<i>Mycobacterium smegmatis</i> str. MC2 155	AFP42026.1	type I	Finnigan et al., 2017
13	NoCAR	<i>Nocardia otitidiscaviarum</i>	WP_029928026.1	type I	Finnigan et al., 2017
14	TpCAR	<i>Tsakamurella paurometabola</i>	WP_013126039.1	type I	Finnigan et al., 2017
15	SruCAR	<i>Segniliparus rugosus</i>	WP_007468889.1	type I	Gahlloth et al., 2017
16	MaCAR1	<i>Mycobacterium abscessus</i> ATCC 19977	CAM63040.1	type I	Khusnutdinova et al., 2017
17	MaCAR2	<i>Mycobacterium abscessus</i> ATCC 19977	CAM64782.1	type I	Khusnutdinova et al., 2017
18	MapCAR	<i>Mycobacterium avium</i> subsp. <i>paratuberculosis</i>	AAS03357	type I	Khusnutdinova et al., 2017
19	McCAR	<i>Mycobacterium chelonae</i>	AKC40871.1	type I	Khusnutdinova et al., 2017
20	MiCAR1	<i>Mycobacterium immunogenum</i>	KPG37443.1	type I	Khusnutdinova et al., 2017
21	MiCAR2	<i>Mycobacterium immunogenum</i>	KPG36834.1	type I	Khusnutdinova et al., 2017
22	NbCAR2	<i>Nocardia brasiliensis</i>	GAJ83027	type I	Khusnutdinova et al., 2017
23	AtCAR	<i>Aspergillus terreus</i>	XP_001212808.1	type II	Wang and Zhao, 2014; Wang et al., 2014
24	SbCAR	<i>Stachybotrys bisbyi</i>	BAV19380.1	type II	Li et al., 2016
25	NcCAR	<i>Neurospora crassa</i>	XP_955820.1	type III	Gross et al., 1968; Schwendenwein et al., 2016
26	TvCAR	<i>Trametes versicolor</i>	XP_008043822.1	type IV	Winkler and Winkler, 2016

from related sequences. This profile was used to search the Protein Database for potential modeling templates. Common protein purification tags were thereby excluded. The templates were ranked based on the alignment score and the structural quality. The NcCAR model was built using the top 10 scoring templates with a maximum oligomeric state setting of 1. None of the newly available CAR structures (Gahlloth et al., 2017) were on this list as they share lower sequence identities and alignment scores to the NcCAR sequence than the templates which were finally used. For each available template, the alignment with the target sequence was obtained using sequence-based profiles of target and templates calculated from related UniProt sequences. Ligands present in the template structures were parameterized and included in the homology modeling process. Side-chain rotamers were generated and accepted or rejected based on a repulsive energy function.

Loops were optimized using different conformations and side-chain re-optimizations (maximum of 50 conformations per loop). Side-chain rotamers were fine-tuned considering electrostatic and knowledge-based packing interactions as well as solvation effects. The hydrogen bonding network was optimized, including possible ligands. After that, an unrestrained high-resolution refinement with explicit solvent molecules was

performed, using knowledge-based force fields. The result was validated to ensure that the refinement did not move the model in a wrong direction. A hybrid model was built out of the top scoring models, where bad regions in the model were iteratively replaced with corresponding fragments from other models.

The final hybrid model of NcCAR-FL is mainly based on PDB structure 5JRH (Han et al., 2016) which was used as template for the assumed A-domain of NcCAR (433 residues aligned, 22% sequence identity). PDB entry 4W4T (Barajas et al., 2015) was used for modeling the assumed R-domain (321 residues aligned, 31% sequence identity) and entry 4DOW (Kuo et al., 2012) was used to model parts of the assumed A-T-domain (311 residues aligned, 26% sequence identity). Most of the found A-domain templates were acetyl-coenzyme A related enzymes which have approximately the same fold as A-domains of CARs. We also created a full-length model of NcCAR (NcCAR-FL) based on structures of CAR from *Segniliparus rugosus* (SruCAR, PDB codes 5MSW, and 5MSV) determined recently (Gahlloth et al., 2017). In order to obtain a full length template to build the NcCAR-FL model on, we first created a fusion model based on the crystal structures of SruCAR (UniProt: E5XP76) where the A-T-domain was in an “open” position and the

TABLE 2 | Highly conserved signature sequences of CARs and summary of results of alanine scan of NcCAR.

Entry	Domain	Consensus sequence	Suggested function
1	A	¹⁸³ I Y T S ¹⁸⁴ G S T G x P ¹⁸⁹ K x A ¹⁹⁰ L H S T T S I	Phosphate binding
2	A	²³⁴ M P M S ²³⁷ H x x G L L F S T N	Substrate binding
3	A	²⁸⁵ P X ₃ E X ₉₋₃₄ R X ₅ L G G ³¹⁰ D S F Q P N V A D	
4	A	³³⁶ Y G x T ³³⁷ E x G F S T	AMP binding
5	A	⁴⁰⁵ D X ₁₄₋₁₇ ⁴²² R X ₂ N x L K L S Q ⁴³² G ⁴³³ E D I V F A N E M T	AMP binding, substrate binding
6	A	⁴⁴¹ F X ₅ L E X ₆ P x V X ₅ Y G ⁴⁵⁷ K I A I F V E	
7	T	⁵⁴² Y X ₇ Y X ₃₁₋₃₄ L X ₁₃₋₁₆ G x D ⁵⁹⁵ S L S A F W N M Q V	PPTase binding site
8	R	⁶⁹¹ V L L T G X ₂ ⁶⁹⁴ G x L G X ₃₋₄ L X ₁₆₋₁₇ R ⁶⁹⁷ V V S I V I	NAD(P)H binding
9	R	⁷⁵⁵ L G L X ₇ L X ₆ I X ₆ V N X ₇₋₈ F ⁷⁸⁷ F V V V M D	
10	R	⁸⁴³ S T V X ₂₆₋₃₂ ⁸⁴⁴ G Y X ₂ S ⁸⁴⁸ K W X ₂ E G S I G L S A Y	NAD(P)H binding
11	R	⁸⁷⁰ R x D M I X ₅₋₇ ⁸⁸² G X ₂ N X ₂ D ⁸⁸⁵ G Q V M E	Substrate binding
12	R	⁹⁰⁴ G x A P X ₁₄₋₂₅ A X ₁₇₋₂₅ N P X ₂₆₋₃₃ ⁹⁷⁸ W X ₁₉₋₂₆ L L R L H A I F	

X no significant effect **X** no activity **X** improved activity on hexanoic acid
X no soluble expression **X** low activity **X** activity solely on piperonylic acid
X low soluble expression

Bold amino acids were exchanged in the mutagenesis study (positions of residues are numbered according to NcCAR).

T part in the structure of the R-domain was present. The A-T and the T-R parts were structurally aligned and connected. This “fusion” template was then energy minimized using the YASARA standard energy minimization protocol to remove remaining clashes formed by the overlay of the two crystal structures, resulting in a full length *SruCAR* model similar to the model obtained by Gahloth et al. (2017). Based on this template, the *NcCAR-FL* model was created using the YASARA homology model protocol restricting it to use the *SruCAR* full length template. We also created models of the A-domain of *NcCAR* using the “closed” structure of the A-T-domain (PDB code 5MSS) as a template. In this “closed” template, the T-domain is in close contact to the A-domain, whereas in the “open” template, the T-domain is flipped away from the A-domain. Again, sequence identities were low, about 21% (793 residues aligned) over the full length *NcCAR* and around 20% (505 residues aligned) using the “closed” template of the A-T-domain of *SruCAR*. As the two models of the full length *NcCAR* were built using different templates, especially the T part consists of different conformations and the models were not aligned directly. Therefore, to compare the two models we aligned the A-domain and the R-domain separately. The R-domains of the two *NcCAR-FL* models (313 residues aligned) have a C α -RMSD of 3.2 Å and the A-domains (365 residues aligned) a C α -RMSD of 4.6 Å.

Site-Directed Mutagenesis

Seventeen Mutants of the *NcCAR* gene were constructed as described previously (Stolterfoht et al., 2017). Another 17 *NcCAR* mutants were generated by site-directed mutagenesis according to Agilent’s QuikChange II protocol using *PfuUltra* High-Fidelity DNA Polymerase (Agilent Technologies). The pETDUET1-*EcPPTaseHTNcCAR* plasmid was used as the template. Primers were ordered at Integrated DNA Technologies, see Supplementary Information (Table S1). PCR was carried out under the following conditions: 50 μ L total volume, 200 μ M of each dNTP, 0.2 μ M of each primer, 5 ng of template plasmid DNA, and 2.5 U of *PfuUltra* High-Fidelity DNA polymerase. Cycling conditions were as follows: initial denaturation at 95°C for 2 min, followed by 16 cycles of denaturation at 95°C for 30 s, annealing at 61°C for 30 s, extension at 72°C for 11 min, and a final extension at 72°C for 15 min. Amplification was controlled by electrophoresis of 10 μ L of the PCR reaction on an agarose gel. After the digestion of parental non-mutated methylated plasmid DNA with *DpnI* at 37°C for 1 h, the samples were desalted and electro-competent *Escherichia coli* Top10F’ cells were transformed with the resulting nicked vector DNA containing the desired mutations.

Plasmid DNA from transformants was checked by restriction site analysis using *XbaI*, *HindIII*, *EcoRV*, *XhoI*, and the correct mutation was confirmed by automated sequencing (Microsynth).

Protein Expression and Purification

E. coli K12 MG1655 RARE [for reduced aldehyde reduction, a strain in which seven genes encoding enzymes with confirmed activity on benzaldehyde were deleted (Kunjapur et al., 2014)]

was transformed with the pETDUET1-*EcPPTaseHTNcCAR* wild-type and mutant vectors, and colonies were selected on LB/ampicillin (Amp). For protein expression, the autoinduction protocol as described by Studier was used (Studier, 2005). Cultivations were performed in HT Multitron shakers (Infors AG). After 4 h at 37°C and 24 h at 20°C, the cells were harvested with an Avanti J-20 XP centrifuge (Beckman Coulter), and stored at –20°C until further use. Thawed cells were resuspended in binding buffer (20 mM KP_i pH 7.4, 0.5 M NaCl, 10 mM imidazole), disrupted by sonication using a 102C converter with a Sonifier 250 (Branson), and the cell-free extract (CFE) was obtained by centrifugation at 20,000 \times g, 4°C, for 1 h. The pellet was used to visualize CAR in insoluble fractions (IFs). Therefore, the pellet was resuspended in 6 M urea and further thoroughly dissolved by heating to 60°C for 45 min with vortexing in between. 1 mL was transferred to a microcentrifuge tube and centrifuged for 30 min at 16,000 \times g and 4°C. The CFE was used for protein purification 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 was exchanged for 50 mM MES buffer, pH 7.5, containing 10 mM MgCl₂, 1 mM EDTA, and 1 mM DTT by size exclusion chromatography using PD-10 columns (GE-Healthcare) with the gravity flow protocol. Protein concentrations of the CFEs, IFs, and Ni-affinity purified proteins were determined with the Pierce™ BCA Protein Assay Kit (Thermo Scientific) using BSA as standard. Aliquots of the purified protein solutions were shock frozen in liquid nitrogen and stored at –80°C. For SDS-PAGE of the purified protein samples (Figure S1, Figures 1A, 2A), 10 μ g of protein were denatured with NuPAGE™ LDS Sample Buffer for 10 min at 80°C and SDS-PAGE was performed using Novex® 4–12% Bis-Tris-gels (50 min at 200 V and 120 mA in MOPS-buffer), followed by staining using SimplyBlue™ SafeStain (Invitrogen). The same procedure was performed for SDS-PAGE of the CFEs (Figure S2). For SDS-PAGE of the IFs (Figure S3), NuPAGE™ Sample Reducing Agent (Thermo Scientific), and NuPAGE™ LDS Sample Buffer were added to the protein sample for denaturation.

Spectrophotometric Assay

The activity of *NcCAR* variants compared to the wild-type toward three different substrates was determined monitoring NADPH depletion. The assay composition was as follows: 10 μ L of NADPH (10 mM in water), 10 μ L of ATP (20 mM in water), 10 μ L of CAR enzyme preparation from Ni-affinity chromatography, 160 μ L of MES buffer (50 mM, pH 6.0, containing 10 mM MgCl₂), and 10 μ L of 100 mM substrate (cinnamic acid, piperonylic acid, or hexanoic acid) in 0.1 M KOH. Enzymes with little or no activity were used without further dilution (0.6–3.3 mg/mL). Active enzymes were used in appropriate dilutions (\geq 0.15 mg/mL). The depletion of NADPH was followed on a Synergy Mx Platerreader (BioTek) at 340 nm and 28°C for 10 min. Blank reactions without ATP were carried out in parallel. For each protein sample, at least three separate experiments were carried out in four technical replicates, respectively.

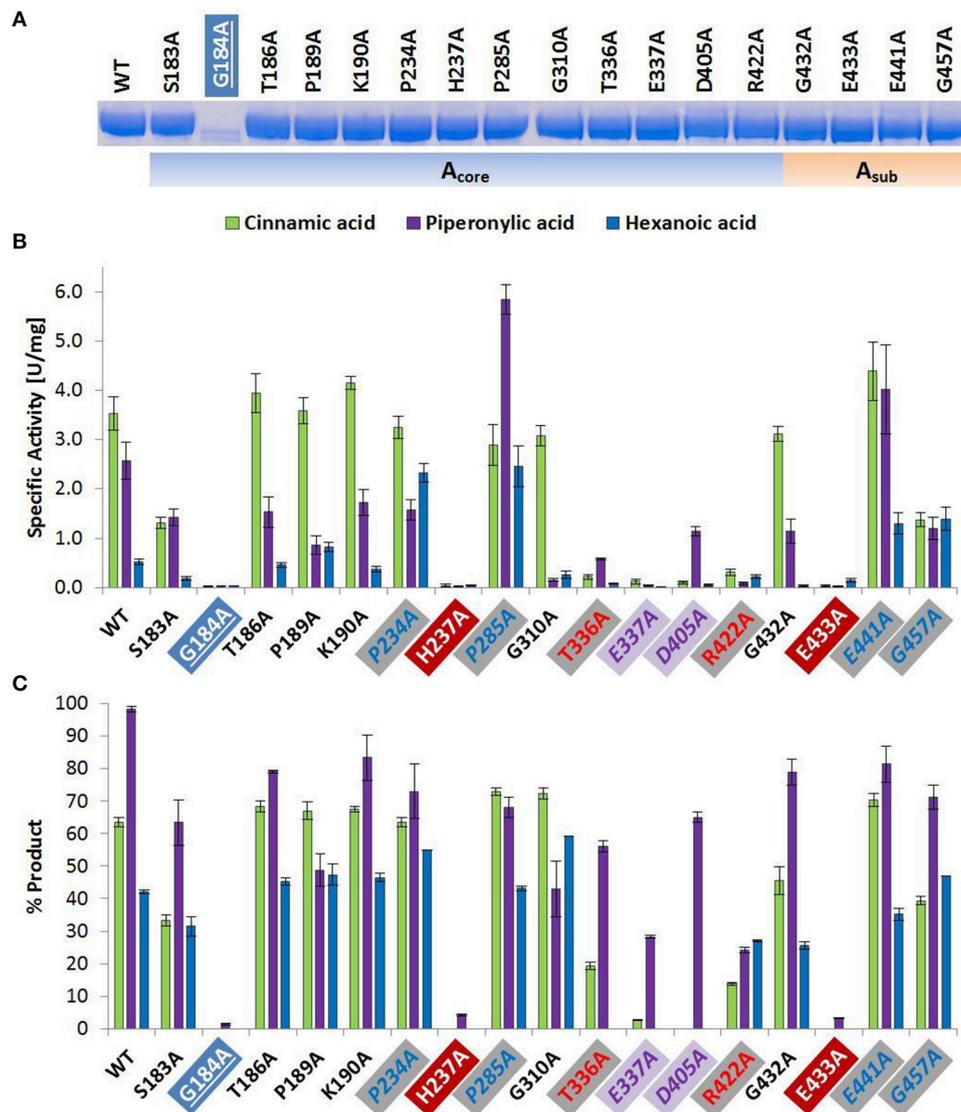


FIGURE 1 | Wild type and alanine mutants of *NcCAR* A_{core} - and A_{sub} -domain (A) SDS-PAGE of Ni-affinity chromatography purified wild-type enzyme and 17 *NcCAR* variants (protein amount loaded: 10 μg). For the full gel see Supporting Information Figure S1 (B) NADPH consumption was followed at 340 nm and 28°C for 10 min. Data are mean values of three separate experiments, each carried out in four technical replicates. (C) Result of whole cell biotransformation after 3 h at 28°C, shown as mean value of triplicates. Cinnamic acid and piperonylic acid reduction was analyzed by HPLC (UV = 254 nm). Green bars represent the sum of cinnamaldehyde and cinnamic alcohol, and violet bars piperonal, respectively. Hexanoic acid reduction was analyzed by GC-FID. Blue bars represent the sum of hexanal and hexanol. For the legend of highlights see **Table 2** footnotes.

Whole Cell Biotransformations

The pETDuet1-*EcPPTaseHTNcCAR* vectors carrying the wild-type and the 34 mutant genes were transformed into electrocompetent *E. coli* K-12 MG1655 RARE cells. Overnight cultures in LB-0.8G medium (LB medium containing 0.8% glucose, 25 mM $(\text{NH}_4)_2\text{SO}_4$, 50 mM KH_2PO_4 , 50 mM Na_2HPO_4 , 1 mM MgSO_4 , and 100 $\mu\text{g}/\text{mL}$ Ampicillin) were inoculated with a single colony, and 2 μL were used to inoculate 100 μL of LB-5052 medium (LB-0.8G medium but with 0.5% glycerol, 0.2% lactose, and 0.05% glucose instead of 0.8% glucose) in 96-well microtiter plates (MTPs). After glucose depletion (6 h

of growth at 37°C, 600 rpm, in an Eppendorf Thermomixer), the temperature was lowered to 20°C for protein expression. After 24 h, 100 μL of reaction mixture were added directly to the cell suspensions in the MTP, and sealed with an adhesive seal. The reactions were carried out in triplicates (three wells for each variant/wild-type) in a total volume of 200 μL containing 50 mM potassium phosphate buffer, pH 7.4, 0.12 M glucose, 21 mM citrate, 4 mM MgSO_4 , and 4 mM cinnamic/piperonylic acid, or 3 mM hexanoic acid, at 28°C, 600 rpm. The cinnamic and piperonylic acid conversions were stopped after 3 h by addition of methanol 1:1 v/v and used for HPLC-UV analysis

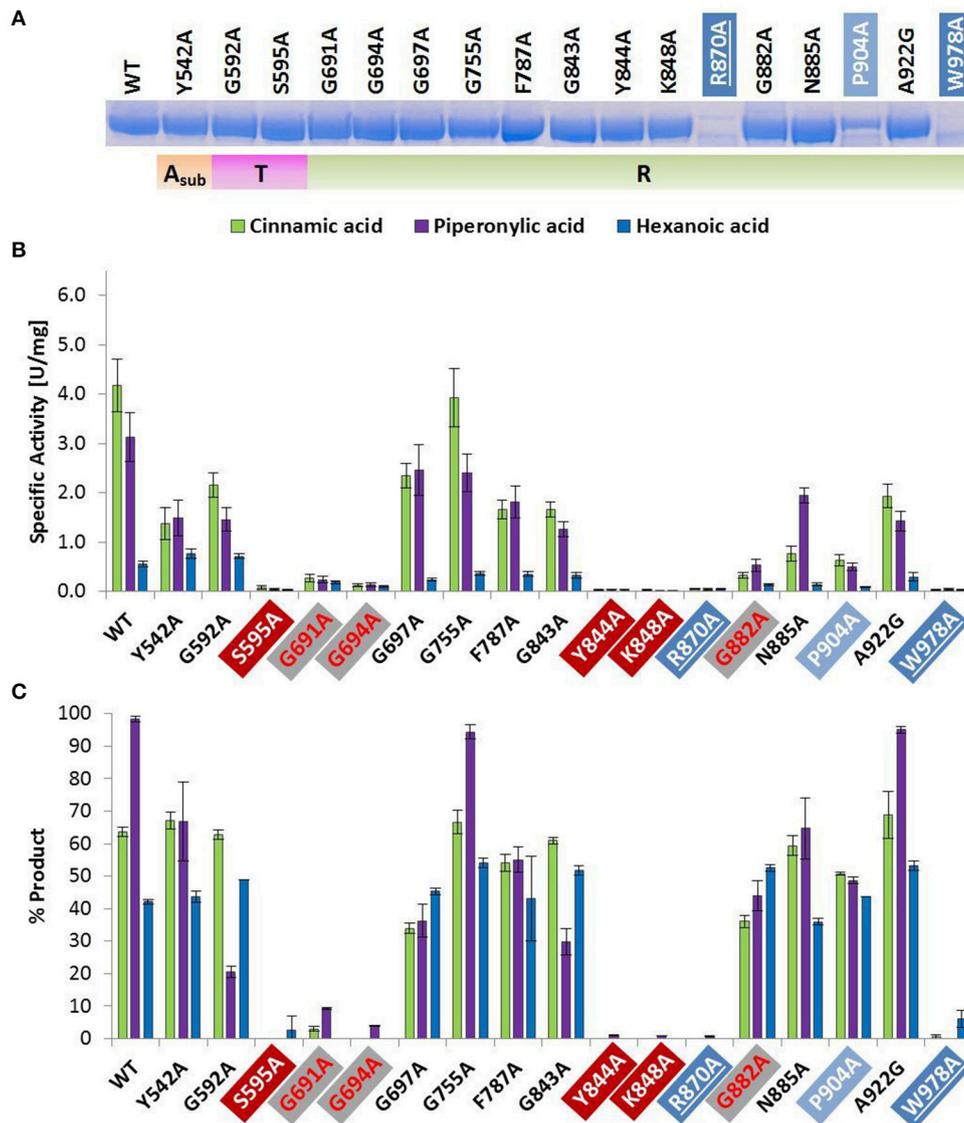


FIGURE 2 | Wild type and alanine mutants of *NcCAR* A_{sub}-, T-, and R-domain (A) SDS-PAGE of Ni-affinity chromatography purified wild-type enzyme and 17 *NcCAR* variants (protein amount loaded: 10 μg). For the full gel see Supporting Information Figure S1 (B) NADPH consumption was followed at 340 nm and 28°C for 10 min. Data are mean values of three separate experiments, each carried out in four technical replicates. (C) Result of whole cell biotransformation after 3 h at 28°C, shown as mean value of triplicates. Cinnamic acid and piperonylic acid reduction was analyzed by HPLC (UV = 254 nm). Green bars represent the sum of cinnamaldehyde and cinnamic alcohol, and violet bars piperonal, respectively. Hexanoic acid reduction was analyzed by GC-FID. Blue bars represent the sum of hexanal and hexanol. For the legend of highlights see **Table 2** footnotes.

after removal of biomass by centrifugation. Reactions with hexanoic acid as the substrate were frozen at -80°C before further sample preparation for GC-FID analysis. After thawing, 5 μL of 3 M HCl were added to each well to protonate hexanoic acid, as well as 95 μL of ddH₂O. 250 μL of each well were transferred to a microcentrifuge tube and extracted into 250 μL of ethyl acetate supplemented with 0.01% tetradecane as internal standard, vortexed, and centrifuged for 2 min at 16,000 × g. The organic phase was dried with sodium sulfate (5–10% v/v), and the supernatant was analyzed by GC-FID. The final substrate/product concentration was 2 mM in all samples.

Chromatographic Analyses

For substrate and product quantification by HPLC, reference acids, aldehydes, and alcohols were measured in different concentrations (0.050–2.5 mM), respectively, to calculate standard curves for each compound. Concentrations were calculated using the linear interpolation. The HPLC-UV measurements were performed with a Kinetex 2.6 μ Biphenyl 100 Å column (Phenomenex) equipped with a Phenylhexyl Security Guard ULTRA cartridge (Phenomenex). The mobile phases were 5 mM ammonium acetate with 0.5% v/v acetic acid in water, and acetonitrile (ACN) at a flow rate of 0.26 mL/min.

A stepwise gradient was used: 25–55% ACN (5 min), 55–70% ACN (5–7.2 min), 70–90% ACN (7.2–7.5 min), 90% ACN was held (7.5–9 min), and then re-equilibrated to starting conditions (25% ACN). The compounds were detected at 254 nm (DAD).

For substrate and product quantification by GC, hexanoic acid, hexanal, and hexanol were measured in different concentrations (0.31–10 mM), respectively, to calculate standard curves for each compound. Concentrations were calculated using the linear interpolation. The samples were analyzed on an Agilent 6890N Network GC System equipped with a flame-ionization detector (FID), a Combi Pal injector (CTC Analytics, Zwingen, Switzerland) and a Chirasil-DEX CB column (25 m × 0.32 mm, 0.25 μm film, Agilent, Vienna, Austria). The following settings were used for the analyses: injector temperature: 200°C, detector temperature: 230°C, 2:1 split ratio, 2.0 mL/min constant N₂ carrier gas flow; temperature programme: 40°C hold 5 min, 10°C/min to 140°C, 20°C/min to 160°C hold 7 min.

RESULTS

Multiple sequence alignments revealed a relatively small number of conserved residues (≤ 37 , depending on the alignment program used) among the currently known four CAR subtypes, given the total number of CAR amino acids (~1,050–1,200) (Stolterfoht et al., 2017). The goal of this study was to identify essential residues for enzymatic carboxylate reduction within these conserved residues to broaden our understanding of carboxylate reductases (E.C. 1.2.1.30) in general, and type III CARs in particular. Our hypothesis was that the replacement of catalytically essential amino acids would result in inactive protein. Based on multiple sequence alignments of 26 recombinant CARs functionally characterized to date (Table 1), and homology models of NcCAR generated to reveal putative active sites, NcCAR variants with replacements of 34 highly conserved amino acids were generated. Specifically, 33 residues were replaced by alanine and one conserved alanine residue was replaced by a glycine. Subsequently, the *in vivo* and *in vitro* activity of these variants was explored, and residues playing a role in one of the catalytic substeps, or in substrate or cofactor binding identified. For the essential post-translational modification, PPTase from *E. coli* was co-expressed with NcCAR from one plasmid. *E. coli* K12 MG1655 RARE served as host and the cells were cultivated as described previously (Schwendenwein et al., 2016).

CAR Expression and Activity

Wild-type NcCAR and each enzyme variant were purified by nickel-affinity chromatography and characterized by SDS-PAGE. Specific activities toward the reduction of cinnamic, piperonylic, and hexanoic acid were determined via a spectrophotometric NADPH depletion assay. Additionally, whole cell biotransformations were analyzed by HPLC-UV and GC-FID to detect the products and to assess the *in vivo* activity of NcCAR variants compared to the wild-type CAR. The impact of replacements of conserved residues is summarized in Table 2.

Amino Acid Replacements That Affect Protein Expression

The amino acid exchange G184A (Table 2, entry 1) in the A-domain and R870A (Table 2, entry 11) as well as W978A (Table 2, entry 12) in the R-domain was detrimental to soluble expression. The variants were found exclusively in the insoluble fraction (Supporting information, Figure S3). Similarly, the variant P904A (Table 2, entry 12) exhibited reduced soluble expression (Figures 1A, 2A).

Amino Acid Replacements That Abolish Activity

In the A-domain, the substitution of three amino acid residues (H237A, E337A, and E433A) abolished CAR activity as determined *in vitro* (Figure 1B). Notably, the variant E337A (Table 2, entry 4) retained some activity on piperonylic acid *in vivo* (Figure 1C). In the T-domain, no activity was found for enzyme variant S595A (Table 2, entry 7). Replacement of R-domain residues Y844 and K848 to alanine also led to activity loss (Table 2, entry 10, Figures 2B,C).

Amino Acid Replacements That Reduce Activity

A moderate decrease in activity was observed for the variants S183A and G457A, while the variants T336A, D405A, and R422A showed a stronger decrease in activity. Interestingly, the variant D405A (Table 2, entry 5) showed activity solely on piperonylic acid (*in vivo* and *in vitro*, Figures 1B,C). Similarly, the variant T336A (Table 2, entry 4) retained activity mainly on piperonylic acid (Figures 1B,C). The exchange of G310 for alanine (Table 2, entry 3) exhibited a specific activity with cinnamic acid similar to that of the wild-type, but a decreased activity with piperonylic and hexanoic acid. Only residual activity was observed for variants G691A and G694A (Table 2, entry 8). Furthermore, the activities of G882A and P904A were relatively low (Figures 2B,C).

Amino Acid Replacements That Increase Activity

The substitution of P285 with alanine seems to increase the activity for the reduction of piperonylic acid (Figure 1B). Furthermore, some substitutions with alanine (P189A, P234A, P285A, E441A, and G457A) enhanced the specific activity toward hexanoic acid compared to the wild-type. Therefore, we screened a spectrum of aliphatic acids from propanoic to nonanoic acid (Figure 3). Whereas wild-type NcCAR showed the highest activity toward butanoic acid, for some variants the preference was shifted to pentanoic acid with an overall increased activity on aliphatic acids with longer chains, especially for variants P234A and P285A.

Modeling of NcCAR

An overview of the models created in this study is depicted in Figure 4. As the underlying sequence identities to the available templates are low, the resulting model needs to be analyzed with care. Figure 5 shows close-ups of the A-domain in a closed and open conformation and the R-domain with the key residues highlighted.

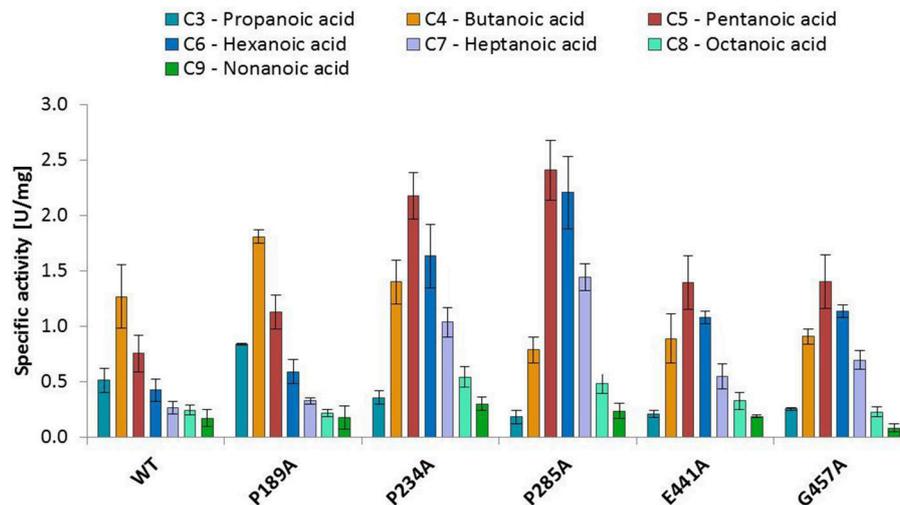


FIGURE 3 | Activities of wild-type *NcCAR* and selected variants for the reduction of aliphatic acids determined by following NADPH consumption at 340 nm for 10 min at 28°C. Data are shown as mean values of three separate experiments, carried out in four technical replicates, respectively. For the wild-type and the variants P234A, and P285A, data from biological duplicates were included.

DISCUSSION

Aligning all experimentally confirmed CAR sequences with A-T-R architecture and broad substrate scope, a particular amino acid pattern becomes apparent that seems to be inherent to the carboxylate reductase function in all E.C. 1.20.1.30 CARs. To gain some insight into the overall structure arrangement and possible locations of the binding pockets, we created *NcCAR* models (Figures 4, 5). We used this information to predict cofactor and substrate binding sites (Table 2, suggested function). We hypothesized that some of the conserved residues would be important key residues for catalysis and therefore performed an alanine scan, assuming that replacement of a key residue by alanine would abolish enzymatic carboxylate reduction activity.

Key Residues in the A-Domain

We assume His 237 of the PxxH motif (Table 2, entry 2), and Glu 433 of the L(F)S(A)xGEK(F) motif (Table 2, entry 5) are essential residues of the A-domain, because their replacement results in complete activity loss. According to the *NcCAR* model, His 237 may interact with a potential substrate and the phosphate of AMP (Figures 5A,B). His 237 (numbering according to *NcCAR*, His 300 in *NiCAR*) was previously recognized to be conserved in CAR enzymes and to be located close to both AMP phosphate and the substrate carboxylate (Gahloth et al., 2017). To the best of our knowledge, the essential Glu 433 was studied experimentally herein for the first time. Glu 433 was interesting as it is positioned at the hinge region (two short beta strands, transparent red circles, Figures 5A,B), which is pointing in opposite directions in the closed and open models. In the closed model, it is near to His 237 and in the open model it is distant from the binding pockets. As explained

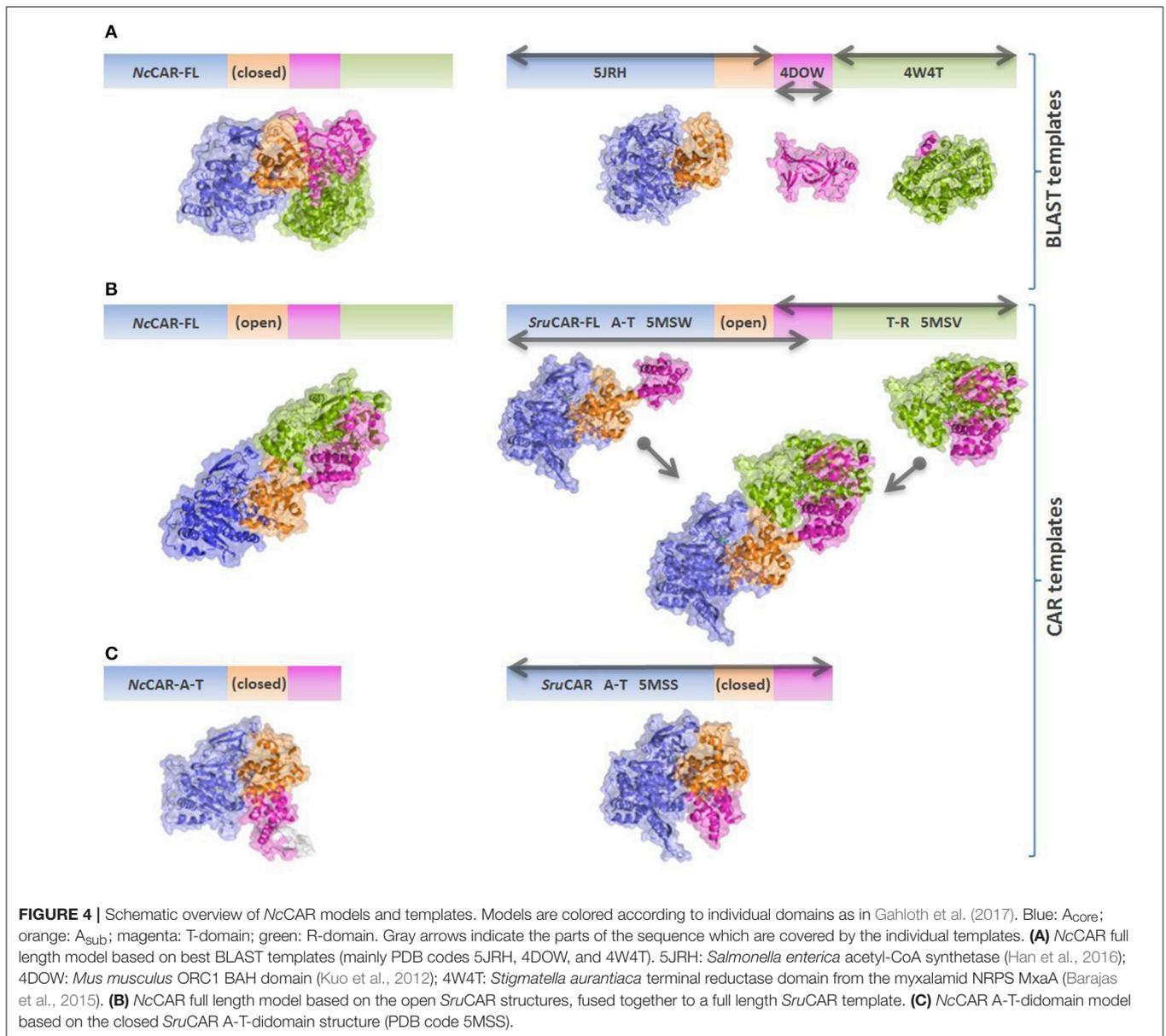
by Gahloth et al. the open form created by rotation of the A-T-domain, may not be as pronounced in solution and is maybe stabilized by crystal contacts upon the crystallization process. The distance of Glu (E533 in *SruCAR*) to the essential His (H315 in *SruCAR*) is also indicated in Figures 2c,d in Gahloth et al. (2017). Our model confirms this observation (Figures 5A,B).

Role of Other Conserved Residues of the A-Domain

Like His 237, Ser 183 and Thr/Ser 182 (Table 2, entry 1) and Thr 336 (entry 4) may interact with the AMP cofactor according to our *NcCAR* model. The S(T)SGS(T)TG(S)xPK motif (Table 2, entry 1) resembles the highly conserved core motif A3 of NRPS defined by Marahiel et al. (1997), to which an interaction with the pyrophosphate leaving group was assigned (Conti et al., 1997). However, the substitution of the Ser, Thr, Pro, or Lys of this motif in CARs did not significantly impact on CAR activity. Other core motifs of NRPS defined by Marahiel et al. share also similarity with signature sequences of CARs defined in Table 2. For example, the core motif A5 is akin to the G(S)TE_xG(T) motif (Table 2, entry 4). Thr 336 and Glu 337 of this motif (Table 2, entry 4) as well as Asp 405 and Arg 422 of the D_x14–17R motif (entry 5) are not considered essential, because their exchange for alanine resulted in active variants (Figures 1B,C). Asp 405 and Arg 422 may be involved in interactions with the adenosine ribose according to our model (Figures 5A,B).

Key Residue of the T-Domain

As already reported by others, the replacement of the phosphopantetheine binding site Ser 689 in *NiCAR* by Ala resulted in an inactive variant (Venkitasubramanian et al., 2007; Wood et al., 2017). The same is true for the equivalent exchange



of Ser 595 in *NcCAR* [Table 2, entry 7; or core motif T (core 6) in Marahiel et al. (1997)], confirming that this residue serves as the phosphopantetheine attachment site and is indispensable for CAR activity.

Key Residue of the R-Domain

In the R-domain, Tyr 844, and Lys 848 of the GYxxxKxxxE(S) motif (Table 2, entry 10) are essential for catalysis, as their substitution with alanine led to complete inactivation of *NcCAR*. Wang and co-workers revealed that the R-domain of *AtCAR* shares the conserved patterns of TGxxGxxG and YxxxK with short-chain dehydrogenase/reductase (SDR) family proteins (Wang et al., 2014). Previous studies on SDRs assigned a structural role in the coenzyme binding region to the TGxxGxxG motif, and suggested YxxxK to be part of the active site in SDR

proteins (Persson et al., 2003). Several site-directed mutagenesis studies on SDR enzymes involved the conserved residues of these motifs and confirmed their importance as their substitution abolished enzyme activity (Chen et al., 1990; Albalat et al., 1992; Obeid and White, 1992; Cols et al., 1993; Chhabra et al., 2012; Barajas et al., 2015). For CARs, a first comparable study to analyse the contribution of these conserved residues to catalytic activity was carried out by Wang and co-workers. Therefore, T690 of the TGxxGxxG motif, as well as the active site Y863 and K867 of *AtCAR* were mutated. The activities of the T690A and Y863F variants on 5-methyl orsellinic acid (5-MOA) were drastically reduced, whereas the exchange of Lys 867 for Ala led to a completely inactive enzyme (Wang et al., 2014), which agrees with our study of the corresponding residues in *NcCAR* (Y844 and K848).

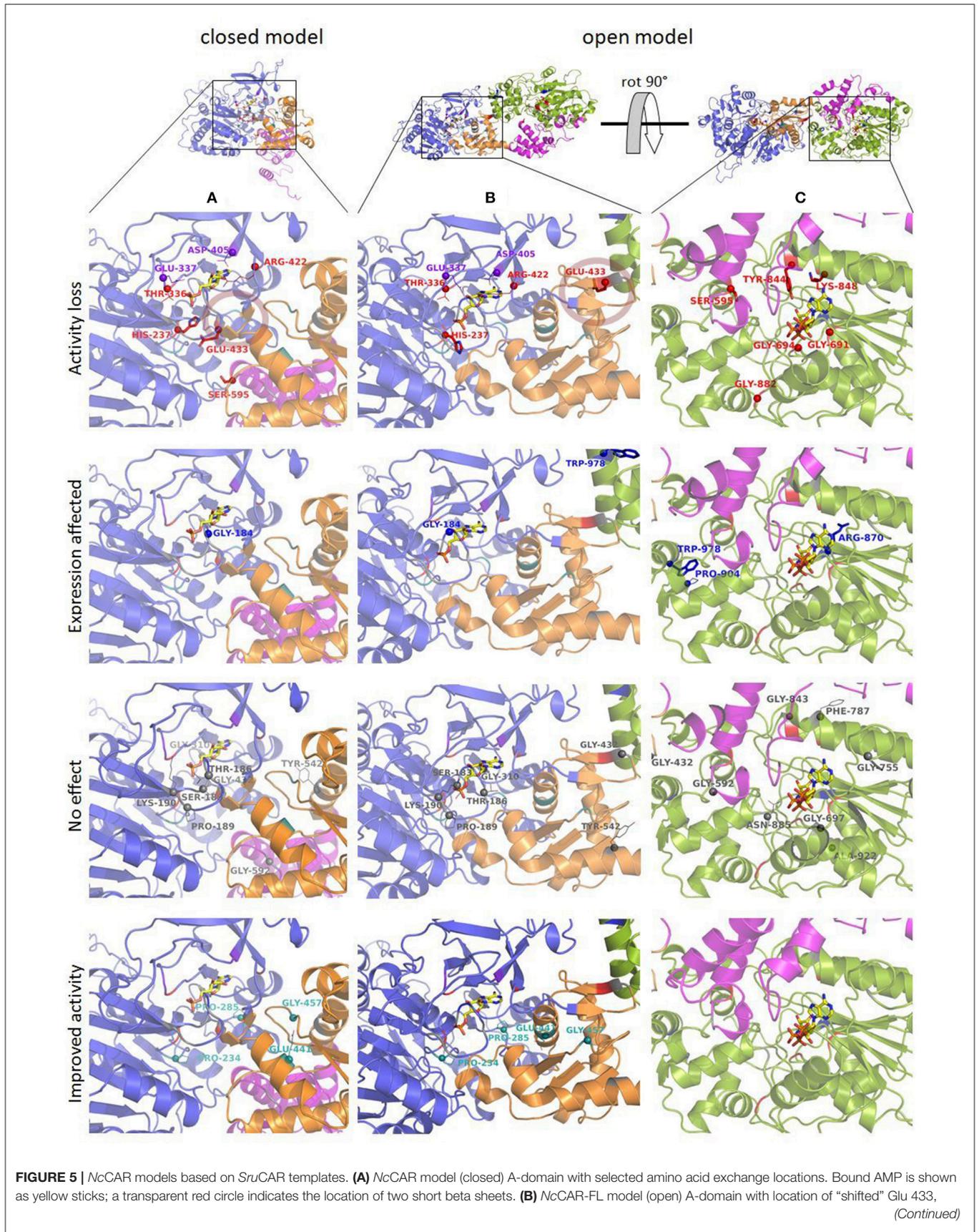


FIGURE 5 | located at the two short beta sheets. **(C)** Close-up of R-domain of *NcCAR*-FL model with selected amino acid exchange locations. Bound NAP is shown as yellow sticks. Red spheres with amino acids as stick representation indicate selected residues which caused activity loss upon replacement by alanine. Amino acids with red line representation indicate variants where little activity was observed and purple spheres indicate locations where replacements with alanine retained activity solely on piperonylic acid. Blue spheres with amino acids as stick representation indicate amino acid exchange locations in the model where expression was abolished and amino acids indicated as blue lines where the expression was very low. Gray spheres indicate locations where, after substitution by alanine, activity was not significantly changed compared to the wild-type *CAR*. Teal spheres indicate amino acid exchanges where improved activity for hexanoic acid was observed.

Role of Other Conserved Residues of the R-Domain

The first two glycines of the T(S)GxxGxxG motif (Table 2, entry 8) are important residues of the R-domain, because their replacement resulted in drastic reduction of activity (Figures 2B,C). This motif is well known as “Rossmann-fold” dinucleotide cofactor binding motif. In the *NcCAR* model, the R-domain (Figure 5C) clearly shows an “extended” type of NADPH-binding α/β Rossmann fold, which is composed of seven parallel beta strands in a twisted beta sheet arranged in the order 3-2-1-4-5-6-7. The signature sequence T/SGxxGxxG at the region of the tight turn at the end of the first β -strand was reported to be in contact with the negatively charged oxygens of the two phosphates of NADPH. The first glycine was proposed to be important for the tightness of the turn, the second allows the dinucleotide to be bound without obstruction from an amino acid side chain at this position, and the third seems to provide space for a close interaction between the β -strand and the succeeding α -helix. Other conserved features of this fingerprint region are (i) a hydrophilic residue at the N-terminus of the first β -strand, (ii) six small hydrophobic residues that form a hydrophobic core of the $\beta\alpha\beta$ unit, and (iii) in case of NADP(H) binding: a basic residue in the loop after the second β -strand, as observed in the extended type of SDRs (Wierenga et al., 1985; Scrutton et al., 1990; Lesk, 1995). Except for the first of the three above-mentioned conserved features, all these characteristics are in accordance with our *NcCAR* model. The basic residue mentioned above is probably a highly conserved Arg, located about 20 residues downstream of the third Gly (R718 in *NcCAR*; Table 2, entry 8).

The comparison of the presumable NADPH binding site of our model with available protein structures of members of the SDR family provided an indication for other residues involved in the reduction mechanism or in NADPH binding. For example Breton et al. reported that a conserved aspartic acid residue, found in the loop between β_3 of the seven-stranded β -sheet and α_3 , interacts with the adenine moiety of NADP⁺ (Breton et al., 1996). In *NcCAR* we also found an Asp in this loop. It seems to be conserved in *CAR*s, except for *TvCAR*, which has a valine at this position. Based on our *NcCAR* models, the side chain of Asp 748 could maybe interact with the adenine.

Consistent with previous structural studies on the SDR superfamily, the nicotinamide moiety of the cofactor is pointing toward the active site (Figure 5C). Besides the YxxxK motif (Table 2, entry 10), a Ser or Thr constitutes the catalytic triad (Jörnvall et al., 1995; Oppermann et al., 1997), which is most likely located at position 815 in *NcCAR*. At this position, there is a Ser exclusively in fungal *CAR*s, whereas in bacterial *CAR*s, Thr is strictly conserved (Table 2, entry 10). In most SDRs an

essential Asn residue (or another Ser), that could be connected through a water molecule to the active site Lys, is extending this triad to form a catalytic tetrad (Filling et al., 2002). Such a residue could not be found in our models or is absent in *CAR*s. According to the concept of the reductive reaction mechanism postulated for SDRs, the Tyr donates a proton to the oxygen of the substrate carbonyl, Ser stabilizes the deprotonated state of Tyr, and the pro-S hydride of the nicotinamide nucleophilically attacks the carbon of the substrate carbonyl. A proton relay is formed that involves the 2'OH of the nicotinamide ribose, which is stabilized by the Lys side chain, and a water molecule bound to the backbone carbonyl of the Asn (Breton et al., 1996; Filling et al., 2002). The presence of at least two, if not all, residues of the catalytic triad/tetrad in *CAR*s lets us assume a similar reduction mechanism. The major role of the Lys was proposed to be the stabilization of the dinucleotide through the interaction with the nicotinamide ribose. In the model discussed herein, the ϵ -amino group of Lys 848 is pointing away from the 2'-hydroxyl of the nicotinamide ribose.

Finally, replacement of Gly 882 of the GxxNxxD/E motif (Table 2, entry 1) by Ala led to significantly decreased activity (Figure 2B). Asn 885 may interact with the pyrophosphate of the dinucleotide and is on the open side of the cavity, therefore also accessible to potential substrate interactions. Another speculation on the basis of the *NcCAR* model is that the acidic residue of this motif interacts with the substrate.

Substrate Specificity Conferring Residues

Variant P234 reduced piperonylic acid less well compared to the wild-type, but showed increased activity for hexanoic acid. Variant P285A was outstanding for both hexanoic acid and piperonylic acid (Figure 1B). We propose that specific residues impart certain specificities, which would allow engineering of *CAR*s for an increased activity toward a desired substrate. To test this hypothesis, we determined specific activities of five alanine variants for the reduction of a range of aliphatic acids and indeed saw distinct substrate preference patterns. Replacement of the two prolines P234 and P285 with alanine elevated the activity for longer aliphatic acids (Figure 3). Since both prolines are possibly located at the margin of the binding pocket – Pro 234 is in proximity of the essential His 237 residue, and Pro 285 is located in a turn between a short β -sheet and an α -helix – without the rigidity or flexion of the peptide chain conferred by proline, both replacements may change the cavity in a way favorable for binding longer fatty acids. In addition, a specificity shift toward longer aliphatic acids was observed for the variants E441A and G457A (Figure 3).

When Gross et al. isolated the *NcCAR* enzyme from the fungus SY7A and characterized its substrate specificity, no activity with aliphatic acids, and in particular with hexanoic acid (caproic acid) was detected (Gross and Zenk, 1969; Gross, 1972). A possible explanation for the discrepancy between our and Gross result regarding substrate scope of *NcCAR* could be that we deal with different enzyme variants in *N. crassa* strain SY7A and OR74A, respectively, or the recombinantly expressed enzyme lacks so far undescribed post-translational modifications.

Our findings support the proposed “gate-keeper”/substrate recognition function of the A-domain (Wang and Zhao, 2014; Moura et al., 2016; Finnigan et al., 2017), as the mutational exchanges in the A-domain led to either increase or decrease in activity while no activity enhancements were observed when residues of the R-domain were replaced.

Amino Acid Substitutions That Affect CAR Expression

The replacement of Gly 184, Arg 870, and Trp 978 by alanine did not result in soluble expression. Pro 904 is located close to Trp 978, and its substitution by Ala yields significantly less soluble protein (Figure 2A). The *in vitro* activity of this variant was very low (Figure 2B), however, the *in vivo* activity of P904A was comparable to the wild-type activity (Figure 2C). This was also observed for variant G882A, and some other variants, especially for E337A, which showed virtually no activity *in vitro*, but clearly exhibited activity on piperonylic acid in the whole cell biotransformation (Figures 1B,C). A reason therefore could be protein stability.

Outlook on Probing Further Residues

We restricted this mutagenesis study to strictly conserved amino acids. Furthermore, only variants with single amino acid exchanges and substitutions only for alanine (or in case of alanine, for glycine) were generated. However, it is to be expected that prospective mutagenesis studies could broaden our understanding of the function of each residue present in signature sequences and could even reveal further essential residues or residues that impart specificities for certain substrates. Additional residues that are suitable for exchange studies are those not strictly conserved, but conserved regarding their biochemical property (e.g., acidic, aromatic, hydroxylic, aliphatic). For example, the threonine in the motif of three glycines (Table 2, entry 8), that is characteristic of NAD(P)(H)-binding enzymes, was not replaced, as it is not completely conserved, albeit only in *MsCAR2* and *MsCAR3* a serine with similar biochemical property occupy this position. The three aliphatic residues upstream of this threonine are likewise conserved. Considering amino acids possessing comparable biochemical property, many more positions than the 34 which were examined in this study are conserved, e.g., the hydroxylic residues neighboring on Ser 183 and Gly 184 (Table 2, entry 1), an acidic amino acid four residues downstream of the conserved Pro 285 (entry 3), or an aromatic amino acid upstream of the conserved Tyr 542 (Table 2, entry 7), to name just a few. Other positions are often only conserved among bacterial CARs. For example while bacterial CARs share an aromatic Phe after the

conserved Gly 432 and Glu 433 in entry 6, there is a basic lysine at the respective position in fungal CARs (*NcCAR*, *TvCAR*, *AtCAR*, and *SbCAR*). Another example is the universally conserved Lys from ANL motif A10 (Marahiel et al., 1997), involved in polar interactions with both the adenosine and the nucleotide ligand, highly conserved in bacterial CARs (Gahlth et al., 2017), but not in fungal CARs, where there is a hydroxylic residue in place of this Lys (Thr 524 in *NcCAR*, not listed in Table 2). As Wang and Zhao engineered *AtCAR* toward their target substrate anthranilate they also found a hydroxylic residue (Ser 540 in *AtCAR*) occupying the same position as the lysine previously reported to be absolutely conserved in other A-domains of NRPSs. Their homology model suggested that the distinct Ser, other than Lys at this position, is not involved in any interaction with the ligand (Wang and Zhao, 2014). However, there is a Lys two amino acids upstream in fungal CARs (Lys 522 in *NcCAR*) which could provide these key interactions as it points toward the adenosine ribose in the open model but is directed outward in the closed model. Another position to be mentioned is the conserved key Asp residue (D984 in *NiCAR*), of which two distinct conformations in the reductase active site were observed [the orientation of this residue was postulated to lead to an on/off state of the reductase (Gahlth et al., 2017)], which is occupied by a Gly in fungal CARs (G872 in *NcCAR*, Table 2, entry 11). Since all characterized CARs show relaxed substrate scope, and no significant preferences for a particular substrate class can be assigned to one phylogenetic subgroup of CARs so far, no specificity conferring code can be deduced from these positions.

CONCLUSION

To expand the knowledge about the structure and molecular functionality of CARs, the identification of key residues for the enzymatic reaction is of importance. The function of conserved residues was studied by a multiple sequence alignment and mutational replacements of these residues in a fungal CAR from *N. crassa*. The effect of these replacements was investigated by analyzing enzymatic activity *in vivo* and *in vitro*. The interpretation of the impact of the amino acid exchanges was supported by homology modeling and comparisons to related structures.

Residues H237, E433, S595, Y844, and K848 were identified to be essential for carboxylate reductase function. Residues T336, E337, D405, R422, G691, G694, and G882 seem to be also important for activity since the exchange to Ala significantly reduced CAR activity. Replacement of P234, P285, E441, and G457 by alanine increased the activity for hexanoic acid reduction, therefore, we assume that these residues play a key role in substrate binding. Our attempts to model *NcCAR* gave a model of rather limited quality at this point, because identities to structural templates were generally very low. Nevertheless, the model allowed us to speculate on the function of several conserved residues.

N. crassa CAR readily reduced all tested aliphatic acids from C3 to C9. The wild type protein showed its highest activity for butanoic acid, but substitution of P234 and P285 resulted in

a shift of the substrate preference toward pentanoic acid, and a higher specific activity for aliphatic acids compared to the wild-type NcCAR in general. This study is a first comprehensive investigation of conserved residues by means of mutagenesis, and the herewith identified signature sequences of E.C. 1.2.1.30 CARs may help in the identification of new proteins with CAR activity.

AUTHOR CONTRIBUTIONS

MW led the project, conceived the research and designed wet lab experiments. GS and KG designed the modeling. HS prepared and purified mutants, carried out in vitro and in vivo assays, analyzed the results and designed figures. DS prepared mutants and developed the protocol for growth in MTPs. GS and TP-K conducted the modeling. HS, GS, and MW co-wrote the manuscript. All authors commented on the manuscript.

REFERENCES

- Akhtar, K. M., Turner, N. J., and 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. doi: 10.1073/pnas.1216516110
- Albalat, R., González-Duarte, R., and Atrian, S. (1992). Protein engineering of *Drosophila* alcohol dehydrogenase. The hydroxyl group of Tyr152 is involved in the active site of the enzyme. *FEBS Lett.* 308, 235–239. doi: 10.1016/0014-5793(92)81282-Q
- Barajas, J. F., Phelan, R. M., Schaub, A. J., Kliewer, J. T., Kelly, P. J., Jackson, D. R., et al. (2015). Comprehensive structural and biochemical analysis of the terminal myxalamid reductase domain for the engineered production of primary alcohols. *Chem. Biol.* 22, 1018–1029. doi: 10.1016/j.chembiol.2015.06.022
- Behrouzian, B., McDaniel, R., Zhang, X., and Clark, L. (2010). *Engineered Biosynthesis of Fatty Alcohols*. WO/2010/135624. Redwood City.
- Breton, R., Housset, D., Mazza, C., and Fontecilla-Camps, J. C. (1996). The structure of a complex of human 17 β -hydroxysteroid dehydrogenase with estradiol and NADP⁺ identifies two principal targets for the design of inhibitors. *Structure* 4, 905–915. doi: 10.1016/S0969-2126(96)00098-6
- Bushley, K. E., Ripoll, D. R., and Turgeon, B. G. (2008). Module evolution and substrate specificity of fungal nonribosomal peptide synthetases involved in siderophore biosynthesis. *BMC Evol. Biol.* 8:328. doi: 10.1186/1471-2148-8-328
- Chen, Z., Lu, L., Shirley, M., Lee, W. R., and Chang, S. H. (1990). Site-Directed mutagenesis of glycine-14 and two “critical” cysteinyl residues in *Drosophila* alcohol dehydrogenase. *Biochemistry* 29, 1112–1118.
- Chhabra, A., Haque, A. S., Pal, R. K., Goyal, A., Rai, R., Joshi, S., et al. (2012). Nonprocessive [2 + 2]-off-loading reductase domains from mycobacterial nonribosomal peptide synthetases. *Proc. Natl. Acad. Sci. U.S.A.* 109, 5681–5686. doi: 10.1073/pnas.1118680109
- Cols, N., Marfany, G., Atrian, S., and González-Duarte, R. (1993). Effect of site-directed mutagenesis on conserved positions of *Drosophila* alcohol dehydrogenase. *FEBS Lett.* 319, 90–94. doi: 10.1016/0014-5793(93)80043-T
- Conti, E., Stachelhaus, T., Marahiel, M. A., and Brick, P. (1997). Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *Embo J.* 16, 4174–4183. doi: 10.1093/emboj/16.14.4174
- Duan, Y., Yao, P., Chen, X., Liu, X., Zhang, R., Feng, J., et al. (2015a). Exploring the synthetic applicability of a new carboxylic acid reductase from *Segniliparus rotundus* DSM 44985. *J. Mol. Catal., B Enzym.* 115, 1–7. doi: 10.1016/j.molcatb.2015.01.014
- Duan, Y., Yao, P., Du, Y., Feng, J., Wu, Q., and 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. doi: 10.3762/bjoc.11.243

ACKNOWLEDGMENTS

The skilful technical assistance of K. Rudnicka and D. Gradischinig is gratefully acknowledged. The Austrian science fund FWF is kindly acknowledged for financial support (Elise-Richter fellowship V415-B21 and P28477-B21). We also thank K. L. J. Prather for generously providing *E. coli* K12 MG1655 RARE. This research has also been supported by the Austrian BMWWF, BMVIT, SFG, Standortagentur Tirol, Government of Lower Austria, and ZIT through the Austrian FFG-COMET- Funding Program.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2018.00250/full#supplementary-material>

- Filling, C., Berndt, K. D., Benach, J., Knapp, S., Prozorovski, T., Nordling, E., et al. (2002). Critical residues for structure and catalysis in short-chain dehydrogenases/reductases. *J. Biol. Chem.* 277, 25677–25684. doi: 10.1074/jbc.M202160200
- Finnigan, W., Thomas, A., Cromar, H., Gough, B., Snajdrova, R., Adams, J. P., et al. (2017). Characterization of carboxylic acid reductases as enzymes in the toolbox for synthetic chemistry. *ChemCatChem* 9, 1005–1017. doi: 10.1002/cctc.201601249
- Gahlth, D., Dunstan, M. S., Quaglia, D., Klumbys, E., Lockhart-Cairns, M. P., Hill, A. M., et al. (2017). Structures of carboxylic acid reductase reveal domain dynamics underlying catalysis. *Nat. Chem. Biol.* 13, 975–981. doi: 10.1038/nchembio.2434
- Galagan, J. E., Calvo, S. E., Borkovich, K. A., Selker, E. U., Read, N. D., Jaffe, D., et al. (2003). The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature* 422, 859–868. doi: 10.1038/nature01554
- Gross, G. G. (1972). Formation and reduction of intermediate acyladenylate by aryl-aldehyde: NADP oxidoreductase from *Neurospora crassa*. *Eur. J. Biochem.* 31, 585–592. doi: 10.1111/j.1432-1033.1972.tb02569.x
- Gross, G. G., Bolkart, K. H., and Zenk, M. H. (1968). Reduction of cinnamic acid to cinnamaldehyde and alcohol. *Biochem. Biophys. Res. Commun.* 32, 173–178. doi: 10.1016/0006-291X(68)90365-3
- Gross, G. G., and Zenk, M. H. (1969). Reduktion aromatischer Säuren zu Aldehyden und Alkoholen im zellfreien System. *Eur. J. Biochem.* 8, 413–419. doi: 10.1111/j.1432-1033.1969.tb00543.x
- Gulick, A. M. (2009). Conformational dynamics in the Acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase. *ACS Chem. Biol.* 4, 811–827. doi: 10.1021/cb900156h
- Hagedorn, S., and Kaphammer, B. (1994). Microbial biocatalysis in the generation of flavor and fragrance chemicals. *Annu. Rev. Microbiol.* 48, 773–800. doi: 10.1146/annurev.mi.48.100194.004013
- Han, X., Shen, L., Wang, Q., Cen, X., Wang, J., Wu, M., et al. (2016). Cyclic AMP inhibits the activity and promotes the acetylation of acetyl-CoA synthetase through competitive binding to the ATP/AMP pocket. *J. Biol. Chem.* 292, 1374–1384. doi: 10.1074/jbc.M116.753640
- He, A., Li, T., Daniels, L., Fotheringham, I., and 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. doi: 10.1128/AEM.70.3.1874-1881.2004
- Hu, Z. (2010). *Methods and Compositions for Producing Fatty Alcohols*. US20100105963 A1. South San Francisco.
- Huber, C., Caldeira, J., Jongejan, J. A., and Simon, H. (1994). Further characterization of two different, reversible aldehyde oxidoreductases from *Clostridium formicoaceticum*, one containing tungsten and the other molybdenum. *Arch. Microbiol.* 162, 303–309. doi: 10.1007/BF00263776

- Jörnvall, H., Persson, B., Krook, M., Atrian, S., González-Duarte, R., Jeffery, J., et al. (1995). Short-Chain Dehydrogenases/Reductases (SDR). *Biochemistry* 34, 6003–6013. doi: 10.1021/bi00018a001
- Kavanagh, K. L., Jörnvall, H., Persson, B., and Oppermann, U. (2008). The SDR superfamily: functional and structural diversity within a family of metabolic and regulatory enzymes. *Cell. Mol. Life Sci.* 65, 3895–3906. doi: 10.1007/s00018-008-8588-y
- Khusnutdinova, A. N., Flick, R., Popovic, A., Brown, G., Tchigvintsev, A., Nocek, B., et al. (2017). Exploring bacterial carboxylate reductases for the reduction of bifunctional carboxylic acids. *Biotechnol. J.* 12:1600751. doi: 10.1002/biot.201600751
- Krieger, E., Koraimann, G., and Vriend, G. (2002). Increasing the precision of comparative models with YASARA NOVA - A self-parameterizing force field. *Proteins Struct. Funct. Genet.* 47, 393–402. doi: 10.1002/prot.10104
- Kunjapur, A. M., Tarasova, Y., and 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. doi: 10.1021/ja506664a
- Kuo, A. J., Song, J., Cheung, P., Ishibe-Murakami, S., Yamazoe, S., Chen, J. K., et al. (2012). ORC1 BAH domain links H4K20 to DNA replication licensing and Meier-Gorlin syndrome. *Nature* 484, 115–119. doi: 10.1038/nature10956
- Lesage-Meessen, L., Delattre, M., Haon, M., Thibault, J.-F., Ceccaldi, B. C., Brunerie, P., et al. (1996). A two-step bioconversion process for vanillin production from ferulic acid combining *Aspergillus niger* and *Pycnoporus cinnabarinus*. *J. Biotechnol.* 50, 107–113. doi: 10.1016/0168-1656(96)01552-0
- Lesk, A. M. (1995). NAD-binding domains of dehydrogenases. *Curr. Opin. Struct. Biol.* 5, 775–783. doi: 10.1016/0959-440X(95)80010-7
- Li, C., Matsuda, Y., Gao, H., Hu, D., Yao, X. S., and Abe, I. (2016). Biosynthesis of LL-Z1272β: discovery of a new member of NRPS-like enzymes for aryl-aldehyde formation. *ChemBioChem* 17, 904–907. doi: 10.1002/cbic.201600087
- Li, T., and Rosazza, J. P. N. (1997). Purification, characterization, and properties of an aryl aldehyde oxidoreductase from *Nocardia* sp. strain NRRL 5646. *J. Bacteriol.* 179, 3482–3487. doi: 10.1128/jb.179.11.3482-3487.1997
- Marahiel, M. A., Stachelhaus, T., and Mootz, H. D. (1997). Modular peptide synthetases involved in nonribosomal peptide synthesis. *Chem. Rev.* 97, 2651–2674. doi: 10.1021/cr960029e
- Moura, M., Pertusi, D., Lenzi, S., Bhan, N., Broadbelt, L. J., and Tyo, K. E. (2016). Characterizing and predicting carboxylic acid reductase activity for diversifying bioaldehyde production. *Biotechnol. Bioeng.* 113, 944–952. doi: 10.1002/bit.25860
- Mukund, S., and Adams, M. W. (1991). The novel tungsten-iron-sulfur protein of the hyperthermophilic Archaeobacterium, *Pyrococcus furiosus*, is an aldehyde ferredoxin oxidoreductase. *J. Biol. Chem.* 266, 14208–14216.
- Napora-Wijata, K., Strohmeier, G. A., and Winkler, M. (2014). Biocatalytic reduction of carboxylic acids. *Biotechnol. J.* 9, 822–843. doi: 10.1002/biot.201400012
- Ni, Y., Hagedoorn, P.-L., Xu, J.-H., Arends, I. W., and Hollmann, F. (2012). A biocatalytic hydrogenation of carboxylic acids. *Chem. Commun.* 48, 12056–12058. doi: 10.1039/c2cc36479d
- Noweck, K., and Grafahrend, W. (2006). “Fatty alcohols,” in *Ullmann's Encyclopedia of Industrial Chemistry*, Vol. 14 (Weinheim: Wiley-VCH), 117–141.
- Obeid, J., and White, P. C. (1992). Tyr-179 and Lys-183 are essential for enzymatic activity of 11β-hydroxysteroid dehydrogenase. *Biochem. Biophys. Res. Commun.* 188, 222–227. doi: 10.1016/0006-291X(92)92373-6
- Oppermann, U. C., Filling, C., Berndt, K. D., Persson, B., Benach, J., Ladenstein, R., et al. (1997). Active site directed mutagenesis of 3β/17β-hydroxysteroid dehydrogenase establishes differential effects on short-chain dehydrogenase/reductase reactions. *Biochemistry* 36, 34–40. doi: 10.1021/bi961803v
- Persson, B., Kallberg, Y., Oppermann, U., and Jörnvall, H. (2003). Coenzyme-based functional assignments of short-chain dehydrogenases/reductases (SDRs). *Chem. Biol. Interact.* 143, 144, 271–278. doi: 10.1016/S0009-2797(02)00223-5
- Qu, G., Guo, J., Yang, D., and Sun, Z. (2018). Biocatalysis of carboxylic acid reductases: phylogenesis, catalytic mechanism and potential applications. *Green Chem.* doi: 10.1039/C7GC03046K
- Schaffer, S., Corthals, J., Wessel, M., Hennemann, H.-G., Häger, H., Volland, M., et al. (2014). Preparation of Amines and Diamines from a Carboxylic Acid or Dicarboxylic Acid or a Monoester Thereof. WO/2014/095986. Essen.
- Schwendenwein, D., Fiume, G., Weber, H., Rudroff, F., and Winkler, M. (2016). Selective enzymatic transformation to aldehydes *in vivo* by fungal carboxylate reductase from *Neurospora crassa*. *Adv. Synth. Catal.* 358, 3414–3421. doi: 10.1002/adsc.201600914
- Scrutton, N. S., Berry, A., and Perham, R. N. (1990). Redesign of the coenzyme specificity of a dehydrogenase by protein engineering. *Nature* 343, 38–43. doi: 10.1038/343038a0
- Stachelhaus, T., Mootz, H. D., and Marahiel, A. (1999). The specificity-conferring code of adenylation nonribosomal peptide synthetases. *Chem. Biol.* 6, 493–505. doi: 10.1016/S1074-5521(99)80082-9
- Stolterfoht, H., Schwendenwein, D., Sensen, W. C., Rudroff, F., and Winkler, M. (2017). Four distinct types of E.C. 1.2.1.30 enzymes can catalyze the reduction of carboxylic acids to aldehydes. *J. Biotechnol.* 257, 222–232. doi: 10.1016/j.jbiotec.2017.02.014
- Studier, F. W. (2005). Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* 41, 207–234. doi: 10.1016/j.pep.2005.01.016
- Thibault, J.-F., Asther, M., Cessaldi, B. C., Couteau, D., Delattre, M., Duarte, J. C., et al. (1998). Fungal Bioconversion of agricultural by-products to vanillin. *LWT Food Sci. Technol.* 31, 530–536. doi: 10.1006/food.1998.0411
- Tripathi, C. M., Agarwal, S. C., and Basu, S. K. (1997). Production of L-phenylacetylcarbinol by fermentation. *J. Ferment. Bioeng.* 84, 487–492. doi: 10.1016/S0922-338X(97)81900-9
- Venkatasubramanian, P., Daniels, L., and Rosazza, J. P. (2007). Reduction of carboxylic acids by *Nocardia* aldehyde oxidoreductase requires a phosphopantetheinylated enzyme. *J. Biol. Chem.* 282, 478–485. doi: 10.1074/jbc.M607980200
- Wang, M., Beissner, M., and Zhao, H. (2014). Aryl-aldehyde formation in fungal polyketides: discovery and characterization of a distinct biosynthetic mechanism. *Chem. Biol.* 21, 257–263. doi: 10.1016/j.chembiol.2013.12.005
- Wang, M., and 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. doi: 10.1021/cs500039v
- White, H., Huber, C., Feicht, R., and Simon, H. (1993). On a reversible molybdenum-containing aldehyde oxidoreductase from *Clostridium formicoaceticum*. *Arch. Microbiol.* 159, 244–249. doi: 10.1007/BF00248479
- White, H., Strobl, G., Feicht, R., and Simon, H. (1989). Carboxylic acid reductase: a new tungsten enzyme catalyses the reduction of non-activated carboxylic acids to aldehydes. *Eur. J. Biochem.* 184, 89–96. doi: 10.1111/j.1432-1033.1989.tb14993.x
- Wierenga, R. K., De Maeyer, M. C. H., and Hol, W. G. J. (1985). Interaction of phosphate moieties with alpha-helices in dinucleotide binding proteins. *Biochemistry* 24, 1346–1357. doi: 10.1021/bi00327a012
- Winkler, M. (2018). Carboxylic acid reductase enzymes (CARs). *Curr. Opin. Chem. Biol.* 43, 23–29. doi: 10.1016/j.cbpa.2017.10.006
- Winkler, M., and Winkler, C. K. (2016). *Trametes versicolor* carboxylate reductase uncovered. *Monatsh. Chem.* 147, 575–578. doi: 10.1007/s00706-016-1676-z
- Wood, A. J. L., Weise, N. J., Frampton, J. D., Dunstan, M. S., Hollas, M. A., Derrington, S. R., et al. (2017). Adenylation activity of carboxylic acid reductases enables the synthesis of amides. *Angew. Chem. Int. Ed Engl.* 56, 14498–14501. doi: 10.1002/anie.201707918
- Yang, S.-T., Huang, H., Tay, A., Qin, W., De Guzman, L., and San Nicolas, E. C. (2007). “Extractive fermentation for the production of carboxylic acids,” in *Bioprocessing for Value-Added Products from Renewable Resources: New Technologies and Applications*, ed S.-T. Yang (Amsterdam: Elsevier B.V.), 421–446.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Stolterfoht, Steinkellner, Schwendenwein, Pavkov-Keller, Gruber and Winkler. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.