

University of Groningen

Identification of a conserved N-terminal domain in the first module of ACV synthetases

Iacovelli, Riccardo; Mózsik, László; Bovenberg, Roel A L; Driessen, Arnold J M

Published in:
 MicrobiologyOpen

DOI:
[10.1002/mbo3.1145](https://doi.org/10.1002/mbo3.1145)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
 Publisher's PDF, also known as Version of record

Publication date:
 2021

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Iacovelli, R., Mózsik, L., Bovenberg, R. A. L., & Driessen, A. J. M. (2021). Identification of a conserved N-terminal domain in the first module of ACV synthetases. *MicrobiologyOpen*, 10(1), [e1145].
<https://doi.org/10.1002/mbo3.1145>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

ORIGINAL ARTICLE

Identification of a conserved N-terminal domain in the first module of ACV synthetases

Riccardo Iacovelli¹  | László Mózsik¹ | Roel A.L. Bovenberg^{2,3} | Arnold J.M. Driessen¹ 

¹Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, The Netherlands

²Synthetic Biology and Cell Engineering, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, The Netherlands

³DSM Biotechnology Centre, Delft, The Netherlands

Correspondence

Arnold J.M. Driessen, Department of Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen 9747 AG, The Netherlands.
Email: a.j.m.driessen@rug.nl

Abstract

The L- δ -(α -aminoadipoyl)-L-cysteinyl-D-valine synthetase (ACVS) is a trimodular non-ribosomal peptide synthetase (NRPS) that provides the peptide precursor for the synthesis of β -lactams. The enzyme has been extensively characterized in terms of tripeptide formation and substrate specificity. The first module is highly specific and is the only NRPS unit known to recruit and activate the substrate L- α -aminoadipic acid, which is coupled to the α -amino group of L-cysteine through an unusual peptide bond, involving its δ -carboxyl group. Here we carried out an in-depth investigation on the architecture of the first module of the ACVS enzymes from the fungus *Penicillium rubens* and the bacterium *Nocardia lactamdurans*. Bioinformatic analyses revealed the presence of a previously unidentified domain at the N-terminus which is structurally related to condensation domains, but smaller in size. Deletion variants of both enzymes were generated to investigate the potential impact on penicillin biosynthesis in vivo and in vitro. The data indicate that the N-terminal domain is important for catalysis.

KEYWORDS

antibiotics, *Nocardia lactamdurans*, nonribosomal peptide synthetases, penicillin biosynthesis, *Penicillium rubens*

1 | INTRODUCTION

The discovery of penicillin has been a crucial turning point in human history, completely revolutionizing the treatment of bacterial infections. Since then, many more classes of compounds have been identified or synthesized, but penicillin and its (semi-) synthetic derivatives are still among the most common antibiotics used worldwide.

The biosynthetic pathway of penicillin (and other β -lactam antibiotics such as cephalosporins and cephamycins) has been extensively characterized in the past decades, resulting in detailed knowledge on the key enzymes involved and the mechanistic aspects of its biosynthesis (Baldwin & Abraham, 1988; Coque et al., 1991; Liras, 1999; Martín,

1998; Martín et al., 1999; Ozcengiz & Demain, 2013; Queener, 1990; Roach et al., 1997; Schofield et al., 1997; Tahlan et al., 2016; Wu et al., 2012). Two distinct enzymatic steps are involved in the production of the β -lactam ring. First, the L- δ -(α -aminoadipoyl)-L-cysteinyl-D-valine synthetase (ACVS) provides the linear tripeptide (L,L,D)-ACV, a precursor of all β -lactam antibiotics. The second step requires the enzyme isopenicillin N synthase (IPNS, formerly also ACV cyclase), which catalyzes the formation of the β -lactam ring. The first enzyme of the pathway, ACVS, belongs to the family of nonribosomal peptide synthetases (NRPS), multimodular enzymes that synthesize small peptides (NRP) by a mechanism that does not require messenger RNA, unlike ribosomal protein synthesis (Marahiel et al., 2009; Reimer et al., 2018).

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2021 The Authors. *MicrobiologyOpen* published by John Wiley & Sons Ltd.

NRP synthesis generally starts in every module at the adenylation (A) domain, which recognizes a specific substrate and catalyzes its activation as (amino) acyl-AMP. In this process, ATP is consumed and inorganic pyrophosphate (PP_i) is released. The activated substrate is then transferred via a thioesterification reaction to the thiolation (T)—or peptide carrier protein (PCP)—domain through the phosphopantetheine (ppant) arm, with AMP being released. The ppant is a CoA (Coenzyme A)-derived cofactor, covalently attached to a highly conserved residue of serine of the T domain. The loaded substrates are subsequently transported to the active site of the condensation (C) domain (or a cyclization domain, Cy). Here, peptide formation occurs via nucleophilic attack of the α -amino group of the downstream substrate to the activated carboxy group of the upstream substrate. As a result, the latter is released from the ppant moiety, and the newly synthesized intermediate is now ready to be transported to the next condensation domain (Concurso & Bruner, 2012; Koglin & Walsh, 2009; Marahiel et al., 2009; Reimer et al., 2018; Schwarzer et al., 2003; Strieker et al., 2010; Walsh, 2016).

The domains A, PCP, and C/Cy are strictly required for NRP synthesis, and therefore, a minimal module contains all three of them (initiation modules can lack a C domain as they represent the starter units of NRPS). Also, optional or modifying domains can be present in one or more modules, such as methyltransferase (De Mattos-Shibley et al., 2018), epimerization (Samel et al., 2014), formylation (Schoenafinger et al., 2006), and oxidase (Schneider et al., 2003) domains. These domains contribute to further increase the chemical diversity of NRPs and often can provide advantageous characteristics such as increased chemical stability or resistance against peptidase. Lastly, when the synthesis is completed, the peptide can be released via the activity of a thioesterase (Te) domain, either by macrocyclization or hydrolysis, resulting in a cyclic or a linear NRP, respectively (Horsman et al., 2016; Kohli & Walsh, 2003; Schneider & Marahiel, 1998; Tanovic et al., 2008).

Because of the relevance of the β -lactam biosynthetic pathway and the extensive research covering it, the ACVS became soon a model to investigate nonribosomal peptide biosynthesis. The protein has been studied in filamentous fungi such as *Penicillium*, *Aspergillus*, and *Cephalosporium* as well as bacterial *Nocardia* and *Streptomyces* species, both in vivo and in vitro (Baldwin et al., 1990, 1991; Coque et al., 1996; Iacovelli et al., 2020; Jensen et al., 1990; van Liempt et al., 1989; Schwecke et al., 1992; Tahlan et al., 2017; Theilgaard et al., 1997). Despite the lack of 3D structures, it is well known that the ACVS is a trimodular NRPS with a reported canonical domain arrangement: AT₁-CAT₂-CATETe₃. Each module is responsible for the activation and incorporation of its specific substrate into the final product. The three substrates L- α -amino adipic acid (L- α -Aaa), L-cysteine, and L-valine (epimerized to D-valine by the E domain) are assembled in a co-linear fashion; thus, the position of the incorporated substrate corresponds to the position of the respective module within the NRPS sequence. Interestingly, the adenylation domain of the first module of ACVS is the only NRPS domain that is known to activate and utilize L- α -amino adipic acid as substrate (Flissi et al.,

2020). Besides, L- α -Aaa is adenylated on the δ -carboxyl group, resulting in a non-canonical peptide bond formation between L- α -Aaa and L-Cys (Iacovelli et al., 2020), suggesting a substrate recognition mechanism different from that of canonical amino acid-activating adenylation domains (Conti et al., 1996, 1997; Stachelhaus et al., 1999).

In this work, we set out to further investigate the architecture of the first module of the ACVS from the fungus *Penicillium rubens* and the soil bacterium *Nocardia lactamdurans* (also known as *Amycolatopsis lactamdurans*). Through bioinformatic analyses, we observed the presence of a previously unidentified domain related to the condensation domain family, in the N-terminal region of the enzyme. The domain appears structurally conserved and shows the typical fold of C domains, despite being approximately half the size. We therefore proceeded to generate variants of both ACVS enzymes that lack this domain, to investigate the potential impact on ACV biosynthesis, and thus on the whole β -lactam biosynthetic pathway.

2 | MATERIALS AND METHODS

2.1 | Strains, plasmids, and general culturing conditions

All cloning procedures in *E. coli* were performed using the strain DH10 β for high transformation efficiencies. Cultures were grown using LB medium at 37°C and 200 rpm, and the antibiotic selection was conducted utilizing 25 μ g/ml Zeocin (phleomycin D1). The *Nocardia lactamdurans pcbAB* gene was cloned using an intermediate gateway vector and subsequently subcloned into a custom-made pBAD-plasmid (pBR322 ori; araC; pBAD, ZeoR) using SbfI x NdeI sites, and including the introduction of a 6xHis-tag on the C-terminal end. This construct was kindly provided by DSM Sinochem Pharmaceuticals (now Centrient Pharmaceuticals), the Netherlands.

Penicillium rubens DS68530 (Δ penicillin biosynthetic cluster, Δ hdfA, derived from DS17690) (Harris et al., 2009; Salo et al., 2015) and DS54468 (1 \times penicillin biosynthetic cluster, Δ hdfB, derived from DS47274) (Nijland et al., 2010) strains were kindly provided by DSM Sinochem Pharmaceuticals. For storage purposes, control and transformant strains were grown on a complex medium for 5–7 days, after which the mycelium was collected and stored on rice grains (Kovalchuk et al., 2012). For liquid culturing, fungal strains were grown in YGG and PPM (penicillin production medium), prepared as described previously (Weber et al., 2012).

Synthetic DNA was purchased from Invitrogen (GeneArt Strings). PCR primers and sequencing primers, as well as chemicals used for the preparation of stock solutions and media, were purchased from Sigma-Aldrich (now MilliporeSigma). PCR reactions were carried out using KAPA HiFi HotStart ReadyMix (Roche), according to the manufacturer's instructions. In silico PCR and cloning procedures, as well as subsequent analyses, were performed using the SnapGene[®] software (from GSL Biotech; available at snapgene.com).

2.2 | Bioinformatic analysis

The sequences of *Ni* ACVS and *Pr* ACVS were acquired from the database Uniprot (Bateman, 2019), with the respective accession numbers being P27743 and P26046. For domain prediction analyses, the web software BLASTP (version 2.10.1+) was used. The search parameters were set as follows: database="non redundant protein sequences (nr)"; search algorithm="blastp (protein-protein BLAST)"; low complexity filter = no; Composition Based Adjustment = yes; E-value threshold = 0.01; maximum number of hits = 500. The results were visualized using the graphical interface of BLAST (conserved domains on). Backbone dynamics prediction—to identify folded and unfolded regions—was performed using the web-based software Dynamine (Cilia et al., 2013, 2014).

Structural models of *Ni* ACVS M1 and *Pr* ACVS M1 were obtained using the webserver i-TASSER (Iterative Threading ASSEmby Refinement), a fold recognition-based protein modeling tool (Roy et al., 2010; Yang et al., 2014; Yang & Zhang, 2015). The highest-ranking model was selected for each of the targets based on their C-score. The C-score (range -5,2) is a confidence score that estimates the quality of the predictions, higher is better. The models were then visualized using the software UCSF Chimera (Pettersen et al., 2004) for imaging. Structural alignments were generated using the MatchMaker suite of UCSF Chimera, with the following parameters: chain pairing=best aligning pair of chains; alignment algorithm=Needleman-Wunsch; Matrix=BLOSUM-62; gap opening penalty=12; gap extension penalty=1.

Ultimately, multiple sequence alignment analyses and identification of the core sequence motifs C1-C7 were achieved using the software Clustal Omega (Sievers et al., 2011).

2.3 | CRISPR/Cas9-mediated deletion of the C* domain in *P. rubens* ACVS

The plasmid pLM-AMA15.0 (AddGene ID #138944) was constructed using Modular Cloning (Weber et al., 2011) and Gibson Assembly (Gibson et al., 2009), as described previously (Mózsik et al., 2020). The sequence encoding *SpCas9*-NLS (nuclear localization sequence) was amplified from pYTK036 (yeast MoClo toolkit, AddGene ID # 65143). The gene was placed under control of the p40S promoter (AN0465, 40S ribosomal protein S8 promoter), amplified from pVE2_10_p40S_QfDBD_VP16AD_cp_penDE_DsRed_0xQUAS (AddGene ID # 154228) (Mózsik et al., 2019). For the sgRNA transcription unit, the hammerhead (HH) and hepatitis delta virus (HDV) ribozyme sequences (self-cleavage), *gpdA* (AN8041) promoter, and *trpC* terminator were amplified from pFC334 (AddGene ID # 87846) (Nødvig et al., 2015). The terbinafine resistance gene, *ergA*, was amplified from pCP1_45 (Pohl et al., 2020). The fungal AMA1-based vector (*bleR*, phleomycin resistance) was repurposed from vector pDSM-JAK-109 (Bovenberg et al., 2016). The resulting vector delivers the expression of both the Cas9 protein and the sgRNA (Figure 3). The DNA insert encoding the sgRNA unit targeting C* *Pr*ACVS was generated via PCR using two primers

with overlapping regions (Table A1). A Golden Gate assembly reaction was performed to replace the *lacZ* gene of pLM-AMA15.0 with the aforementioned insert and cloned into DH10 β *E. coli* competent cells. Positive transformants were identified using blue-white screening and were confirmed by sequencing (Macrogen Europe B.V.). Marker-free donor DNA for the deletion of C* was generated by PCR using primers with 106 bp homology toward pACVS (5' of ATG codon) and ACVS ORF downstream of C*ACVS (Figures 3b, 4c, Table A1).

Protoplasts of *P. rubens* DS54468 were obtained as described previously (Pohl et al., 2016). For the transformation, 25 μ g of marker-free donor DNA was mixed with 1 μ g pLM-AMA15.x CRISPR/Cas9 plasmid targeting C*ACVS (Figure 3b). The fungal selection was carried out on solid medium plates supplemented with 1.1 μ g/ml terbinafine hydrochloride (MilliporeSigma). The genotype of positive transformants was confirmed by colony PCR using Phire Green Hot Start II PCR Master Mix (Thermo Fisher Scientific) and sequencing (Macrogen Europe B.V.). For storage and further analysis, the strains were purified by performing two rounds of sporulation on the terbinafine selection medium.

2.4 | Antibacterial plate assays

Agar plate-based antibacterial assays were set up as follows: an overlay of soft LA-agar (1%) inoculated with *Micrococcus luteus* (sensitive to β -lactams) to an OD600 of 0.125 was poured on top of a pre-solidified agar (1%) bottom layer. Before that, Oxford Towers (8 \times 10 mm) were spaced out evenly on top of the bottom agar layer to generate four wells in the top layer. The Oxford Towers were then removed aseptically, and the wells were inoculated with 10 μ l of supernatant samples. The plates were then incubated at 30°C for 24 h. For imaging, a LAS-4000 system (Fujifilm Life Science) was utilized.

2.5 | Liquid culturing of *P. rubens* and sampling

For penicillin production, spores (immobilized on 25 rice grains) of both control and transformant strains were inoculated in 25 ml YGG medium and grown for 24 h in shake flasks (25°C, 200 rpm). The cultures were then diluted eight times into 25 ml PPM containing 2.5 g/L phenoxyacetic acid (POA) and grown in the same conditions for 5 days. Supernatant samples for HPLC/MS analysis were taken and filtered using 0.45- μ m syringe filters. The samples were subsequently stored at -20 or -80°C for short-term and long-term storage, respectively, before HPLC/MS analysis. The experiment was performed in triplicate.

2.6 | Engineering the C* deletion variant of the *N. lactamdurans* ACVS

The deletion variant of *Ni* ACVS was obtained via a two-step cloning strategy. Firstly, we amplified the deletion insert from a Golden Gate

(GG) intermediate vector bearing the pBAD promoter region and a fragment of the *pcbAB* gene. The forward primer was designed to bind just outside of the 3' end of the C* domain (AA position 237). On the 5' tail, homologous sequence to the RBS (ribosome-binding sequence) of the promoter and an appropriate restriction enzyme site were added. The reverse primer only contained a homologous sequence and an appropriate restriction site (Table A1).

The deletion insert was then subcloned into the same template intermediate vector, resulting in a Golden Gate-ready vector now lacking the C* domain (AA positions 2–236). The product was checked by restriction analysis and sequencing (Macrogen Europe B.V.). Lastly, the pBAD-*Nl* Δ C*-ACVS plasmid was assembled using the Golden Gate assembly reaction (Engler et al., 2009) as described previously (Iacovelli et al., 2020) and cloned into DH10 β chemically competent *E. coli* cells for storage.

2.7 | Overexpression and purification of *NL* ACVS and *NL* Δ C*-ACVS

For overexpression and purification, the plasmids were cloned into the expression strain *E. coli* HM0079 (Gruenewald et al., 2004). The strains carrying both the wild-type and the deletion variant were inoculated overnight in 3 ml selective LB. The cultures were then diluted 1:100 into fresh 2xPY medium (10 g/L NaCl, 10 g/L yeast extract, 16 g/L Bacto™ tryptone), grown to an OD₆₀₀ of 0.6, transferred to 18°C and 200 rpm for 1 h and subsequently induced using 0.2% L-arabinose. After 18 h, the cells were harvested by centrifugation at 4000 g for 15 min. After resuspension in lysis buffer (HEPES 50 mM pH 7.0, 300 mM NaCl, 2 mM DTT, complete EDTA free protease inhibitor 1 cp/10 ml; Roche no. 04693159001), cells were disrupted using sonication (6 s/15 s on/off; 50x; 10 μ m amplitude; Soniprep 150 MSE) and the cell-free lysate was obtained by centrifugation at 4°C, 17,000 g, 15 min. Purified enzymes were extracted using Ni²⁺ affinity purification (1 ml slurry–500 μ l resin–per 1 L of culture) using disposable gravity-flow columns. Wash steps were performed using 10 column bed volumes each of wash buffer (HEPES 50 mM pH 7.0, 300 mM NaCl, 20 mM imidazole) followed by a three-step elution using four-bed volumes of each elution buffer (HEPES 50 mM pH 7.0, NaCl 300 mM, imidazole 50–150 or 250 mM). Fractions were then analyzed on SDS-PAGE 5% mini-gels and imaged. 2D densitometry analyses (AIDA Image Analysis software, Elysia-raytest) were followed to determine the concentration of the deletion variant relative to the wild-type ACVS. Lastly, the purified wild-type and deletion variants were, respectively, diluted and concentrated using Amicon U-100 spin filters (Merck). The final concentration for wild-type ACVS was determined using A280 (NanoDrop 1000; Thermo Fisher Scientific).

2.8 | In vitro peptide formation assays

Diluted/concentrated *Nl* ACVS and *Nl* Δ C* ACVS were subjected to in vitro peptide synthesis assays, to determine product formation.

Reaction mixture was set up as follows: HEPES 50 mM pH 7.0, 300 mM NaCl, 5 mM ATP pH 7.0, 5 mM L- α -aminoadipic acid, 2 mM L-cysteine, 2 mM L-valine, 5 mM MgCl₂, 2 mM DTT, and ~0.25 μ M ACVS/ Δ C* ACVS. The total volume was set at 200 μ l. Control reactions were set up analogously, but omitting ATP or L- α -aminoadipic acid. Reactions were run at 30°C and sampling took place at 0 (T₀) and 300 (T_{5h}) minutes for endpoint value determination. NaOH was added to each sample to a final concentration of 0.1 M to quench the reactions. Samples were subsequently stored at –80°C. Before HPLC/MS analysis, the samples were thawed in ice, reduced by adding DTT to a final concentration of 10 mM, and filtered.

2.9 | High-performance liquid chromatographic and mass spectrometric analysis (HPLC/MS)

Samples (50 μ l) were obtained from fungal liquid cultures and the in vitro reactions were subjected to HPLC/MS analysis, with an injection volume of 5 μ l each. The analysis was performed using an HPLC/MS Orbitrap (Thermo Scientific) in combination with a Reverse Phase-C18 column (ACQUITY UPLC BEH C18, 130 Å, 1.7 μ m, 2.1 mm \times 150 mm). Scan range was set at 100–2000 m/z in positive ion mode (ion optics voltages: 4.1 kV spray, 24.4 V capillary, and 73.7 V tube lens), with the capillary temperature set at 325°C. For the in vitro samples, a 24-min gradient program with Milli-Q water (A), acetonitrile (B) and 2% formic acid (C) was utilized: 0 min, A 90%, B 5%, C 5%; 4 min, A 90%, B 5%, C 5%; 13 min, A 0%, B 95%, C 5%; 16 min A 0%, B 95%, C 5%; 16.01 min, A 90%, B 5%, C 5%; 24 min A 90%, B 5%, C 5% at a flow rate of 0.150 ml/min. For the fungal (in vivo) samples, a 60-min gradient program with Milli-Q water (A), acetonitrile (B) and 2% formic acid (C) was utilized: 0 min, A 90%, B 5%, C 5%; 5 min, A 90%, B 5%, C 5%; 30 min, A 35%, B 60%, C 5%; 35 min, A 0%, B 95%, C 5%; 45 min, A 0%, B 95%, C 5%; 45.01 min, A 90%, B 5%, C 5%; 60 min, A 90%, B 5%, and C 5% at a flow rate of 0.300 ml/min. The Bis-ACV standard was obtained from Bachem, reduced to (L,L,D)-ACV, and used for peak identification. The Penicillin V standard was purchased from Sigma-Aldrich (now MilliporeSigma) and used for quantification in a standard curve at concentrations of 0.0005, 0.0025, 0.0125, 0.0625, 0.3125, 0.625, and 1.25 g/L. The CV dipeptide and the pigment chrysogine (control for fungal experiments, Figure A1) were identified according to their predicted monoisotopic masses, in the absence of synthetic standards.

3 | RESULTS AND DISCUSSION

3.1 | Identification and modeling of a conserved domain in the first module of ACV synthetases

Through domain prediction and backbone dynamics prediction (Figure A2), we identified a structurally organized region at the N-terminus of the first module of a bacterial and a fungal ACVS, respectively, from *Nocardia lactamdurans* (*Nl*) and *P. rubens* (*Pr*) (formerly annotated as *Penicillium chrysogenum*). Initial BlastP analyses

show the significant alignment of the amino acids sequence 2–236 (*Nl* ACVS) and 222–308 (*Pr* ACVS) to the condensation domain family (pfam00668), with a respective E-value of 2.35E-03 and 3.68E-03. The N-terminal regions vary in size between 230 amino acids for *Nl* ACVS and 300 amino acids for *Pr* ACVS, while the size of a full condensation domain is generally around 450 amino acids. Because of their alignment with the condensation domain protein family, we refer to the two regions as C* domain from now on (Figure 1a). A similar region has been previously observed in a fungal NRPS (Berry et al., 2019) and in fungal α -amino adipic acid (Lys2-type) reductases such as Lys2 from *Penicillium chrysogenum* and NPS3 from *Ceriporiopsis subvermispota* (Hijarrubia et al., 2003; Kalb et al., 2014, 2015). Interestingly, the reductases and the first module of of

ACVS activate the same substrate, and it has been shown that the N-terminal domain of NPS3 is critical for adenylation activity; hence, it was termed the adenylation activating (ADA) domain (Kalb et al., 2015).

To gain further insight on the C* domains from ACVS, we submitted the sequences of the first module of the two enzymes to DynaMine, an online tool for the prediction of backbone dynamics. Starting from the primary sequence as input, DynaMine predicts backbone flexibility at the residue-level and outputs backbone N-H S² order parameter values. These values can range between 0 and 1, indicating, respectively, a highly dynamic residue (therefore part of an unstructured/unfolded region) or a residue with a stable conformation (meaning the residue is likely part of a highly structured/

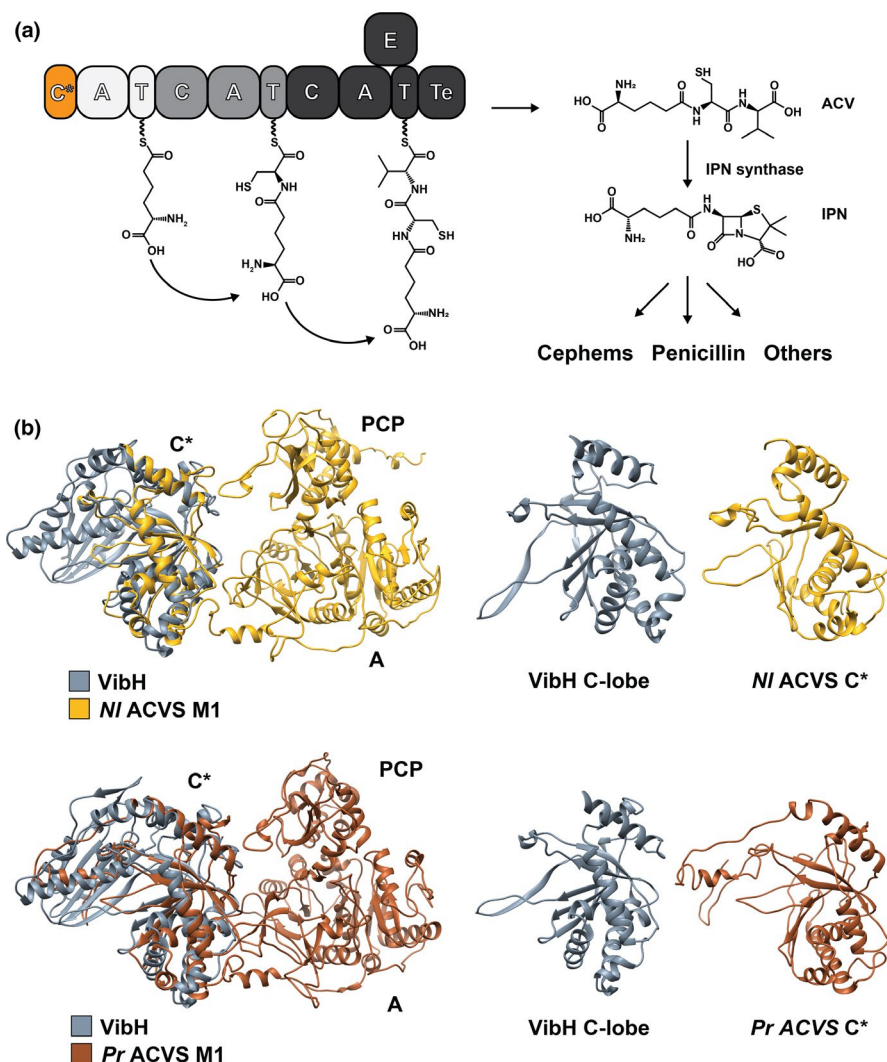


FIGURE 1 (a) Domain organization and reaction scheme of ACVS. The ACV tripeptide is synthesized from L- α -amino adipic acid, L-cysteine, and L-valine. In the next step, ACV is converted to isopenicillin N (IPN) by the IPN synthase, which catalyzes the formation of the β -lactam ring. Following different biocatalytic routes, IPN is further converted into several classes of β -lactam antibiotics. The newly identified N-terminal C* domain is highlighted in orange. (b) Modeling of *Nl* ACVS M1 and *Pr* ACVS M1. In both cases, all the canonical domains can be observed, with their expected structural organization. Also, at the N-terminus of the modules, we observe a highly structured domain (C*) that aligns very well with model condensation domain VibH (PDB: 1L5A). Specifically, C* domains align with the C-lobe of VibH and completely lack elements from the N-lobe, including the active site motif HHxxxDG. *Pr* ACVS M1 C* shows an additional disorganized region at the N-terminal portion (panel B, bottom right), which correlates well with the domain and backbone prediction analyses

folded region). The result is a graph where the S^2 values are plotted against the respective residues, making it possible to identify (predicted) rigid or flexible regions (Figure A2). In accordance with what we previously observed with BlastP, two structured regions were predicted in the N-terminal portion of the proteins, with *Pr* ACVS showing an extra stretch predicted to be more flexible and disorganized. Furthermore, we extended the domain prediction analysis to other ACVSs from different organisms, which all showed similar N-terminal regions (Figure A3). For this work, we focused on *Nl* ACVS and *Pr* ACVS.

We therefore proceeded to generate structural models of the initiation modules of the two ACV synthetases, utilizing the web software i-TASSER, a fold recognition-based protein modeling tool. For each of the targets, i-TASSER generated several models based on a database of fold templates. Only the model with the highest C-score (0.83 for *Nl* ACVS M1 and 0.80 for *Pr* ACVS M1) was selected for each target (see corresponding methods section for further details). Given the strong structural conservation of NRPS enzymes, in both cases, the adenylation domains and the PCPs can easily be identified and appear organized in a canonical manner (Figure 1b). In the N-terminal regions, a third domain can be observed, with well-conserved structural motifs. The structures were then aligned with the crystal structure of VibH (Keating et al., 2002) (PDB ID: 1L5A), a model condensation domain, using the structure alignment tool of UCSF Chimera (Pettersen et al., 2004). The C* domain of ACVS superimposes almost perfectly with half of VibH, specifically with the C-terminal lobe (Figure 1b). The C* domain of *Pr* ACVS has an extra sequence at the N-terminus—about 60 amino acids in length—which appears to be rather disorganized and does not align with VibH (Figure 1b). C domains are pseudo-dimers constituted of two lobes, N-lobe and C-lobe, organized in a V-shape structure (Bloudoff et al., 2013; Bloudoff & Schmeing, 2017; Keating et al., 2002). It seems that the C* domains from ACVS enzymes conserve the structure of an isolated C-lobe, but completely lack the N-lobe. The canonical active site motif of condensation domains, HHxxxDG (De Crecy-Lagard et al., 1995; Marshall et al., 2002), as well as the acceptor PCP binding site, are found in the N-lobe (Bloudoff et al., 2013; Bloudoff & Schmeing, 2017; Keating et al., 2002) and therefore are missing in C* domains. These results are in accordance with the observations of Kalb et al. (2015) on a structural model of the ADA domain of NPS3, further confirming the similarities between these domains. Subsequent multi-alignment sequence analysis using Clustal Omega revealed that many of the other conserved motifs of C domains (Marahiel et al., 1997; Marxen et al., 2015; Meyer et al., 2016; Rausch et al., 2007) are also absent in C* and ADA (C1–C5), while some are partially retained (C6–C7) (Figure 2). These findings strongly suggest that the C* domains do not exert

the catalytic function typical for condensation domains during the enzymatic synthesis of ACV. This region is, however, highly conserved structurally and omnipresent in ACVS enzymes and thus likely fulfills another role.

3.2 | CRISPR/Cas9-mediated deletion of C* in *P. rubens* ACVS abolishes the production of penicillin antibiotics

Because of the striking presence of this structural domain in the family of ACVS proteins, we investigated the potential function of the C* domain of *Pr* ACVS. For this purpose, we created a *P. rubens* strain lacking the C* domain. To do so, we used a novel CRISPR/Cas9 approach based on the delivery of a custom-made plasmid carrying both the genes encoding Cas9 and the target-specific sgRNA, as well as an anti-fungal (terbinafine) resistance marker (Figure 3). The transformants were first selected for the presence of the plasmid and further analyzed for the correct introduction of the designed C* domain genome deletion by targeted PCR and DNA sequencing. Two strains carrying the desired deletion, C1, and C2, were selected for liquid culturing experiments and HPLC/MS analysis. Also, two control strains were included: the parental penicillin producer strain containing one copy of the penicillin biosynthetic cluster (DS54468), and a non-producer strain (DS68530) lacking the entire penicillin biosynthetic cluster.

For liquid culturing, all the strains were grown in shake flasks in PPM containing POA, a precursor of penicillin V. After 5 days, the supernatant was collected, filtered, and stored. Before HPLC/MS analysis, the samples were subjected to an antibacterial plate assay to screen for the presence of β -lactam compounds (Figure 4a). The assay indicated the absence of antimicrobial activity in the samples from strains C1, C2, and the negative control, whereas the positive control showed the formation of a large halo. Analysis via HPLC/MS confirmed these results (Figure 4b). The production of penicillin V is abolished in the strains C1 and C2, suggesting that the deletion of C* rendered ACVS nonfunctional, preventing the formation of the antibiotic compound.

3.3 | Deletion of C* in *N. lactamdurans* ACVS abolishes in vitro synthesis of ACV, while the dipeptide CV is produced at wild-type levels

The study in *P. rubens* does not reveal whether the deletion of the C* domain causes a functional defect or rather disrupts the expression of the ACVS. To investigate this possibility, we then focused on *Nl* ACVS, which is very well overexpressed in *E. coli* and has already

FIGURE 2 Multiple sequence alignment performed using Clustal Omega. The sequences of *Nl* ACVS M1 C* *Pr* ACVS M1 C* were aligned to those of several known condensation domains (obtained from the Uniprot database). The ADA domain from the α -amino adipate reductase NPS3 was also included in the alignment. The main residues of core motifs C1–C7 of the condensation domain family are highlighted in red. The active site motif, C3, and motifs C1–C5 are completely lacking in both C* domains and ADA, while C6 is retained only in the case of NPS3 and C7 in all cases

C1

Pc_ACVS_M1_C*
 NocB_C1
 SrFAC_C1
 GrSB_C1
 NI_ACVS_M1_C*
 NI_ACVS_M2_C
 NPS3_ADA
 EntF_C
 VibH
 CDA_C1

-----VVDVYPLAPMQEGMLFRGLFWP--DSDAYFNQNVLELVGGTDEDVLRWARVRADR 54
 -----VQDMYYLSPMQEGMLFHAILNP--GQSFYLEQITMKVKSGLNKLGLSESNVMDR 54
 -----VQDMYRLSPMQEGMLFHALLDK--DKNAHLVQMSIAIEGTVDVELLESNILDR 54
 -----VARVLANGLQQGLLYHHLKTAGDDAYVQSVHRYHAPTRPELMDQAAARQT 54
 -----MSQHLPLVAQPGIWMAEKLE--LPSAMSAHYVELTGEVDSPLLARAVVAGLAQ 54
 -----MLLAQKPFMQRHLAYP--HINLDTVAHSLRGTPLDHTLLRALHLTVSE 48
 GMSNSVVRHGLTSAQHEVWMLAQQLDP--RGAHYRTGSCLEIDGPLDHAVALSRALLRTVAG 59

C2

Pc_ACVS_M1_C*
 NocB_C1
 SrFAC_C1
 GrSB_C1
 NI_ACVS_M1_C*
 NI_ACVS_M2_C
 NPS3_ADA
 EntF_C
 VibH
 CDA_C1

-----YEILRTGFWDRAPLQYVR--DDLAP-----PWALLDWSGED--PDRRA 96
 YDVFRTVIEHKVRRPVQVWL--KRRQF-----HIEEIDLTHLT--GSEQT 96
 YDVFRTTFLHEKIQPLQVVL--KERPV-----QLQFKDISSLD--EEKRE 96
 YPALRLRFDMAE--EPVQIVDNDKPF-----DMRFVDSLATADDAEQE 96
 ADTLRMRFTEDMG--EVMQWVD--DALTF-----ELPEIDLRNTDHPGT-- 96
 IDLFRARSAGQE--LYNHPPS--PPIDYQ-----DLSHLEAPEL-- 85
 TETLCSRF LTDEEGRPYRAYC--PPAPEGSAVEDDGGVPTVPLLRHIDLSDGHEDPEGE-- 117

C3

Pc_ACVS_M1_C*
 NocB_C1
 SrFAC_C1
 GrSB_C1
 NI_ACVS_M1_C*
 NI_ACVS_M2_C
 NPS3_ADA
 EntF_C
 VibH
 CDA_C1

-----MTLQKPPNGTPIGFS--A-----TTLSLMSGSSVKNKGTIK 35
 ELLERLADRAAGLSLA--DCPLRYRLTLVRLGGRRHYLLWSHHILLDGMWCLSLTGDVDF 155
 AKINYEKQDKIRGFDLT--RDIPMRAAIFKKAEEFEMWMSYHHIILDGMWCFVQVDFL 155
 QAIQEQYQDGETVFDLT--RDPMLRVAIFQTGKVNQWMSYFHIIILMDGWCNFIIFNDLF 155
 ARVRELQERDRTEPYDLA--GGRLFRVYLKQREDLFSLIFSCHHILLDGMWLSLPHVDEHV 155
 --AQAQMQADLQDRLVDSGKPLVFLHQLIQVADNRWYQRYHLLVDGFSFPATRQIA 154
 --AMRQIEQDLQRSSTLI--DAPITSHQVYRLSHSEHLIYTRAHVHVLIDGTYGMFLQERLS 142
 --AQRMDRDRATPLD--RPGLSHALFTLLGGRRHYLLVGLVHVHVIDGTSMWLFYERLA 174

C4

Pc_ACVS_M1_C*
 NocB_C1
 SrFAC_C1
 GrSB_C1
 NI_ACVS_M1_C*
 NI_ACVS_M2_C
 NPS3_ADA
 EntF_C
 VibH
 CDA_C1

PSNG-----IFKPSTRDMDPCGNAADGGSIRV---RFRGGIERMKECVQV--PERCDLSGL 88
 RYVEELAHGREPRLEVPVRYDYV--RMLTRARAARSEESERFWREYLRLDTPSPFSTAPA 213
 KYVNALREQKPYSLPPVKPYKDYI--KMLLEKQD--KQASLRMYREYLEGFEQGTFAEQR 211
 NZYLSLKEKPLQLEAVQPYKQFI--KMLLEKQD--KQALRYMKEHLMWYDQSVTLPKKK 211
 RNYLALRAGQIESDVNAVAAQ--RYMEAH---RNDHAAYVVEQLGRID--ERGFAGL 209
 NIYCTMLRGEPTPASPTFADWVEEYQYRESEAMQDAAAFMAEQRRQL--PPPASLSPA 213
 QHYQSLSGQT--PTAAFQYQSYLEEAAYLTSRHYQDKQFMQGYLREA--PDLTITSAT 200
 EYVRRALRDRGAVPAAAFGOTDRMVAAGEEYRARSARYERDRAYMTGLFTDR--PEPVSLTGR 233

C5

Pc_ACVS_M1_C*
 NocB_C1
 SrFAC_C1
 GrSB_C1
 NI_ACVS_M1_C*
 NI_ACVS_M2_C
 NPS3_ADA
 EntF_C
 VibH
 CDA_C1

TTDSTRYQLASTGFGDASAAVQ---ERLMTVPVDVHA--ALQELCLERRVSVGSVNFVSVH 144
 -----DREGEF-----D---TVWHEPEQLTR--LVNEAARGVYTTNAISQAAMS 254
 -----K--KQKDYEP-----K---ELLFSPSEAEATK--AFTELAKSQHTLSTALQAVMS 255
 -----AAINNTTYP-----A---QFRFAFDKVLTO--QLRIZANQSQVTLNIVFQITWG 256
 LKDEMRHVRVSDYDSTAVRATQ---EKELTISGQDYT--ALKQA--LGAMPLEAFALATLH 82
 LNEKSRVSLVGDYHVQRHRT---RKLVLADTG--ALKAGCAADQVTLHSLVQFVWH 264
 PRPTSGKRFVEAAFTADLSEQTCLSLKALYNEGEDEEEDVPNKRTPSAPHLLLSSTFS 85
 PLP--GRSA-----SAD---ILRKLKLEFDGFE--RQLATQLSGVQRDLALALAA 257
 YDP---QLS-----HAY---SLSYTLNSQLNHV--LLKLANAQIGMPDALVALCA 243
 GGG---RAL-----AVT---VRSGLNPPERTVE--LGRAAEATGAHWARVVIAGVA 276

C5

Pc_ACVS_M1_C*
 NocB_C1
 SrFAC_C1
 GrSB_C1
 NI_ACVS_M1_C*
 NI_ACVS_M2_C
 NPS3_ADA
 EntF_C
 VibH
 CDA_C1

QMLKGFNGTHTTITASLHREQLNQNSPSSW-----VVSPTIVTHNRDQMSVAQAVESI 198
 MLLALQGTGARDVHGLTIAGRPPELEGSEHMGVIFINTPIRLRL--DPDQTVLGLLETTI 312
 VLISRYQSGDGLAFGTWVSGRPAEIKGVEHMGVGLFINVPRRKL--SEGITFNGLLKRL 313
 IYVLYQYNSHNWVYGVSVSGRPSSEISGIEKMGVLFINTLPLRIQT--QKQDSFIELVKTV 314
 SVLHAYGHGQITVAVFLRDKGKVL-----PVVVDHLEQAGLTCAEAAEQ 126
 KVLHATGGTNTWGTIVSGRNLVDPGIENSAAGLFINLPLIVDHIHQOAGONVAEAVRDI 324
 VLLHRYTGDTLVGS-----S--SAS---AREPLILRLAV--NPPDPFWAVRVK 129
 LMLGRLCNRMWYAAGFFMRRLGSA--ALATGPPVNLWPLGHIH--AAQETLPELATRL 313
 LYLES-A--EPDAPMLWLPFMRWGSV--AANVPLMWSNLPLRLSL--AQQTSLGNVLYKQS 298
 AFLHRTTGARDVWVSVVPTGRYGAN--ARITTPGMVNSNRLPLRLAV--RPGESFARVVEV 332

C6

Pc_ACVS_M1_C*
 NocB_C1
 SrFAC_C1
 GrSB_C1
 NI_ACVS_M1_C*
 NI_ACVS_M2_C
 NPS3_ADA
 EntF_C
 VibH
 CDA_C1

EAGRSEKE--SV--TAI---DSGSSLV---KMGLFDLLLVSVFYDADADARP----- 239
 HRDLATTAAGHGVPLAQIKSWASGERLS---GGRVFDNLVAFENVPEDNLPG---DDL 364
 QEQVLSQEPHQVPLYDIQSQAD-----QPKLIDHITVFNVPYLPQDAKN--EESSEN 363
 HQNVLFSSQHEVYFLYETQNHTEL-----KQNLIDHIMVNIENYPLVEELQKNSIMQKV 367
 EDVAREDM-----YLPPEELL---QRGLFDALLVLADGHLG----- 160
 QAAVNTMNSKSTVLEGRLOQSGEM-----KRRLFDTLVLENYRLLDEEE--ELAHQE 375
 QQVESEA--EADHLPYDITIMHALGRDKAEPDMSGPLFRVRFDETD-----APQ-- 177
 AAQLKKMRRHQRYDAEQIVRDSGRAA--GDEPLFGVLIKIVFYDQLDIPDQA----- 365
 GQAIRSLYLHGRYRIEQIEQDQGLMA--EQSYMSPFINILPFESEPH--FADCCQT----- 349
 SEAMSGLLAHSRFRGEDLDRELGGAG-----VSGTYVNVMPYIRPVDFGGPVG----- 380

C7

Pc_ACVS_M1_C*
 NocB_C1
 SrFAC_C1
 GrSB_C1
 NI_ACVS_M1_C*
 NI_ACVS_M2_C
 NPS3_ADA
 EntF_C
 VibH
 CDA_C1

-----CFDFPLAVIVR-----ECDANLSLTLRFSDCLFNEETICNFTDAL 279
 SIQITDLHA--QEKTEYPIGLIVL-----PLQKLAHFHYDVTTHTRAQVDRFIEVF 414
 GFDVMDVHV--FEKSNYDLNLMAS-----PGDEMILIKLAYNENVDFEAFILRLK5QL 413
 GFTVRDVKM--FEPTNYDMYVMVL-----PRDEISVRLDYNAAYVDFIDFKIKTEGHM 417
 ---FTEL-----PPAPLVTIVRD-----DPAAGLHMRYAYAGEFFEDKIAGVLDA 205
 ALRFKAYD--ADKVDYPIAVVAR-----EEGDELTVTLWYAGELFDEIDITDLDDVA 426
 ---ETFMQSTSTSDLVFVZTRQZSSRSLTPHISLRLYNSLFTPARLTCIIDQL 232
 ---QHTLATGPVNDLELALFP-----DVHGDLSIELANKQRYDEPTLIQHAERL 413
 ---ELKVLASGSAEGINTFRG-----SPQHELCLDITADLASYPQSHQSHCHERF 397
 ---LMSISSGPTTDLNIVLTG-----TPESGLRVDFEGNPQYVGGQDLTVLQERF 428

NIILAEAVTGRVTPVADIELLSAEQKQQL----- 308
 RLLTEQLCRRPSARLSELELLPAADRELL----- 443
 LTAIQQLIQNPDPQVSTINLVDDRERFLL----- 443
 KEVALCVANPHVLVQDVPVLLTKQEKQHL----- 447
 REVLGQFVIGRPEQLVADIDLVSAAEQELQL----- 234
 RTLFRKVTEDTARVRELDLISPMRA----- 453
 SVLLRKAANPLAQV----- 247
 KMLIAQFADPALLCGDVMPLPGEYAQLA----- 443
 PRFFEQLARQEQVQDVARLAEAPALAAATSTRAIAS----- 436
 VRFLEAALDPAATVDEVALT----- 450

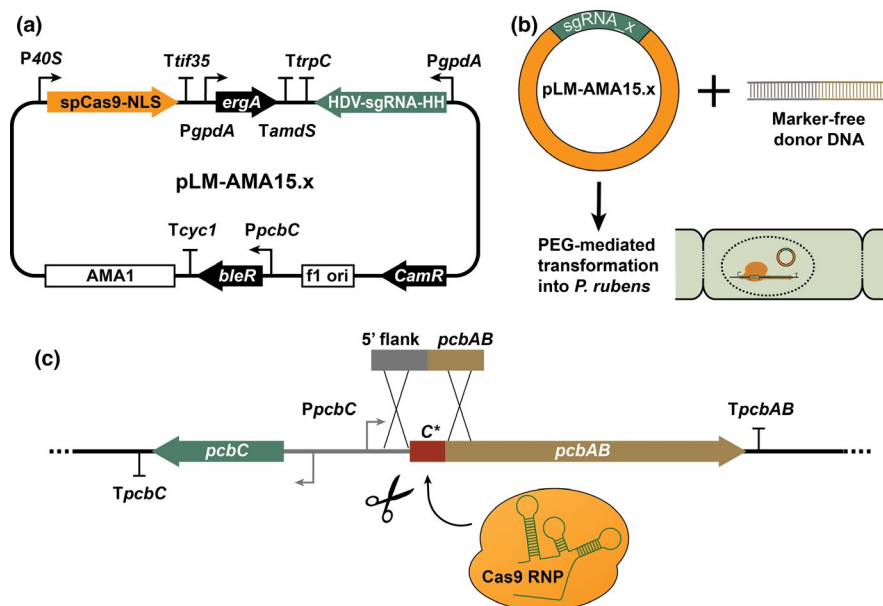


FIGURE 3 CRISPR/Cas9 plasmid-mediated deletion of the C^* domain in *Penicillium rubens* ACVS. (a) Schematic representation of pLM-AMA15.x plasmid encoding the components of CRISPR/Cas9 genome editing system, SpCas9-NLS, and the sgRNA targeting the genomic region of interest. (b) Representation of CRISPR/Cas9 plasmid and donor DNA delivery into fungal protoplasts. (c) Illustration of in vivo expressed Cas9-RNP performing marker-free genome engineering and removal of C^* domain from *pcbAB* of *P. rubens* DS54468

been extensively characterized as a purified protein in previous studies (Coque et al., 1996; Iacovelli et al., 2020). To generate the NI ACVS ΔC^* variant, we employed a Golden Gate-based assembly strategy previously described (Iacovelli et al., 2020). The construct was subsequently cloned and overexpressed in *E. coli* HM0079 (Gruenewald et al., 2004), a strain that natively expresses a phosphopantetheinyl transferase, responsible for the conversion of apo-PCP to holo-PCP domains, necessary for the function of NRPS. The NI ACVS and its ΔC^* variant, containing a 6xHis C-terminal tag, were purified via Ni²⁺-affinity chromatography and analyzed via SDS-PAGE (Figure 5a). Despite the lower expression levels, sufficient amounts of NI ACVS ΔC^* could be purified for activity assays. Further, the purified protein remained fully stable for 5 h necessary to conduct the enzyme assays, indicating that the overall organization and fold of the enzyme are not compromised.

The wild-type ACVS and ACVS ΔC^* were, respectively, diluted or concentrated to comparable concentrations based on 2D densitometry analysis on Coomassie Brilliant Blue stained SDS-PAGE (Figure 5a) and subjected to in vitro peptide formation assays. In these assays, all the amino acid substrates (L- α -amino adipic acid, L-cysteine, L-valine) are provided along with ATP and Mg²⁺, necessary for peptide biosynthesis. Two control reactions were included: one where ATP was omitted from the ΔC^* -ACVS reaction mixture and one where L- α -amino adipic acid was omitted from the wild-type ACVS reaction mixture. Reactions were analyzed via HPLC-MS for the presence of ACV and the dipeptide CV. In line with the results obtained in *P. rubens*, NI ACVS ΔC^* is completely unable to synthesize the tripeptide ACV (Figure 5b). Importantly, the enzyme synthesized significant amounts of the CV dipeptide. CV is normally a by-product or intermediate of ACV synthesis. When the substrate of module 1,

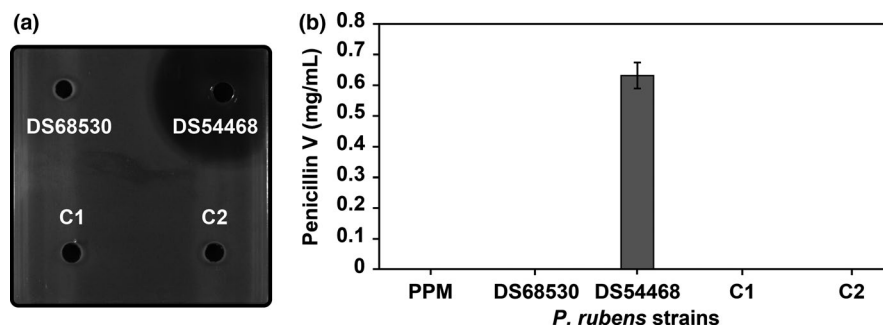


FIGURE 4 Deletion of the C^* domain in Pr ACVS. (a) Antibacterial plate assays showing that the transformants (C1 and C2) are not able to produce β -lactam antibiotics anymore. Positive control (producer strain DS54468) and negative control (non-producer strain DS68530) are included to validate the assay. (b) HPLC/MS analysis of penicillin V production, confirming the loss of penicillin V production by the transformants (C1 and C2). Producer (DS54468) and non-producer (DS68530) strains, as well as sterile PPM medium, are included

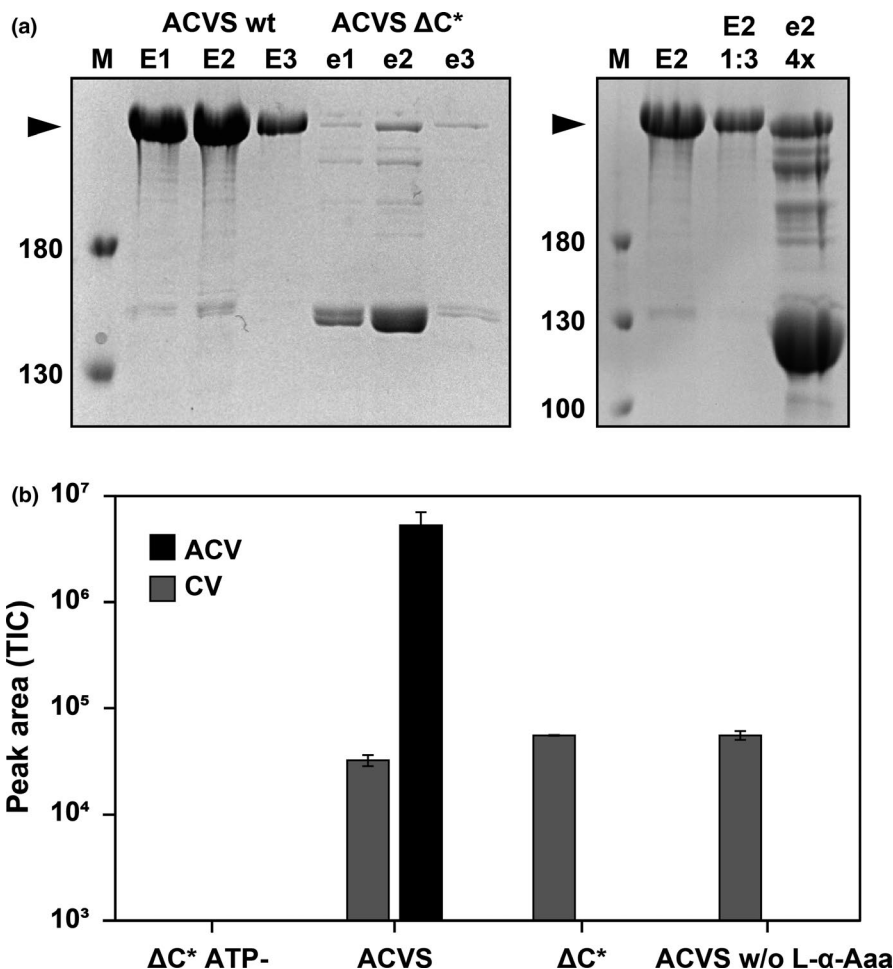


FIGURE 5 Deletion of the C* domain in *NI* ACVS. (a) SDS-PAGE analysis showing expression and purification of the wild-type and deletion variant of the *NI* ACVS. ACVS ΔC^* is expressed at lower levels but the purified protein fractions were diluted or concentrated to achieve comparable concentrations for catalytic activity measurements (top right gel). (b) In vitro assay of ACV (black bar) and CV (gray bar) production as detected by HPLC/MS analysis for ACVS (lanes 2 and 4) and its ΔC^* variant (lanes 1 and 3). Reactions were performed with the three amino acids L- α -aminoadipic acid, L-cysteine, and L-valine, except in lane 4 where L- α -Aaa was omitted

L- α -aminoadipic acid, is omitted from the reaction with the wild-type ACVS, therefore bypassing the function of module 1, we observe the production of CV at the same levels as with the ΔC^* variant. Expectedly, no ACV was produced in this case. Further, in the control reaction where ACVS ΔC^* was not supplemented with ATP neither peptides were detected. The data indicate that the deletion of the C* domain renders module 1 of *NI* ACVS unable to incorporate L- α -aminoadipic acid. However, the enzyme remains active in the production of CV, an activity that can be attributed to fully functional modules 2 and 3. These results are analogous to those obtained with ADA-less variants of NPS3, which also showed a drastic loss of adenylation activity. This suggests that C* and ADA domains carry out the same function during the activation of L- α -aminoadipic acid.

4 | CONCLUSIONS

Here, we report for the first time on the identification of a previously undiscovered conserved domain in the first module of bacterial and

fungal ACVS enzymes, which we named the C* domain. A similar domain, called ADA domain, was previously described in α -aminoadipic acid reductases from fungi, where its presence is essential for the adenylation function (Kalb et al., 2015). Following bioinformatic analysis and molecular modeling, we investigated the importance of the C* domain on the functionality of the ACVS enzymes, by engineering deletion variants and analyzing their enzymatic activity in vitro and in vivo, respectively. The results suggest that the domain arrangement of ACVS should be now regarded as C*AT₁-CAT₂-CATETe₃. Further, our experimental data indicate that the presence of the C* domain in the first module of *P. rubens* and *Nocardia lactamdurans* ACVS is critical for its functionality as a unit. Without the C* domain, the activation and subsequent incorporation of L- α -aminoadipic acid do not occur and the biosynthesis of ACV is stalled. However, the exact function of this newly identified domain remains at the moment elusive. As hypothesized for the ADA domain of NPS3, it is possible that the C* domains have a structural role, perhaps maintaining the correct fold of module 1, which is essential for its activity. Given the uniqueness of the first module of

ACVS enzymes and fungal α -amino adipic acid reductases in terms of substrate specificity—they are the only adenylating enzymes known to activate L- α -amino adipic acid, and they do so in a non-canonical manner (Coque et al., 1996; Iacovelli et al., 2020; Jensen et al., 1990; Kudo et al., 2019; Tahlan et al., 2017; Theilgaard et al., 1997)—it is also possible that the C* domain of ACVS and ADA domain of NPS3 are somehow involved in the recruitment and/or positioning of the substrate. Additionally, the presence and importance for catalysis of C* in the first module of ACVS and ADA in NPS3 (and other Lys2-type reductases) suggest that these enzymes might share a common origin, perhaps an elongation module that later evolved into an initiation module and stand-alone reductases, respectively.

Further efforts are required to unravel the evolutionary history and exact role of C* domains during the biosynthesis of ACV. New insights about this mechanistic aspect of the enzyme could prompt engineering efforts targeted at modifying the substrate specificity of the first module, to generate modified penicillin derivatives and replace the procedures for semi-synthetic penicillin with the complete fermentative production of such derivatives.

ETHICS STATEMENT

None required.

ACKNOWLEDGMENTS

We thank DSM Sinochem Pharmaceuticals (now Centriant Pharmaceuticals) for the NI ACVS construct. We would also like to thank Dr. Marten Exterkate and Niels de Kok for their help with HPLC/MS analysis. The research leading to these results was financially supported by the European Commission's Marie Curie Initial Training Network Quantfung (FP7-People-2013-ITN, grant no. 607332).

CONFLICT OF INTEREST

None declared.

AUTHOR CONTRIBUTIONS

Riccardo Iacovelli lead the formal analysis, investigation (lead), methodology, visualization, conceptualization, writing—original draft, and writing—review and editing. László Mózsik lead the investigation and methodology and supported the visualization, writing—original draft, and writing—review and editing. Roel A. L. Bovenberg supported the conceptualization, supervision, writing—original draft, and writing—review and editing. Arnold J. M. Driessen lead the conceptualization, supervision, and project administration and supported the writing—original draft and writing—review and editing.

DATA AVAILABILITY STATEMENT

All data generated or analyzed during this study are included in this published article.

ORCID

Riccardo Iacovelli  <https://orcid.org/0000-0003-4819-2251>

Arnold J.M. Driessen  <https://orcid.org/0000-0001-9258-9104>

REFERENCES

- Baldwin, J. E., & Abraham, E. (1988). The biosynthesis of penicillins and cephalosporins. *Natural Products Reports*, 5, 129. <https://doi.org/10.1039/np9880500129>
- Baldwin, J. E., Bird, J. W., Field, R. A., O'Callaghan, N. M., & Schofield, C. J. (1990). Isolation and partial characterisation of ACV synthetase from *Cephalosporium acremonium* and *Streptomyces clavuligerus*. *The Journal of Antibiotics (Tokyo)*, 43, 1055–1057.
- Baldwin, J. E., Bird, J. W., Field, R. A., O'Callaghan, N. M., Schofield, C. J., & Willis, A. C. (1991). Isolation and partial characterisation of ACV synthetase from *Cephalosporium acremonium* and *Streptomyces clavuligerus*. Evidence for the presence of phosphopantothenate in ACV synthetase. *The Journal of Antibiotics (Tokyo)*, 44, 241–248.
- Bateman, A. (2019). UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Research*, 47, D506–D515. <https://doi.org/10.1093/nar/gky1049>
- Berry, D., Mace, W., Grage, K., Wesche, F., Gore, S., Schardl, C. L., Young, C. A., Dijkwel, P. P., Leuchtman, A., Bode, H. B., & Scott, B. (2019). Efficient nonenzymatic cyclization and domain shuffling drive pyrrolopyrazine diversity from truncated variants of a fungal NRPS. *Proceedings of the National Academy of Sciences of the United States of America*, 116, 25614–25623. <https://doi.org/10.1073/pnas.1913080116>
- Bloudoff, K., Rodionov, D., & Schmeing, T. M. (2013). Crystal structures of the first condensation domain of CDA synthetase suggest conformational changes during the synthetic cycle of nonribosomal peptide synthetases. *Journal of Molecular Biology*, 425, 3137–3150. <https://doi.org/10.1016/j.jmb.2013.06.003>
- Bloudoff, K., & Schmeing, T. M. (2017). Structural and functional aspects of the nonribosomal peptide synthetase condensation domain superfamily: Discovery, dissection and diversity. *Biochimica Et Biophysica Acta - Proteins Proteomics*, 1865, 1587–1604. <https://doi.org/10.1016/j.bbapap.2017.05.010>
- Bovenberg, R. A. L., Kiel, J. A. K. W., Wenzel, T. J., & Los, A. P. (2016). EP 2 683 732 B1. European Patent Office. <https://data.epo.org/publication-server/rest/v1.0/publication-dates/20160824/patents/EP2683732NWB1/document.pdf>
- Cilia, E., Pancsa, R., Tompa, P., Lenaerts, T., & Vranken, W. F. (2013). From protein sequence to dynamics and disorder with DynaMine. *Nature Communications*, 4, <https://doi.org/10.1038/ncomms3741>
- Cilia, E., Pancsa, R., Tompa, P., Lenaerts, T., & Vranken, W. F. (2014). The DynaMine webserver: Predicting protein dynamics from sequence. *Nucleic Acids Research*, 42, 264–270. <https://doi.org/10.1093/nar/gku270>
- Concurso, H. L., & Bruner, S. D. (2012). Structure and noncanonical chemistry of nonribosomal peptide biosynthetic machinery. *Natural Products Reports*. <https://doi.org/10.1039/c2np20023f>
- Conti, E., Franks, N. P., & Brick, P. (1996). Crystal structure of firefly luciferase throws light on a super-family of adenylate-forming enzymes. *Structure*, 4, 287–298. [https://doi.org/10.1016/S0969-2126\(96\)00033-0](https://doi.org/10.1016/S0969-2126(96)00033-0)
- Conti, E., Stachelhaus, T., Marahiel, M. A., & Brick, P. (1997). Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *EMBO Journal*, 16, 4174–4183.
- Coque, J. J., de la Fuente, J. L., Liras, P., & Martin, J. F. (1996). Overexpression of the *Nocardia lactamdurans* alpha-amino adipylyl-cysteine-valine synthetase in *Streptomyces lividans*. The purified multienzyme uses cystathionine and 6-oxopiperidine 2-carboxylate as substrates for synthesis of the tripeptide. *European Journal of Biochemistry*, 242, 264–270.
- Coque, J. J., Martin, J. F., Calzada, J. G., & Liras, P. (1991). The cephamycin biosynthetic genes pcbAB, encoding a large multidomain peptide synthetase, and pcbC of *Nocardia lactamdurans* are clustered together in an organization different from the same genes in *Acremonium chrysogenum* and *Penicillium chrysogenum*. *Molecular Microbiology*, 5, 1125–1133.

- De Crecy-Lagard, V., Marliere, P., & Saurin, W. (1995). Multienzymatic non ribosomal peptide biosynthesis: identification of the functional domains catalysing peptide elongation and epimerisation. *Comptes Rendus L'Académie Des Sciences, III*, 927–936.
- De Mattos-Shiple, K. M. J., Greco, C., Heard, D. M., Hough, G., Mulholland, N. P., Vincent, J. L., Micklefield, J., Simpson, T. J., Willis, C. L., Cox, R. J., & Bailey, A. M. (2018). The cycloaspeptides: Uncovering a new model for methylated nonribosomal peptide biosynthesis. *Chemical Sciences*, 9, 4109–4117. <https://doi.org/10.1039/c8sc00717a>
- Engler, C., Gruetzner, R., Kandzia, R., & Marillonnet, S. (2009). Golden gate shuffling: A one-pot DNA shuffling method based on type IIS restriction enzymes. *PLoS One*, 4, <https://doi.org/10.1371/journal.pone.0005553>
- Flissi, A., Ricart, E., Campart, C., Chevalier, M., Dufresne, Y., Michalik, J., Jacques, P., Flahaut, C., Lisacek, F., Leclère, V., & Pupin, M. (2020). Norine: Update of the nonribosomal peptide resource. *Nucleic Acids Research*, 48, D465–D469. <https://doi.org/10.1093/nar/gkz1000>
- 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. *Nature Methods*, 6, 343–345. <https://doi.org/10.1038/nmeth.1318>
- Gruenewald, S., Mootz, H. D., Stehmeier, P., & Stachelhaus, T. (2004). In vivo production of artificial nonribosomal peptide products in the heterologous host *Escherichia coli*. *Applied and Environment Microbiology*, 70, 3282–3291. <https://doi.org/10.1128/AEM.70.6.3282-3291.2004>
- Harris, D. M., van der Krogt, Z. A., Klaassen, P., Raamsdonk, L. M., Hage, S., van den Berg, M. A., Bovenberg, R. A. L., Pronk, J. T., & Daran, J. M. (2009). Exploring and dissecting genome-wide gene expression responses of *Penicillium chrysogenum* to phenylacetic acid consumption and penicillinG production. *BMC Genomics*, 10, 1–20. <https://doi.org/10.1186/1471-2164-10-75>
- Hijarrubia, M. J., Aparicio, J. F., & Martín, J. F. (2003). Domain structure characterization of the multifunctional α -aminoacidate reductase from *Penicillium chrysogenum* by limited proteolysis: Activation of α -aminoacidate does not require the peptidyl carrier protein box or the reduction domain. *Journal of Biological Chemistry*, 278, 8250–8256. <https://doi.org/10.1074/jbc.M211235200>
- Horsman, M. E., Hari, T. P. A., & Boddy, C. N. (2016). Polyketide synthase and non-ribosomal peptide synthetase thioesterase selectivity: Logic gate or a victim of fate? *Natural Products Reports*, 33, 183–202. <https://doi.org/10.1039/C4NP00148F>
- Iacovelli, R., Zwahlen, R. D., Bovenberg, R. A. L., Arnold, J., & Id, M. D. (2020). Biochemical characterization of the *Nocardia lactamdurans* ACV synthetase. *PLoS One*, 15, 1–16. <https://doi.org/10.1371/journal.pone.0231290>
- Jensen, S. E., Wong, A., Rollins, M. J., & Westlake, D. W. (1990). Purification and partial characterization of delta-(L-alpha-aminoacidipyl)-L-cysteinyl-D-valine synthetase from *Streptomyces clavuligerus*. *Journal of Bacteriology*, 172, 7269–7271.
- Kalb, D., Lackner, G., & Hoffmeister, D. (2014). Functional and phylogenetic divergence of fungal adenylate-forming reductases. *Applied and Environment Microbiology*, 80, 6175–6183. <https://doi.org/10.1128/AEM.01767-14>
- Kalb, D., Lackner, G., Rappe, M., & Hoffmeister, D. (2015). Activity of α -aminoacidate reductase depends on the N-terminally extending domain. *ChemBioChem*, 16, 1426–1430. <https://doi.org/10.1002/cbic.201500190>
- Keating, T. A., Marshall, C. G., Walsh, C. T., & Keating, A. E. (2002). The structure of VibH represents nonribosomal peptide synthetase condensation, cyclization and epimerization domains. *Natural Structural Biology*, 9, 522–526. <https://doi.org/10.1038/nsb810>
- Koglin, A., & Walsh, C. T. (2009). Structural insights into nonribosomal peptide enzymatic assembly lines. *Natural Products Reports*, 26, 987–1000. <https://doi.org/10.1039/b904543k>
- Kohli, R. M., & Walsh, C. T. (2003). Enzymology of acyl chain macrocyclization in natural product biosynthesis. *Chemical Communications*, 3, 297–307. <https://doi.org/10.1039/b208333g>
- Kovalchuk, A., Weber, S. S., Nijland, J. G., Bovenberg, R. A. L., & Driessen, A. J. M. (2012). Fungal ABC transporter deletion and localization analysis. *Methods in Molecular Biology*, 1–16. https://doi.org/10.1007/978-1-61779-501-5_1
- Kudo, F., Miyanaga, A., & Eguchi, T. (2019). Structural basis of the non-ribosomal codes for nonproteinogenic amino acid selective adenylation enzymes in the biosynthesis of natural products. *Journal of Industrial Microbiology and Biotechnology*, 46, 515–536. <https://doi.org/10.1007/s10295-018-2084-7>
- Liras, P. (1999). Biosynthesis and molecular genetics of cephamycins. Antonie van Leeuwenhoek. *International Journal of General Molecular Microbiology*, 75, 109–124. <https://doi.org/10.1023/A:1001804925843>
- Marahiel, M. A. & Essen, L. O. (2009). Chapter 13 Nonribosomal peptide synthetases. In *Mechanistic and structural aspects of essential domains, Methods in enzymology*, 1st ed. Elsevier Inc. [https://doi.org/10.1016/S0076-6879\(09\)04813-7](https://doi.org/10.1016/S0076-6879(09)04813-7)
- Marahiel, M. A., Stachelhaus, T., & Mootz, H. D. (1997). Modular peptide synthetases involved in nonribosomal peptide synthesis. *Chemical Reviews*, 97, 2651–2674. [https://doi.org/10.1016/S0006-3495\(02\)75175-8](https://doi.org/10.1016/S0006-3495(02)75175-8)
- Marshall, C. G., Hillson, N. J., & Walsh, C. T. (2002). Catalytic mapping of the vibriobactin biosynthetic enzyme VibF. *Biochemistry*, 41, 244–250. <https://doi.org/10.1021/bi011852u>
- Martín, J. F. (1998). New aspects of genes and enzymes for β -lactam antibiotic biosynthesis. *Applied Microbiology and Biotechnology*, 50, 1–15. <https://doi.org/10.1007/s002530051249>
- Martín, J. F., Casqueiro, J., Kosalková, K., Marcos, A. T., & Gutiérrez, S. (1999). Penicillin and cephalosporin biosynthesis: Mechanism of carbon catabolite regulation of penicillin production. Antonie van Leeuwenhoek. *International Journal of General Molecular Microbiology*, 75, 21–31. <https://doi.org/10.1023/A:1001820109140>
- Marxen, S., Stark, T. D., Rüttschle, A., Lücking, G., Frenzel, E., Scherer, S., Ehling-Schulz, M., & Hofmann, T. (2015). Depsipeptide intermediates interrogate proposed biosynthesis of cereulide, the emetic toxin of *Bacillus cereus*. *Scientific Reports*, 5, 1–14. <https://doi.org/10.1038/srep10637>
- Meyer, S., Kehr, J., Mainz, A., Dehm, D., Petras, D., Su, R. D., & Dittmann, E. (2016). Biochemical dissection of the natural diversification of microcystin provides lessons for synthetic biology of NRPS article biochemical dissection of the natural diversification of microcystin provides lessons for synthetic biology of NRPS. *Cell Chemical Biology* 23, 462–471. <https://doi.org/10.1016/j.chembiol.2016.03.011>
- Mózsik, L., Büttel, Z., Bovenberg, R. A. L., Driessen, A. J. M., & Nygård, Y. (2019). Synthetic control devices for gene regulation in *Penicillium chrysogenum*. *Microbial Cell Factories*, 18, 1–13. <https://doi.org/10.1186/s12934-019-1253-3>
- Mózsik, L., Hoekzema, M., De Kok, N. A. W., Bovenberg, R. A. L., Nygård, Y., & Driessen, A. J. M. (2020). CRISPR-Based transcriptional activation tool for silent genes in filamentous fungi. *BioRxiv*. 10.13.338012
- Nijland, J. G., Ebbendorf, B., Woszczyńska, M., Boer, R., Bovenberg, R. A. L., & Driessen, A. J. M. (2010). Nonlinear biosynthetic gene cluster dose effect on penicillin production by *Penicillium chrysogenum*. *Applied and Environment Microbiology*, 76, 7109–7115. <https://doi.org/10.1128/AEM.01702-10>
- Nødvig, C. S., Nielsen, J. B., Kogle, M. E., & Mortensen, U. H. (2015). A CRISPR-Cas9 system for genetic engineering of filamentous fungi. *PLoS One*, 10, 1–18. <https://doi.org/10.1371/journal.pone.0133085>
- Ozcengiz, G., & Demain, A. L. (2013). Recent advances in the biosynthesis of penicillins, cephalosporins and clavams and its regulation. *Biotechnology Advances*, 31, 287–311. <https://doi.org/10.1016/j.biotechadv.2012.12.001>

- Petterson, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25, 1605–1612. <https://doi.org/10.1002/jcc.20084>
- Pohl, C., Kiel, J. A. K. W., Driessen, A. J. M., Bovenberg, R. A. L., & Nygård, Y. (2016). CRISPR/Cas9 based genome editing of *Penicillium chrysogenum*. *ACS Synthetic Biology*. <https://doi.org/10.1021/acssynbio.6b00082>
- Pohl, C., Polli, F., Schütze, T., Viggiano, A., Mózsik, L., Jung, S., de Vries, M., Bovenberg, R. A. L., Meyer, V., & Driessen, A. J. M. (2020). A *Penicillium rubens* platform strain for secondary metabolite production. *Scientific Reports*, 10, 1–16. <https://doi.org/10.1038/s41598-020-64893-6>
- Queener, S. W. (1990). Molecular biology of penicillin and cephalosporin biosynthesis. *Antimicrobial Agents and Chemotherapy*, 34, 943–948. <https://doi.org/10.1128/AAC.34.6.943>
- Rausch, C., Hoof, I., Weber, T., Wohlleben, W., & Huson, D. H. (2007). Phylogenetic analysis of condensation domains in NRPS sheds light on their functional evolution. *BMC Evolutionary Biology*, 15, 1–15. <https://doi.org/10.1186/1471-2148-7-78>
- Reimer, J. M., Haque, A. S., Tarry, M. J., & Schmeing, T. M. (2018). Piecing together nonribosomal peptide synthesis. *Current Opinion in Structural Biology*, 49, 104–113. <https://doi.org/10.1016/j.sbi.2018.01.011>
- Roach, P. L., Clifton, I. J., Hensgens, C. M., Shibata, N., Schofield, C. J., Hajdu, J., & Baldwin, J. E. (1997). Structure of isopenicillin N synthase complexed with substrate and the mechanism of penicillin formation. *Nature*, 387, 827–830. <https://doi.org/10.1038/42990>
- Roy, A., Kucukural, A., & Zhang, Y. (2010). I-TASSER: A unified platform for automated protein structure and function prediction. *Nature Protocols*, 5, 725–738. <https://doi.org/10.1038/nprot.2010.5>
- Salo, O. V., Ries, M., Medema, M. H., Lankhorst, P. P., Vreeken, R. J., Bovenberg, R. A. L., & Driessen, A. J. M. (2015). Genomic mutational analysis of the impact of the classical strain improvement program on β -lactam producing *Penicillium chrysogenum*. *BMC Genomics*, 16, 1–15. <https://doi.org/10.1186/s12864-015-2154-4>
- Samel, S. A., Czodrowski, P., & Essen, L. O. (2014). Structure of the epimerization domain of tyrocidine synthetase A. *Acta Crystallographica, Section D: Biological Crystallography*, 70, 1442–1452. <https://doi.org/10.1107/S1399004714004398>
- Schneider, A. & Marahiel, M. A. (1998). Genetic evidence for a role of thioesterase domains, integrated in or associated with peptide synthetases, in non-ribosomal peptide biosynthesis in *Bacillus subtilis*. *Archives of Microbiology*, 169, 404–410. <https://doi.org/10.1007/s002030050590>
- Schneider, T. L., Shen, B., & Walsh, C. T. (2003). Oxidase domains in epothilone and bleomycin biosynthesis: Thiazoline to thiazole oxidation during chain elongation. *Biochemistry*, 42, 9722–9730. <https://doi.org/10.1021/bi034792w>
- Schoenafinger, G., Schracke, N., Linne, U., & Marahiel, M. A. (2006). Formylation domain: An essential modifying enzyme for the non-ribosomal biosynthesis of linear gramicidin. *Journal of the American Chemical Society*, 128, 7406–7407.
- Schofield, C. J., Baldwin, J. E., Byford, M. F., Clifton, I., Hajdu, J., Hensgens, C., & Roach, P. (1997). Proteins of the penicillin biosynthesis pathway. *Current Opinion in Structural Biology*, 7, 857–864. [https://doi.org/10.1016/S0959-440X\(97\)80158-3](https://doi.org/10.1016/S0959-440X(97)80158-3)
- Schwarzer, D., Finking, R., & Marahiel, M. A. (2003). Nonribosomal peptides: From genes to products. *Natural Products Reports*, 20, 275–287. <https://doi.org/10.1039/b111145k>
- Schwecke, T., Aharonowitz, Y., Palissa, H., von Dohren, H., Kleinkauf, H., & van Liempt, H. (1992). Enzymatic characterisation of the multifunctional enzyme delta-(L-alpha-aminoadipyl)-L-cysteiny-D-valine synthetase from *Streptomyces clavuligerus*. *European Journal of Biochemistry*, 205, 687–694.
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., & Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 7, <https://doi.org/10.1038/msb.2011.75>
- Stachelhaus, T., Mootz, D. H., & Marahiel, M. A. (1999). The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chemistry & Biology*, 6, 493–505. [https://doi.org/10.1016/S1074-5521\(99\)80082-9](https://doi.org/10.1016/S1074-5521(99)80082-9)
- Strieker, M., Tanović, A., & Marahiel, M. A. (2010). Nonribosomal peptide synthetases: Structures and dynamics. *Current Opinion in Structural Biology*, 20, 234–240. <https://doi.org/10.1016/j.sbi.2010.01.009>
- Tahlan, K., Moore, M. A., & Jensen, S. E. (2016). δ -(L- α -aminoadipyl)-L-cysteiny-D-valine synthetase (ACVS): Discovery and perspectives. *Journal of Industrial Microbiology and Biotechnology*. <https://doi.org/10.1007/s10295-016-1850-7>
- Tahlan, K., Moore, M. A., & Jensen, S. E. (2017). δ -(L- α -aminoadipyl)-L-cysteiny-D-valine synthetase (ACVS): Discovery and perspectives. *Journal of Industrial Microbiology and Biotechnology*, 44, 517–524. <https://doi.org/10.1007/s10295-016-1850-7>
- Tanovic, A., Samel, S. A., Essen, L.-O., & Marahiel, M. A. (2008). Crystal structure of the termination module of a nonribosomal peptide synthetase. *Science*, 321, 659–663. <https://doi.org/10.1126/science.1159850>
- Theilgaard, H. B., Kristiansen, K. N., Henriksen, C. M., & Nielsen, J. (1997). Purification and characterization of delta-(L-alpha-aminoadipyl)-L-cysteiny-D-valine synthetase from *Penicillium chrysogenum*. *The Biochemical Journal*, 327 (Pt 1), 185–191.
- van Liempt, H., von Dohren, H., & Kleinkauf, H. (1989). delta-(L-alpha-aminoadipyl)-L-cysteiny-D-valine synthetase from *Aspergillus nidulans*. The first enzyme in penicillin biosynthesis is a multifunctional peptide synthetase. *Journal of Biological Chemistry*, 264, 3680–3684.
- Walsh, C. T. (2016). Insights into the chemical logic and enzymatic machinery of NRPS assembly lines. *Natural Products Reports*, 33, 127–135. <https://doi.org/10.1039/C5NP00035A>
- Weber, E., Engler, C., Gruetzner, R., Werner, S., & Marillonnet, S. (2011). A modular cloning system for standardized assembly of multigene constructs. *PLoS One*, 6, <https://doi.org/10.1371/journal.pone.0016765>
- Weber, S. S., Polli, F., Boer, R., Bovenberg, R. A. L., & Driessen, A. J. M. (2012). Increased penicillin production in *Penicillium chrysogenum* production strains via balanced overexpression of isopenicillin n acyltransferase. *Applied and Environment Microbiology*, 78, 7107–7113. <https://doi.org/10.1128/AEM.01529-12>
- Wu, X., García-Estrada, C., Vaca, I., & Martín, J. F. (2012). Motifs in the C-terminal region of the *Penicillium chrysogenum* ACV synthetase are essential for valine epimerization and processivity of tripeptide formation. *Biochimie*, 94, 354–364. <https://doi.org/10.1016/j.biochi.2011.08.002>
- Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., & Zhang, Y. (2014). The I-TASSER suite: Protein structure and function prediction. *Nature Methods*, 12, 7–8. <https://doi.org/10.1038/nmeth.3213>
- Yang, J., & Zhang, Y. (2015). I-TASSER server: New development for protein structure and function predictions. *Nucleic Acids Research*, 43, W174–W181. <https://doi.org/10.1093/nar/gkv342>

How to cite this article: Iacovelli R, Mózsik L, Bovenberg RA, Driessen AJ. Identification of a conserved N-terminal domain in the first module of ACV synthetases. *MicrobiologyOpen*. 2021;10:e1145. <https://doi.org/10.1002/mbo3.1145>

APPENDIX

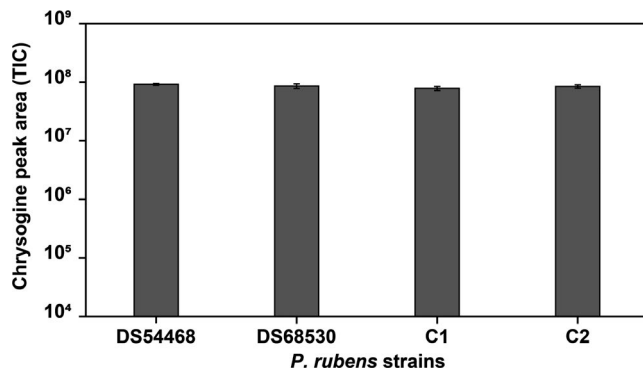


Figure A1 Production of chrysogine was detected and measured via HPLC/MS in the cultures' supernatant after 5 days of growth, to serve as a control

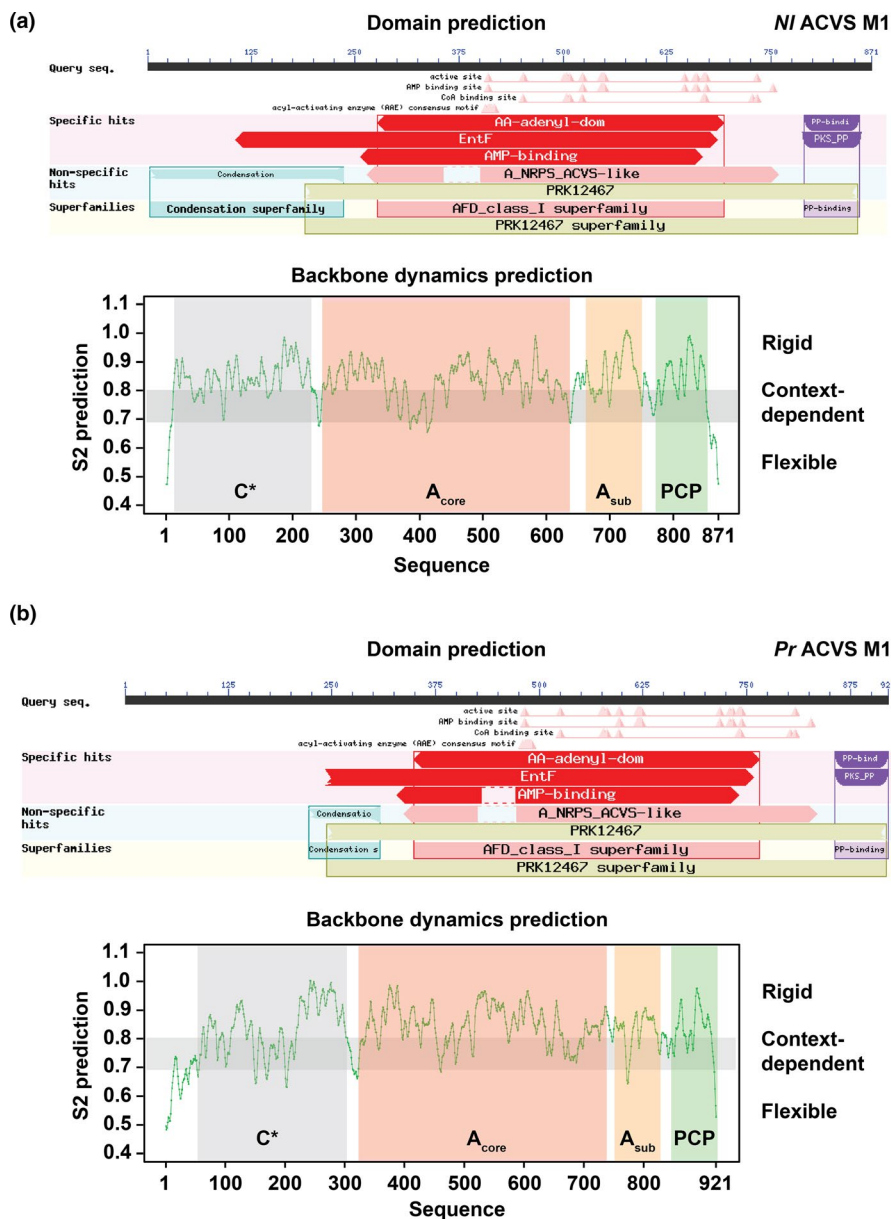


Figure A2 Domain and backbone dynamics prediction of NI ACVS M1 (A) and Pr ACVS M1 (B). The upper panels represent the graphic summary of the BlastP domain prediction, where all the canonical domains are identified, as well as the N-terminal regions that align (partly) to the condensation domain family. The lower panels depict the backbone dynamics prediction obtained with DynaMine. In addition to the three canonical domains of module 1, predicted folded regions can be identified also in the N-terminal stretch of the two sequences, termed C* domains

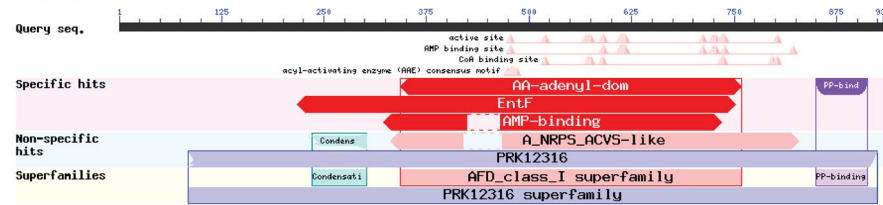
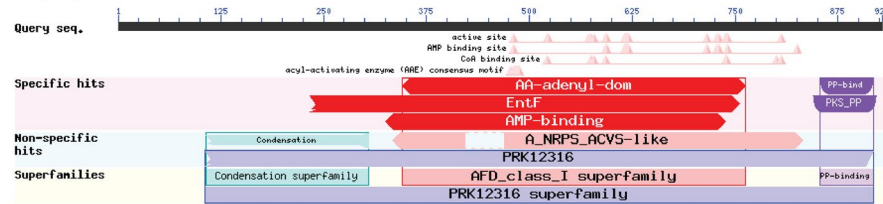
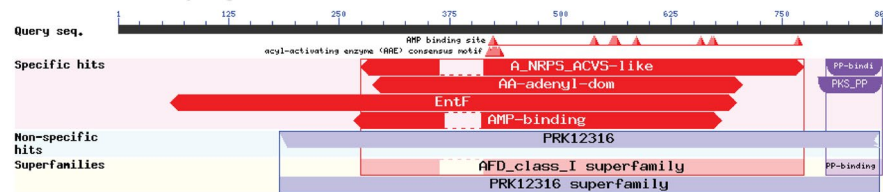
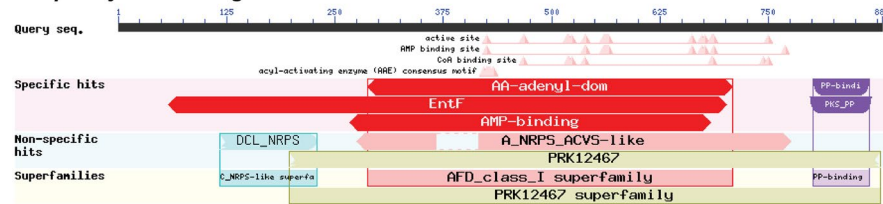
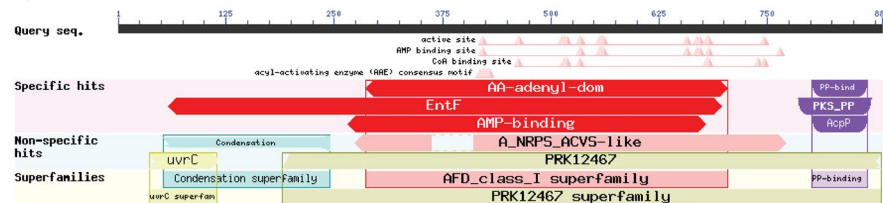
Aspergillus flavus* ACVS M1**Aspergillus nidulans* ACVS M1*****Acremonium chrysogenum* ACVS M1*****Streptomyces clavuligerus* ACVS M1*****Lysobacter lactamdurans* ACVS M1**

Figure A3 BlastP domain predictions of ACV synthetases from different organisms show the omnipresence of the conserved N-terminal region. The ACVS from *Acremonium chrysogenum* does not show alignment to the condensation domain family but retains an equally sized region

TABLE A1 List of oligonucleotides used for PCR amplification of sgRNA insert and marker-free donor DNA for deletion of C* in *Pr ACVS* and for amplification of *NI ACVS* Δ C* deletion fragment

Description	Template	Primer Sequences (5' → 3')
sgRNA insert C* <i>PrACVS</i>	Overlapping oligonucleotides	FW: ATGGTCTCACCGACACTACCTGATGAGTCCGTGAGGACGAAACGAG RV: ATGGTCTCTAAACATAGGCTTCTCGGCCACTACGACGAGCTTACTCGTTTCGTCCTCACGGACTCA
Marker-free donor DNA Δ C* <i>PrACVS</i>	Overlapping oligonucleotides	FW: CCCAGTATAAGGAATTCCCCTCGAGCTTGTCTGTGATTGCGTTTTTCTAACACTTGTGTTG CATCCGATCCGTCCCTACCAATTATTGGTCATTGACAGACATGGAAGAGTGGAACAA RV: CGCAGACAACGGCTATTTTGTCTTCATGCCGTTCAACCACCTTCAATGAGATGGTGCAGTCGCT TTGATGAAGGGTACTCGCCATCCGTGTTGTCCACTTCCATGTCTGTCAATG
Deletion fragment Δ C* <i>NIACVS</i>	pBAD- <i>NI ACVS</i> intermediate vector	FW: ATCGTGTACAATCGATTACCTAGGAGGAATTAACCATGCACCAGTGGAAACGGCACCGAC RV: ATCGAAGCTTGCAATGGCGCATGTC

TABLE A2 Sequences of C* domains deleted from *Nocardia lactamdurans* and *Penicillium rubens ACVS*

Enzyme	Amino acid positions	Sequence
<i>NI ACVS</i>	2-236	TSARHLKSAADWCARIDAIAGQRCDLEM LLKDEWRHRVAVRSDTAVRATQEK ELTISGQDYTALKQALGAMPEAFA LATLHSV LHAYGHGHQTVVAFLRDG KVLPPVVVDHLEQAGLTCAEAAEQLE DAVAREDMYLPPEELLQRGLFDALL VLADGHLGFTLPPAPLV TIVRDDP AAGCLHWRIAYAGEFFEDKIIAGV LDVAREVLGQFIGRPEQLVADIDL VSAEQELQL
<i>Pr ACVS</i>	2-308	TQLKPPNGTTPIGFSATSLNASGSSSVKN GTIKPSNGIFKPSTRDTMDPCSGNAAD GSIRVFRGGIERWKECVNQVPERCDL SGLTTDSTRYQLASTGFGDASAAYQE RLMTVPVDVHAALQELCLERRVSVGSV INFSVHQMLKGFNGTHTITASLHREQ NLQNSSPSWVVSPTIVTHENRDGWSV AQAVESIEAGRGSEKESVTAIDSGSSLV KMGLFDLLVSFVDADDARIPCFDFPLAV IVRECDANLSLTLRFSDCLFNEETICNF TDALNILLAEAVIGRVTPVADIELLSAE QKQQL