


ORIGINAL ARTICLE

A novel way to synthesize pantothenate in bacteria involves β -alanine synthase present in uracil degradation pathway

Mariana López-Sámano¹ | Luis Fernando Lozano-Aguirre Beltrán¹ |
 Rosina Sánchez-Thomas² | Araceli Dávalos¹ | Tomás Villaseñor³ |
 Jorge Donato García-García² | Alejandro García-de los Santos¹ 

¹Programa de Ingeniería Genómica, Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, México

²Departamento de Bioquímica, Instituto Nacional de Cardiología "Ignacio Chávez", Tlalpan, México

³Departamento de Medicina Molecular y Bioprocesos, Instituto de Biotecnología, UNAM, Cuernavaca, México

Correspondence

Alejandro García-de los Santos, Centro de Ciencias Genómicas, UNAM. Av. Universidad s/n Col. Chamilpa 62210, Cuernavaca, Morelos, México.
 Email: alex@ccg.unam.mx

Abstract

Pantothenate is an indispensable vitamin precursor of the synthesis of coenzyme A (CoA), a key metabolite required in over 100 metabolic reactions. β -Alanine (β -ala) is an indispensable component of pantothenate. Due to the metabolic relevance of this pathway, we assumed that orthologous genes for β -alanine synthesis would be present in the genomes of bacteria, archaea, and eukaryotes. However, comparative genomic studies revealed that orthologous gene replacement and loss of synteny occur at high frequency in *panD* genes. We have previously reported the atypical plasmid-encoded location of the pantothenate pathway genes *panC* and *panB* (two copies) in *R. etli* CFN42. This study also revealed the unexpected absence of a *panD* gene encoding the aspartate decarboxylase enzyme (ADC), required for the synthesis of β -ala. The aim of this study was to identify the source of β -alanine in *Rhizobium etli* CFN42. In this study, we present a bioinformatic analysis and an experimental validation demonstrating that the source of β -ala in this *R. etli* comes from β -alanine synthase, the last enzyme of the uracil degradation pathway.

KEYWORDS

β -alanine, pantothenate, CoA, comparative genomics, uracil degradation, pantothenate, vitamin

1 | INTRODUCTION

β -Alanine is a nonproteinogenic β -amino acid that occurs in all living organisms. In prokaryotes, β -ala is indispensable for the synthesis of pantothenate, the precursor of the essential cofactor coenzyme A (CoA). CoA is the source of 4'-phosphopantetheine for fatty acid and polyketide synthesis (Leonardi & Jackowski, 2007). In eukaryotes, β -amino acids and β -peptides play important roles in the regulation

of nutritional metabolism, immunity, and the central nervous system (Naveed Riaz, Rehman M, & Mahboob Ahmad, 2017).

The major pathway for β -ala synthesis in *Escherichia coli* is the decarboxylation of aspartate by aspartate decarboxylase (ADC; Cronan, 1980). The ADC protein is a pyruvoyl-dependent enzyme that is initially synthesized as a zymogen (pro-ADC). A cleavage of pro-ADC occurs between Gly24 and Ser25, creating the active-site pyruvoyl moiety. Stuecker (Stuecker, Bramhacharya,

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

© 2020 The Authors. *MicrobiologyOpen* published by John Wiley & Sons Ltd.

Hodge-Hanson, Suen, & Escalante-Semerena, 2015) proposed two classes of ADC based on the type of cleavage of the zymogen (pro-ADC). Class I of the ADC cleavage requires the MRF (Maturation Regulatory Factor) acetyl-CoA sensor and has been found only in gammaproteobacteria. ADC Class II is an autocatalytic cleavage and is found in a wide number of bacterial phyla. Since the majority of archaea lack homologues of the *E. coli* K12 acetyl-CoA synthesis pathway genes, the mechanism of pantothenate/CoA biosynthesis has not been completely deduced in these organisms.

The pantothenate synthesis pathway, which includes a glutamate decarboxylase (GAD) that substitutes for ADC and uses pyridoxal 5'-phosphate (PLP) as a cofactor, was reported in archaea (Tomita, Yokooji, Ishibashi, Imanaka, & Atomia, 2014). Curiously, GAD prefers aspartate (Asp) rather than glutamate (Glu), as its substrate, although commonly GAD catalyzes the decarboxylation of Glu to γ -aminobutyrate (GABA).

Although prokaryotes and eukaryotes have an indispensable requirement for β -ala for the synthesis of coenzyme A (CoA), the pathways involved in its synthesis are very diverse. The uracil fermenting bacterium *Clostridium uracilicum* degrades uracil to β -ala. Uracil or thymine is first converted to dihydrouracil. The dihydropyrimidinase enzyme catalyzes the hydration of dihydrouracil to produce *N*-carbamoyl- β -ala, which is hydrolyzed to β -ala, CO₂, and NH₃, by β -ala synthase (Campbell, 1957).

The reductive degradation of pyrimidine as a source of β -ala was supported by genetic and biochemical analyses in several bacteria, including *Clostridium uracilicum* (Campbell, 1957) and *Clostridium botulinum* (Hilton, Mead, & Elsdén, 1975). Although the reductive degradation of pyrimidines has also been implicated as the de novo source of β -ala in *E. coli* auxotrophs, lack of response to dihydrouracil indicated that in these bacteria, the major pathway for β -ala synthesis was the decarboxylation of aspartate catalyzed by ADC.

Genschel (Genschel, 2004) performed a phyletic analysis for the occurrence of *E. coli* and human genes for in pantothenate and CoA synthesis across 47 completely sequenced genomes, 20 from the Bacteria, 16 from Archaea, and 11 from Eukarya. This study revealed a mosaic of orthologues with 20 to 70% amino acid identities. At least one protein was missing from each of the 47 analyzed genomes.

Comparative genomics using the *E. coli* pantothenate pathway genes as query against the 20 sequenced bacterial genomes revealed multiple gaps that may represent distantly related homologues due to the absence of, at least, one gene per surveyed bacterial genome (Genschel, 2004).

The *Rhizobiales* order is a heterogeneous group of Gram-negative bacteria, taxonomically located within the alphaproteobacteria division. Some of its members are facultative diazotrophs that associate with leguminous plants to carry out symbiotic nitrogen fixation. Others are pathogens of plants or animals (Martínez-Romero & Caballero-Mellado, 1996). Our model is *Rhizobium etli* CFN42, which was originally isolated from bean root nodules (Martínez-Romero, 2003). Its genome consists of a circular chromosome and six large plasmids ranging in size from 194 to 642 Kb (Gonzalez et al., 2006).

In the course of examining *Rhizobium etli* CFN42 plasmids for the presence of housekeeping genes encoding essential functions, we found that both *panC* and *panB* genes were clustered together on the 642-kb replicon p42f. We demonstrated that both are indispensable for the synthesis of pantothenate (Villaseñor et al., 2011; Figure A1). Surprisingly, we did not find homologues of the *E. coli panD* gene in the genome of *R. etli* CFN42. Since strain CFN42 grows in minimal medium without exogenous pantothenate or β -ala, it was assumed that it is a pantothenate prototroph.

Agrobacterium fabrum C58 (formerly *A. tumefaciens* C58), a plant pathogen that induces tumors in numerous plants, was the only member of the *Rhizobiales* order included in Genschel's study. According to this analysis, *A. fabrum* C58 lacks ketopantoate reductase (KPR, EC 1.1.1.169) but has a putative ADC detected by a BlastP search. We performed BlastP searches in order to gain insight on the presence or absence of ADC in the genomes of rhizobial reference strains.

Several questions arise from the presence or absence of ADC in *R. etli* CFN42. Is the absence of ADC an exclusive characteristic of strain CFN42 or is it a widespread characteristic of the *Rhizobiales* order or perhaps the alphaproteobacteria?

The aim of this work was to identify the enzyme that synthesizes β -ala and replaces the function of ADC, allowing *R. etli* CFN42 to be a β -ala prototroph. We also performed an in silico analysis of the alphaproteobacteria group to understand the occurrence, diversity, and evolution of the enzymes involved in β -ala synthesis.

2 | MATERIAL AND METHODS

2.1 | Bacterial strains, media, and growth conditions

The characteristics of the bacterial strains and plasmids used in this study are listed in Table 1. Bacterial growth was started from glycerol stocks (20%, stored at -70°C) propagated in plates of PY-rich medium (per L, 5 g peptone, 3 g yeast extract, 1 ml of CaCl₂ and 15 g agar). *Rhizobium* strains were grown at 30°C in three different media: (a) PY-rich medium, (b) chemically defined mineral medium (MM), and (c) chemically defined mineral medium plus 1 μM calcium pantothenate (MMP) or 1 μM β -ala, added from filter sterilized stocks. Base MM containing 10 mM succinate as carbon source, 10 mM NH₄Cl as nitrogen source, 1.26 mM K₂HPO₄, and 0.83 mM MgSO₄ was adjusted to pH 6.8 and sterilized. After sterilization, the following components were added to the final concentration indicated: 1.49 mM CaCl₂·2H₂O (autoclaved separately), 0.0184 mM FeCl₃·6H₂O, 10 $\mu\text{g}/\text{ml}$ biotin, and 10 $\mu\text{g}/\text{ml}$ thiamine (all filter sterilized). MMP contains the same components plus 1 μM calcium pantothenate. *Rhizobium* strains were grown at 30°C for 20 hr in PY medium. *Escherichia coli* K12 MG1655 and *E. coli* BL21 (DE3) were used for cloning and to express the *R. etli* β -alanine synthase, respectively. *E. coli* strains were grown at 37°C for 20 hr in Luria-Bertani (LB) medium (1% tryptone, 0.5% yeast extract, 0.5% NaCl, pH 7.2).

TABLE 1 Bacterial strains and plasmids used in this study

	Relevant Genotype	References
<i>Rhizobium etli</i> strains		
CFN42	Wild type, Nalr	Segovia, Young and Martinez Romero 1993
CFNX186	CFN42 cured of plasmid p42f; Nalr	Brom et al., 1992
CFN42	CFN42 pfΔ308–637	Brom et al., 1992
CFN42 RHE_CH02599-	CFN42 RHE_CH02599::pK18mob Kmr	This study
CFN42 amaB-	CFN42 amaB::pK18mob Kmr	This study
CFN42 amaB-/amaB <i>R. etli</i>	CFN42 amaB::pK18mob/complemented with amaB into pFAJ1708 Tc	This study
CFN42 amaB-/aam <i>R. etli</i>	CFN42 amaB::pK18mob/complemented with aam into pSRK Gm	This study
CFN42 amaB-/amaB <i>A. fab</i>	CFN42 amaB::pK18mob/complemented with amaB/ <i>A. fab</i> into pFAJ1708 Tcr	This study
CFN42 amaB-/panD <i>A. fab</i>	CFN42 amaB::pK18mob/complemented with panD/ <i>A. fab</i> into pFAJ1708 Tcr	This study
CFN42 amaB-/bioA <i>A. fab</i>	CFN42 amaB::pK18mob/complemented with bioA/ <i>A. fab</i> into pFAJ1708 Tcr	This study
CFN42 amaB-/panD <i>E. coli</i>	CFN42 amaB::pK18mob/complemented with panD/ <i>E. coli</i> into pFAJ1708 Tcr	This study
CFN42 amaB-	CFN42 amaB::pK18mob/complemented with pFAJ1708 Tc	This study
<i>Rhizobium tropici</i>		
CIAT 899	CIAT 899 aam::pK18mob Kmr	This study
CIAT 899	CIAT 899 gabt::pK18mob Kmr	This study
<i>Agrobacterium fabrum</i> C58		
fabrum C58	panD::pK18mob Kmr	This study
fabrum C58	bioA::pK18mob Kmr	This study
fabrum C58	amaB::pK18mob Kmr	This study
<i>Escherichia coli</i> and plasmid		
K-12 substr. MG1655 ΔpanD	MG1655 ΔpanD::Kan	This study
K-12 substr. MG1655 ΔpanD/panD <i>A. fab</i>	MG1655 ΔpanD::Kan/panD <i>A. fab</i> into pUC19 Cbr	This study

(Continues)

TABLE 1 (Continued)

	Relevant Genotype	References
K-12 substr. MG1655 ΔpanD/amaB <i>A. fab</i>	MG1655 ΔpanD::Kan/amaB <i>A. fab</i> into pUC19 Cbr	This study
K-12 substr. MG1655 ΔpanD/bioA <i>A. fab</i>	MG1655 ΔpanD::Kan/bioA <i>A. fab</i> into pUC19 Cbr	This study
K-12 substr. MG1655 ΔpanD/amaB <i>R. etli</i>	MG1655 ΔpanD::Kan/amaB <i>R. etli</i> into pUC19 Cbr	This study
K-12 substr. MG1655 ΔpanD/aam <i>R. etli</i>	MG1655 ΔpanD::Kan/aam <i>R. etli</i> into pUC19 Cbr	This study
K-12 substr. MG1655 ΔpanD/RHE_CH02599 <i>R. etli</i>	MG1655 ΔpanD::Kan/RHE_CH02599 <i>R. etli</i> into pUC19 Cbr	This study
K-12 substr. MG1655 ΔpanD	MG1655 ΔpanD::Kan/complemented with pUC19 Cbr	This study
DH5α	Host for recombinant plasmids; Nalr	
pK18mob	pK18, derivative mob; Kmr	Schäfer et al. 1994
pUC19	Cloning vector Cbr	
pSRK	pBBRMC-5-derived expression vector lac promoter, lacIq, lacZ α+, Gmr	Khan, Gaines, Roop and Farrand, 2008
pFAJ1708	Broad Host range cloning vector, Tcr	
pETSUMO	Protein and Peptide Expression System; His Tag (6x), SUMO Tag; Kmr	Hanington, Barreda and Belosevic, 2006
<i>Escherichia coli</i> BL21(DE3)	Host for recombinant plasmids;	Thermo fisher Scientific
BL21(DE3) AmaB	pETSUMO with AmaB recombinant protein	
BL21(DE3) pETSUMO	pETSUMO empty vector	

2.2 | DNA manipulations

Standard techniques were used for plasmid and total DNA isolation, restriction digests, ligations, transformations, and agarose gel electrophoresis (Sambrook, Fritsch, & Maniatis, 1989). Plasmid mobilization from *E. coli* to *Rhizobium* was done by conjugation on PY plates at 30°C by using overnight stationary phase cultures. Donors

(*E. coli* DH5) and recipients (*R. etli* CFN42 wild-type and mutant strains) were mixed at a 1:2 ratio, and suitable markers were used for transconjugant selection.

2.3 | Analysis for the occurrence of 12 proteins involved in pantothenate synthesis and phylogenetic analysis of putative ADC enzymes found in alphaproteobacteria

We selected 204 alphaproteobacteria to analyze for the presence and absence of 12 proteins related to the pantothenate synthesis and transport. The protein FASTA files (*faa*) for each of the genomes were downloaded from the RefSeq NCBI database. Protein sequences with an expectation value (*E*) of 10^{-3} or less were considered as putative homologues. We used Proteinortho v5.15 to obtain the clusters of orthologous proteins from the 204 protein FASTA files. Next, we used the Pfam v31.0 database to determine which protein ortho clusters represent the 12 proteins of interest analyzed in this work. The proteins we searched for were PYD1, PYD2, PYD3, GAD, KPHMT, PS, ADC, KPR, MRF, KAR, Aam, and GabT. Finally, we determined which alphaproteobacteria were represented in each protein cluster (Table S1).

For phylogenetic analysis, we used the Pfam v31.0 database to determine which proteinortho clusters represent the ADC proteins. A total of 37 homologues belonging to alphaproteobacteria sequences were tested with a group of nine external sequences listed in Table S2 and were aligned against Muscle v3.8.31.

The resulting data set containing 46 putative ADC homologues was used to infer the evolutionary relationships. We used ProtTest3 v3.4.2 for the evolutionary model, and the best result was LG + G model, using amino acid alignment. The phylogenetic analysis was performed with PhyML v3.3.20170530 under (-d aa -m LG -a e -o ltr) parameters (Figure A2).

2.4 | Cloning and sequence analysis of *amaB* gene, mutants, and complemented mutant

The *amaB* (RHE_CH03290) gene was overexpressed in *E. coli* DH5-alpha. The coding region of the *amaB* gene was amplified from the genomic DNA of *R. etli* CFN42 by PCR. The amplified fragment was inserted into the pET-SUMO expression vector (Ni-NTA Purification System; Sigma-Aldrich). After confirming the absence of mutations, the plasmid was introduced into *E. coli* strain BL21 (DE3). Primer set list is in Table A1.

2.5 | Overexpression and purification of wild-type β -ureidopropionase AmaB

The transformant BL21(DE3) strain was grown in LB medium supplemented with 100 mg/ml of carbenicillin. A single colony was

transferred into 10 ml of LB medium with carbenicillin at the above-mentioned concentration in a 100-ml flask. This culture was incubated overnight at 37°C with shaking. Five hundred milliliters of LB medium with 100 mg/ml of carbenicillin was inoculated with 5 ml of the overnight culture in a 1-liter flask. After 3 hr of incubation at 37°C with vigorous shaking, the optical density at 600 nm (OD₆₀₀) of the culture was 0.3–0.5. For induction of β -alanine synthase gene expression, isopropyl- β -D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.1 mM and incubation was continued at 30°C for an additional 6 hr.

The cells were collected by centrifugation (8,000 \times g, 4°C, 10 min), washed twice in wash buffer (2.5 M NaCl, 250 mM NaH₂PO₄, 20 mM imidazole pH 8.0), and resuspended in 50 ml of the same buffer. The cells were disrupted by sonication using a UP200S ultrasonic processor, in ice for four periods of 15 s at pulse mode 0.5 and 40% sonic power. The cell debris was pelleted by centrifugation (8,000 \times g, 4°C, 10 min), and the supernatant was applied to a 2-ml column of nickel metal-affinity resin (Ni-NTA Purification System; Sigma-Aldrich) and β -ureidopropionase purified as recommended by the manufacturer. The purified enzyme was dialyzed against 20 mM sodium phosphate buffer, pH 8.0, and stored at 4°C.

2.6 | Enzyme assays

The standard enzymatic reaction was carried out with purified AmaB at a final concentration of 1 mg/ml along with 125 mM 3-ureidopropionic acid and 10 mM MgCl₂ dissolved in 100 mM sodium phosphate buffer, pH 8.0, in a 3 ml reaction volume (Martinez-Gomez et al., 2008). The reaction mixture was incubated at 30°C for 60 min, with the apoenzyme preincubated (1 hr) at 4°C with 2 mM of NiCl₂, and 500 μ l samples were stopped for every 15 min, by the addition of 50 μ l of 3% TCA. After centrifugation, the presence of β -ala in the resulting supernatants was estimated by high-performance liquid chromatography (HPLC).

2.7 | Determination of β -ala by HPLC/fluorescence

Determination of β -ala was carried out by HPLC coupled to a Multi γ -fluorescence detector (Waters 1525/2475) using a reverse-phase C-18 Spherisorb ODS2 column of 5 μ m particle size and 150 \times 4.6 mm (Waters; García-García, Peña-Sanabria, Sánchez-Thomas, & Moreno-Sánchez, 2018). Enzymatic reactions were stopped with perchloric acid (3% v/v) at the indicated times and immediately frozen in liquid nitrogen and kept at -70°C. The acidic supernatants were neutralized with 3 M KOH/0.1 M Tris and centrifuged to remove KClO₄. Supernatant was recovered and used for β -ala determination by derivatization with 37 mM ortho-phthalaldehyde (OPA). β -ala and α -alanine standards (Sigma-Aldrich, Saint Louis, MO, USA) were used for identifying of chromatographic peaks.

3 | RESULTS

3.1 | Orthologues of the canonical L-aspartate- α -decarboxylase enzyme are predominantly absent in α -proteobacteria

A previous study on ADC phylogeny and amino acid conservation analyses revealed that ADCs are present in γ -proteobacterial genomes and most maintain the *panCBD* synteny (Stuecker et al., 2015). We noticed the absence of the *panCBD* gene cluster while functionally characterizing *panC* and *panB* in rhizobia (Villaseñor et al., 2011). In the present study, BlastP and Psi Blast searches using ADC from *E. coli* and *A. fabrum* C58 as query revealed the absence of ADC homologues in *R. etli* CFN42 and other reference strains (Table S1).

To generalize the absence of ADC homologues in α -proteobacteria, we assessed the occurrence of putative ADCs in the proteome of 204 alphaproteobacteria, 84 rhizobia and 120 members of seven families of alphaproteobacteria (Table S1). The complete proteome of each bacterium was obtained from the NCBI reference sequence collection (RefSeq) and clustered with Proteinortho v5.15, a large-scale Blast-based orthology detection tool (Lechner et al., 2011; Figure A2). This analysis only showed 37 putative ADCs from 204 α -proteobacteria genomes.

3.2 | Unrooted maximum-likelihood-based tree inferred from the alpha- and gammaproteobacteria ADCs revealed high divergence among them

An important characteristic of the alphaproteobacteria is its genome plasticity, which allows different genome rearrangements, including deletions or duplications (Prell & Poole, 2006; Tiwari & Lata, 2018). We made a phylogenetic analysis to get a wider view of the evolutionary relationship among the ADCs from the α -, β -, γ -, and ϵ -proteobacteria (Table S2).

The resulting maximum-likelihood-based tree is shown in Figure 1, and the data set is presented in Table S2. To determine if this ADC phylogeny maintains the coherence of species phylogeny, it was compared to the previously reported species trees performed by the Bayesian analysis of 104 concatenated alignments (Williams, Sobral, & Dickerman, 2007) and with the most recent robust species tree; this was done under the maximum-likelihood framework with a data set of 200 single-copy and conserved genes for the alphaproteobacteria (Muñoz-Gómez et al., 2019).

The majority of ADCs belonging to the *Rhizobiales* order were grouped in a single cluster (Figure 1, green oval bottom). Unexpectedly, we found two ADCs from *Bradyrhizobium japonicum* and *Afipia* sp close to γ - and β -proteobacteria (Figure 1, upper green oval). Two ADCs of the *Rhodospirillales* order (*Azospirillum brasilense* and *Magnetospirillum magneticum*) were located as orthologues of *Rhizobiales* (Figure 1, purple oval). The ADC from *Maritimibacter alkaliphilus* was located within the *Rhizobiales* order, whereas in the species tree, *M. alkaliphilus* belongs to the *Rhodobacterales* order (Muñoz-Gómez et al., 2019).

This heterogeneous cluster of *Rhizobiales* ADCs links through a long branch with remote orthologues from class alphaproteobacteria belonging to the following orders: *Rhizobiales* (*B. japonicum* and *Afipia*), *Caulobacterales* (*C. crescentus*), *Rhodobacterales* (*Hyphomonas neptuniou* and *Hirschia baltica*), and *Parvularculales* (*Parvularcula bermudensis* γ -proteobacteria (outgroup, *E. coli*, *Salmonella*, *Pseudomonas aeruginosa*), β -proteobacteria (*B. pertussis* and *R. solanacearum*), and ϵ -proteobacteria (*Corynebacterium glutamicum*)).

3.3 | Presence, absence, duplications, and functional redundancy of the six *pan* genes involved in pantothenate synthesis

In addition to ADC, the 84 rhizobial genomes were surveyed for the presence of the enzymes that catalyze pantothenate synthesis. This revealed that the KPHMT (ketopantoate hydroxymethyl transferase, PanB) is highly conserved in the *Rhizobiales* order and was absent in only 8.4% of the analyzed genomes. The genera lacking KPHMT were *Bradyrhizobium* sp ORS 278, *Candidatus liberobacter* (4 strains), and *Hoeflea* (2 species). KPHMT was predicted to be present in the other members of *Rhizobiales*, which have a diversity of habitats. Two copies of this enzyme were present in 17.8% of rhizobia, mostly in the *Rhizobium* and *Sinorhizobium* species. Three copies of KPHMT were found in 4.8% of the *Rhizobiales* order: three *Mesorhizobium* species and one in *Rhizobium leucaenae* (Table S1).

The next step in the pathway is the reduction of α -ketopantoate to produce pantoate. Two enzymes can perform this reduction: KPR (α -ketopantoate reductase, PanE) was found in only 57% of the rhizobial genomes, while KAR (acetohydroxy acid reductoisomerase, *ilvC*) was present in 95.2% of the genomes (Table S1). Most human, plant, and mammalian pathogens have lost the KPR enzyme. Interestingly, *Candidatus* genera lacked both KAR and KPR enzymes in their genomes.

In the last step of the pathway, pantothenate synthetase (PS, PanC) catalyzes the ATP-dependent condensation of D-pantoate with β -ala to form pantothenate. This enzyme was absent in 7.1% of the surveyed genomes, some of which belong to parasites such as *Hoeflea* and *Candidatus* (Table S1). Strains with a single copy were found in 88% of the analyzed rhizobia genomes. Two genes were found in 4.7% of *Mesorhizobium* and *Bradyrhizobium* species.

The occurrence of putative ADC enzymes is shown in Table 2. An ADC encoding gene was present in 19% of the genomes, and two strains had a second copy of ADC in their genomes. *Mesorhizobium japonicum* MAFF303099 had one in a chromosome and the other one in a plasmid, and *M. loti* NZP 2037 had both in a chromosome.

The search for MRF (Maturation Regulatory Factor, *panM*) homologues revealed that 78.5% of the genomes lacked an MRF homologue; 15.4% had one copy and 6% encoded two copies. However, only five genomes coded for both ADC and MRF (*A. fabrum* C58, *Bradyrhizobium japonicum* USDA11, *R. etli* bv *mimosa* str. Mim1, *Rhizobium gallicum*, and *Sinorhizobium fredii* HH130; Table S1).

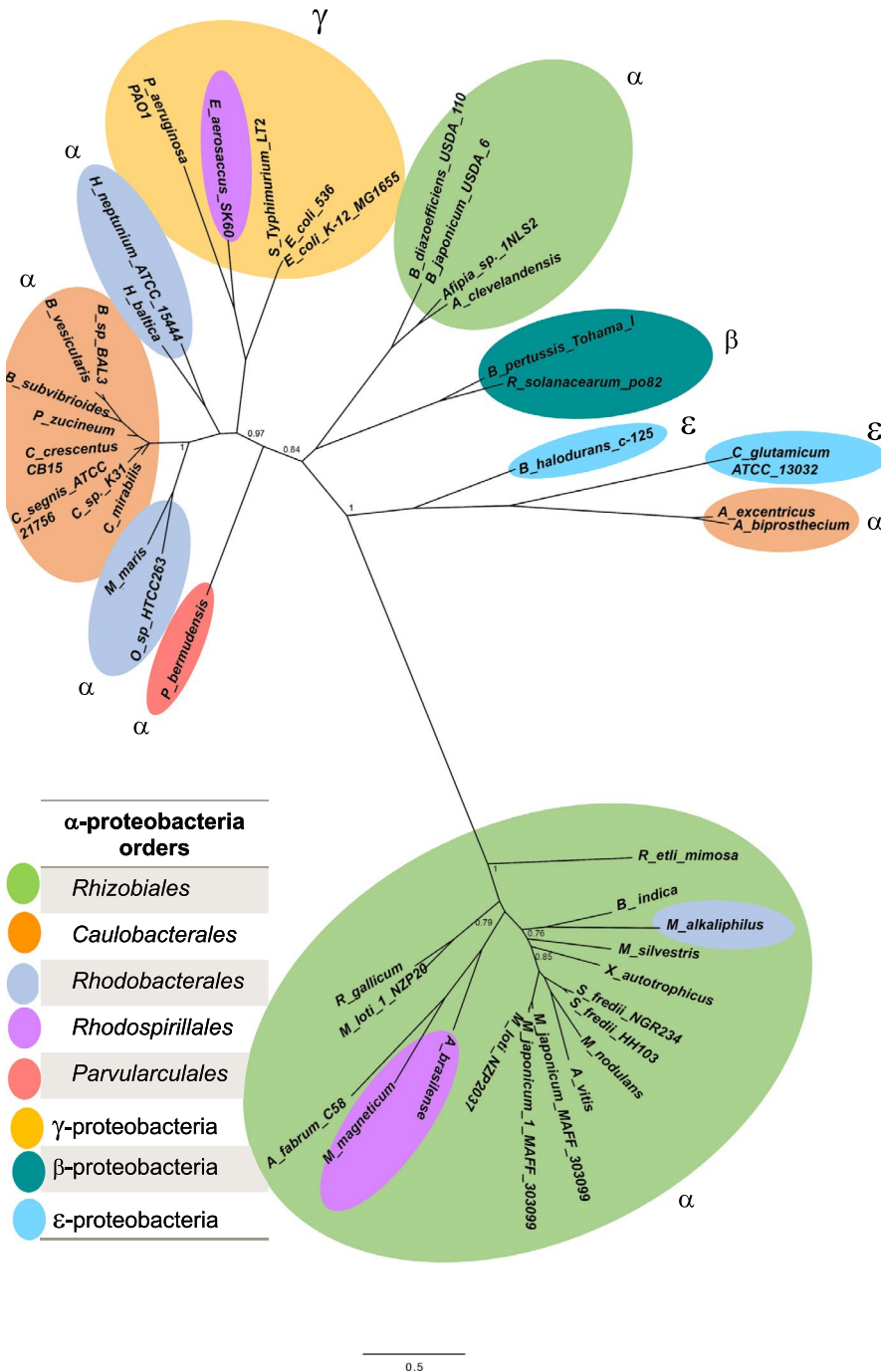


FIGURE 1 The putative ADCs of alphaproteobacteria found in our occurrence analysis. A maximum-likelihood phylogenetic tree inferred from a subset of 204 genomes, where we extracted only 37 ADCs. The tree shows a monophyletic clade of proteins distantly related to those from γ -proteobacteria and other α -proteobacteria

Our results showed that only 16 of the 84 analyzed genomes (19.04%) encoded the complete pantothenate pathway. Of the 68 genomes with gaps in the pathway, the predominant deficiencies were a lack of ADC in 80.95% of the genomes and the absence of both ADC and KPR in 38%.

3.4 | *Rhizobium etli* CFN42 is a pantothenate prototroph

The model of pantothenate synthesis established in *E. coli* (Cronan, 1980; Leonardi & Jackowski, 2007) indicates that the enzymes

missing in rhizobia should cause auxotrophy. Growth assays were done in liquid chemically defined medium with *R. etli* CFN42 (lacks *panD*) and *Sinorhizobium meliloti* 1,021 (lacks *panD* and *panE*) wild-type strains, and an *R. etli* CFN42 plasmid p42f-cured strain (CFNX186) that is defective for growth in chemically defined medium without pantothenate. We found that the wild-type strains were able to grow through three subcultures in minimum medium without β -ala or pantothenate. This shows that even with the absence of *panE* and/or *panD*, rhizobia are still able to synthesize β -ala and pantothenate (Figure 2). This prototrophy contrasts with the auxotrophy exhibited by *R. etli* CFNX186, which lacks *panC* and *panB*, as well as plasmid p42f (Brom et al., 1992).

TABLE 2 Occurrence* of pantothenate synthesis genes on Rhizobiales order

Gene	Enzyme	Occurrence (%)
<i>panD</i>	ADC	19.04
<i>panM</i>	MRF	21.42
<i>panB</i>	KPHMT	91.66
<i>panE</i>	KPR	57.14
<i>ilvC</i>	KAR	95.23
<i>panC</i>	PS	92.85

Note: *The percentage was calculated based on the number of rhizobia bacteria that covered the sample ($n = 84$).

3.5 | Occurrence analysis revealed different pathways that would replace ADC in rhizobia

To identify which enzyme(s) might be responsible for the synthesis of β -ala, we performed bioinformatic analyses of 204 alphaproteobacterial genomes to find possible pathways or genes that could potentially produce this metabolite. Based on a literature search, we selected six genes of interest that encode enzymes of the pyrimidine degradation pathway (AmaB, Dht, PyrD), glutamate decarboxylase (GAD), and the Aam and GabT transaminases (Figure 3).

It is believed that β -ala synthesis in bacteria results only through decarboxylation of aspartate by ADC (Cronan, Littel, & Jackowski, 1982; David & Lichstein, 1950). Other ways of producing β -ala exist in eukaryotes. Two routes occur in fungi: *Saccharomyces cerevisiae* produces β -ala by the degradation of spermine (White, Gunyuzlu, & Toyn, 2001), and *Schizosaccharomyces pombe* and *Saccharomyces kluyveri* obtain it from uracil degradation (Lundgren, Gojković, Piškur, & Dobritzsch, 2003; Table S1, Figure 3).

The pyrimidine degradation pathway involves three enzymatic steps from uracil to produce β -ala, CO_2 , and NH_3 (Campbell, 1957). In the final step of the pathway, β -ala synthase (AmaB) uses *N*-carbamoyl- β -alanine as substrate. In rhizobia, the in vitro activity of AmaB has been detected in *A. fabrum* C58 and *S. meliloti* 1021. The authors showed the production of β -ala from 3-ureidopropionic acid in vitro, in the last step of the pathway (Martínez-Rodríguez, Martínez-Gómez,

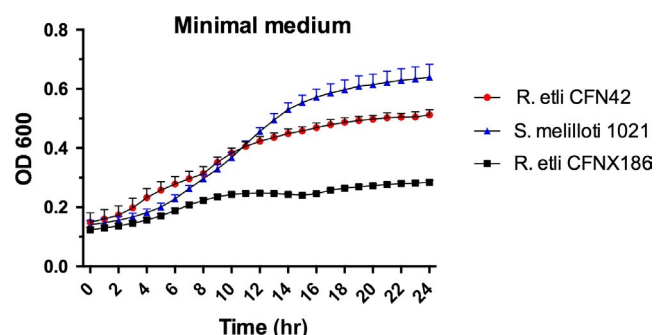


FIGURE 2 Growth test for prototrophy of wild-type *Rhizobium etli* CFN42, its p42f-cured derivative CFNX186, and in wild-type *Sinorhizobium meliloti* 1,021. Tested in minimal medium without β -alanine or pantothenate

Rodríguez-Vico, Clemente-Jiménez, & Las Heras-Vázquez, 2010). In archaea, β -ala can be synthesized by a GAD that uses Asp as a substrate. In these studies, it was shown that two enzymes annotated as GADs had higher affinity for Asp than for Glu and they demonstrated the in vitro activity of the enzymes in *Methanocaldococcus jannaschii* and *Thermococcus kodakarensis* (Tomita et al., 2014; Wang, Xu, & White, 2014).

We also included in the study two transaminases that in bacteria, insects, and mammals produce β -ala in a single-step reaction. The first one, Aam, acts on L-alanine and 3-oxopropanoate to produce pyruvate and β -ala (Dalluge, Liao, Gokarn, & Jessen, 2005; Yun, Lim, Cho, & Kim, 2004). The second, GabT, performs a transamination of malonate semi-aldehyde and L-glutamate (Nanaya, Hidenori, Keiko, tatsuhiko, Ikeda, & Takao, 1982; Wilding, Peat, Newman, & Scott, 2016).

In summary, our bioinformatic analysis showed that two transaminases and the pyrimidine degradation pathway were encoded in the *R. etli* CFN42 genome. We did not find any candidate genes for ADC or GAD, nor a complete polyamine degradation pathway.

3.6 | AmaB functionally complements strains lacking ADC

In our study, we tested the function of different genes in *R. etli*, by inactivating those that encode two transaminases (Aam and GabT) and the *amaB* gene for pyrimidine degradation (Table A1). Following with the canonical decarboxylation pathway, we found a putative ω -amino acid decarboxylase that was different from the ADC and GAD enzymes. The genes were interrupted using a suicide plasmid, and the resulting mutants were tested for growth in defined medium without β -ala or pantothenate. From this screening, we found that the *amaB* mutant was auxotrophic for β -ala, while inactivation of the other genes caused no growth deficiency (data not shown).

amaB (RHE_CH03290) is a chromosomal gene annotated as β -alanine synthase. It belongs to the pyrimidine degradation pathway and transforms 3-ureidopropionic acid to β -ala, CO_2 , and ammonia. We disrupted this gene in *R. etli* CFN42 and grew the ReAM-1 (*amaB*⁻) mutant in mineral medium (MM) without β -ala or pantothenic acid. The mutant was deficient in growth, indicating a β -ala auxotrophy, and its growth was restored by exogenous β -ala or by introducing the *amaB* gene in a plasmid (Figure 4).

Similarly, the mutant was complemented with a plasmid-borne copy of the *amaB* gene from *A. fabrum* C58. The product of this gene has been shown to have β -alanine synthase activity in vitro (Martínez-Gómez et al., 2008).

3.7 | Purified AmaB produces β -ala from 3-ureidopropionic acid in vitro

The his₅-tag enzyme was purified in an immobilized nickel affinity column under native conditions and had a molecular mass of 60 kDa,

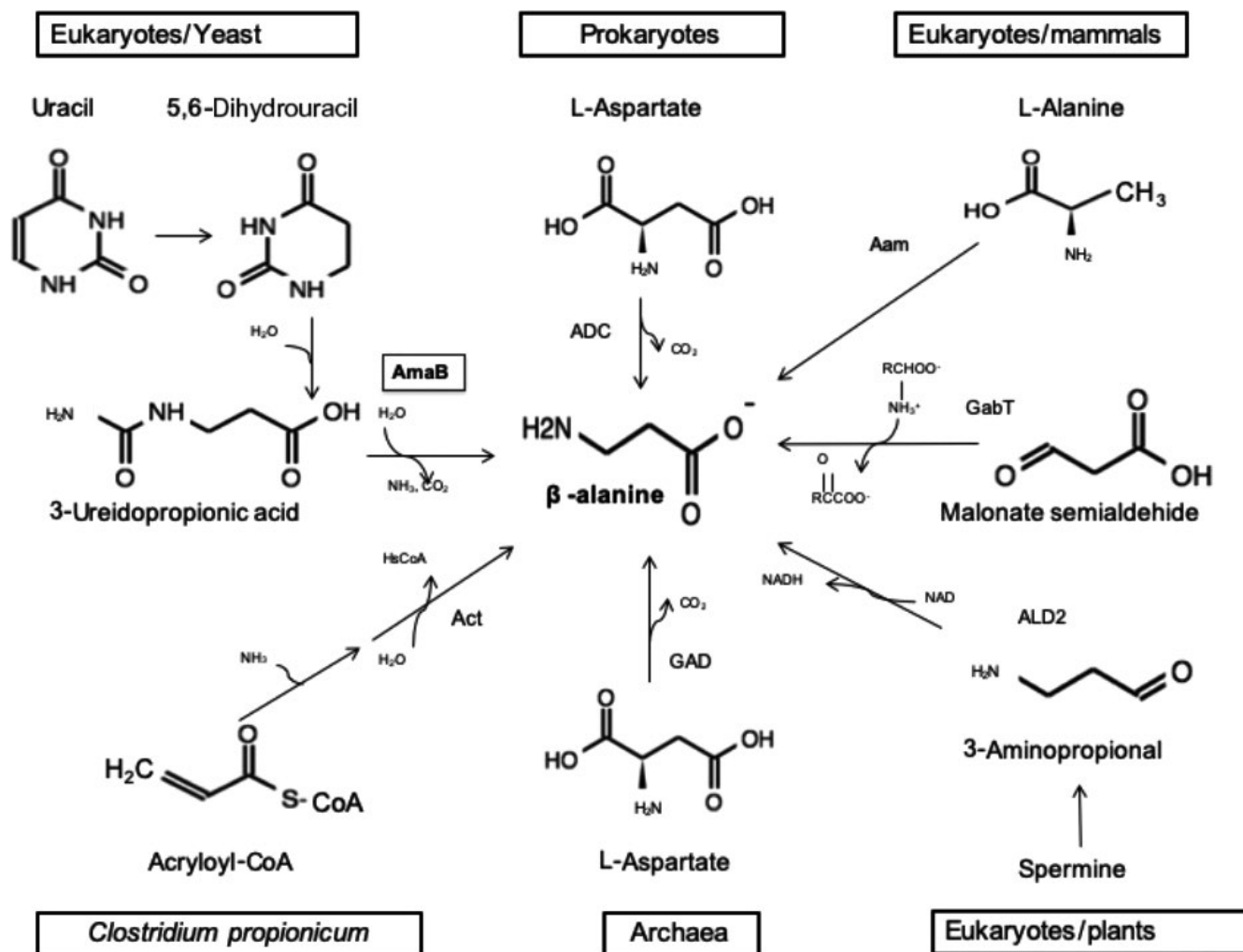


FIGURE 3 β-Alanine biosynthesis in different domains of life. (AmaB) β-alanine synthase; (ADC) 1-aspartate decarboxylase; (Aam) 2,3-aminomutase; (gabT) 4-aminobutyrate transferase; (ALD2) amine oxidase; (GAD) glutamate decarboxylase; (Act) β-alanyl-CoA:ammonia lyase

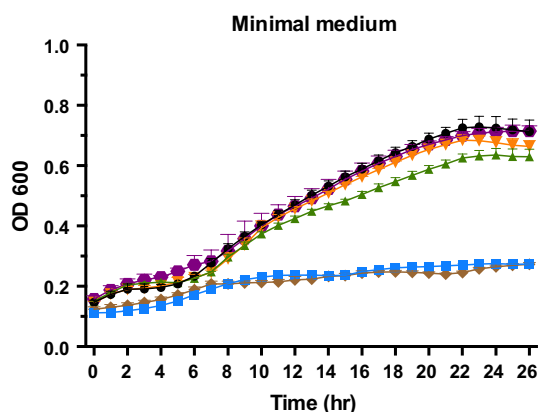


FIGURE 4 Growth in minimal medium of the *Rhizobium etli* CFN42 wild-type (●) strain and its derivative complements. *R. etli* CFN42 *amaB* mutant (■); CFN42 *amaB/amaB*⁺ of *R. etli* CFN42 (▲); CFN42 *amaB/amaB*⁺ of *A. fabrum* C58 (▼); CFN42 *amaB* complemented with 1 μM of β-alanine (●); CFN42 *amaB/pFAJ1708* empty vector (◆)

consistent with β-ala synthase (45 kDa), plus the 15-kDa 6His-Sumo tag (Figure A3).

Enzymes of this type are characterized as metalloenzymes that use Ni²⁺ and Co²⁺ as cofactors in enzyme assays. The reaction mixture contained purified AmaB preincubated with Ni²⁺ or Co²⁺, 10 mM MgCl₂, 100 mM sodium phosphate buffer, and 3-ureidopropionic acid as a substrate. We initially used a TLC system with ninhydrin detection to identify the presence of β-ala (Niederwieser et al., 1971; Figure A4). We observed enzymatic activity with both metal ions, and no product was formed in their absence.

As described below, we also performed our enzymatic assays using an HPLC system to obtain a better resolution.

3.8 | Synthesis of β-ala by recombinant AmaB

The *R. etli* CFN42 *amaB* gene was heterologously expressed in *E. coli* strain BL21 (DE3) and recovered by Ni²⁺ affinity chromatography, as

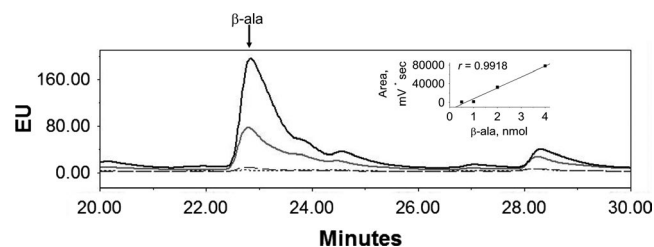


FIGURE 5 Standard of β -ala was analyzed by HPLC/fluorescence at 0.5 (dotted line), 1 (dashed line), 2 (gray line), and 4 (black line) nmols following the protocol detailed in Material and Methods. Inset shows the linear analysis of areas from each peak

previously described (Martinez-Gomez et al., 2008). Production of β -ala by recombinant AmaB was analyzed by HPLC. The fluorescence response of β -ala had a linear relation with concentration (Figure 5). The time course of recombinant AmaB activity using 3-ureidopropionic acid as substrate and Ni^{2+} as cofactor showed that β -ala is synthesized at a linear rate for up to 30 min (Figure 6a). β -ala was not detected in a reaction assay without recombinant AmaB protein (Figure 6b). The standard of β -ala overlapped with the peak of the compound

synthesized by AmaB, while the α -ala standard did not (Figure 6c). These results indicated that recombinant AmaB is able to synthesize β -ala.

4 | DISCUSSION

The relevance of β -ala as a key component of pantothenate synthesis has been well established. However, the diversity of mechanisms described in bacteria and eukaryotes suggests that the synthesis of β -ala has not been totally elucidated.

Pioneer studies performed in *E. coli* and γ -proteobacteria defined that β -ala was synthesized by the decarboxylation of L-aspartate in a one-step reaction catalyzed by ADC. The concept of a canonical one-step decarboxylation reaction was for many years assumed to be the sole source of β -ala in bacteria.

The genomic era facilitates the comparison of pathways among numerous species (Genschel, 2004); this bioinformatic approach helped us determine the diversity of mechanisms involved in β -ala synthesis. In this study, we found several differences

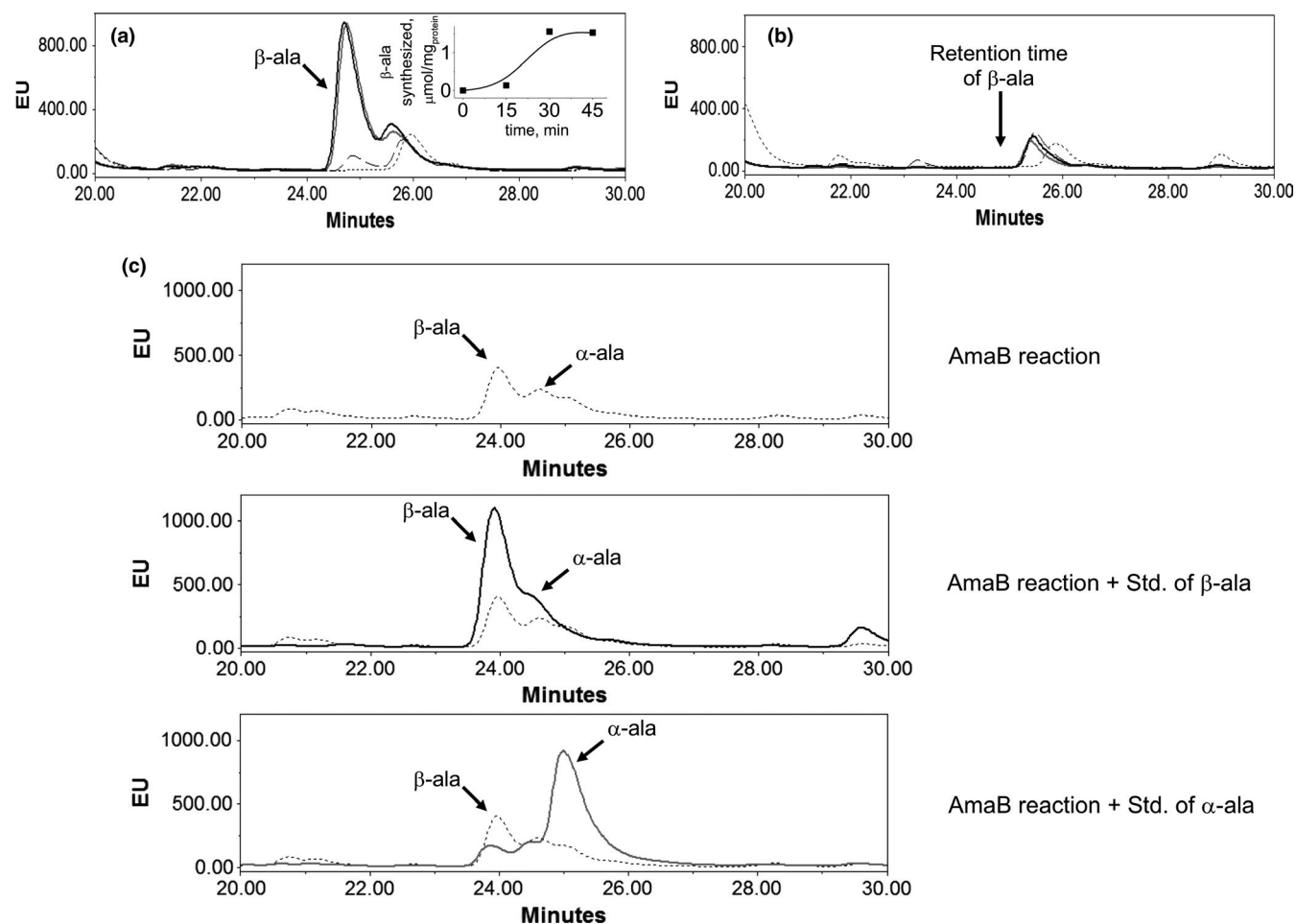


FIGURE 6 Analysis β -alanine synthesized by AmaB. (a) Activity of recombinant AmaB, at 0 (dotted line), 15 (dashed line), 30 (gray line), and 45 min (black line). Peak of β -ala was observed at 24.8 min. (b) Reaction without enzyme; β -ala peak was not observed. (c) To confirm peak of β -ala, internal standard of β -ala (black line) and L-ala (gray line) was added independently to neutralized AmaB reaction carried out for 15 min (dotted line)

between *R. etli* (alphaproteobacteria) and *E. coli* (γ -proteobacteria); the most intriguing was the absence of an ADC homologue in rhizobia. In previous studies, analyses of *E. coli* and other γ -proteobacteria revealed that β -ala was produced by the decarboxylation of aspartate by aspartate decarboxylase enzyme (ADC; Cronan et al., 1982); in several archaea, β -ala was synthesized by a glutamate decarboxylase (GAD) able to decarboxylate both aspartate and glutamate (Tomita et al., 2014). These data confirm the relevance of one-step decarboxylases, not only in bacteria but also in archaea.

An unusual alternative source of β -ala synthesis is the reductive degradation of pyrimidine. This three-step reaction was found in *Clostridium uracilicum* (Campbell, 1957) and *C. botulinum* (Hilton et al., 1975), as well as in *E. coli* strains: *E. coli* W, *E. coli* D2, *E. coli* 99-1, and *E. coli* 99-2 (Table S1); in contrast to previous studies, none of them was able to grow in the presence of dihydrouracil and β -ureidopropionic acid (Slotnick & Weinfeld, 1956).

In bacteria belonging to the *Rhizobiales* order, little is known about the metabolism of β -ala and pantothenate (Villaseñor et al., 2011). The occurrence analysis performed in this work indicates that our model, *R. etli* CFN42, lacks ADC and GAD, the most common one-step reaction used in bacteria to synthesize β -ala. We suggest that there can be functional redundancy in certain rhizobia strains. As part of our work, we constructed different single and double mutants in *A. fabrum* C58 to try to get an auxotrophic strain, but in all cases, the mutants continue to be β -ala prototrophic (data not shown).

Particularly for the *Rhizobiales* order, we constructed a heat map with their most representative genomes; here, we can associate the loss and prevalence of different pathways, assuming that the decarboxylation pathway is missing in most of rhizobia genomes (Figure A5).

Sinorhizobium meliloti and *A. fabrum* C58 have been tested for production of β -amino acids through the uracil degradation pathway because of their pharmaceutical relevance (Martínez-Rodríguez et al., 2010). Unexpectedly, the research only showed the ability to produce β and ω amino acids in vitro; we do not know if these strains synthesize β -ala through this pathway or if these strains have a functional redundancy with another β -ala synthesis pathway.

As part of our occurrence analysis, we extended our work to alphaproteobacteria with 120 more genomes from seven different orders (Table S1). We found a correlation between the rhizobia order and alphaproteobacteria. In general, we observed that the pyrimidine degradation pathway (37%) and Aam transaminase (56%) are widely distributed in alphaproteobacteria, as well as in rhizobia (Table S1). We also observed that ADC and GAD enzymes are poorly represented in alphaproteobacteria, with 17% and 6.8%, respectively. This analysis suggests a strong correlation between the loss of the decarboxylation pathway and predominance of the pyrimidine degradation pathway in the *Rhizobiales* order and in alphaproteobacteria.

We also tested the activity of recombinant AmaB in vitro, by HPLC, to confirm the catalytic activity of this protein by showing

that it produces β -ala from 3-ureidopropionic acid; this corroborates the presence of alternative pathway in which bacteria produce this essential metabolite.

5 | CONCLUSIONS

Prokaryotes and eukaryotes require β -ala to synthesize CoA; however, the source of β -ala is quite variable even in bacteria.

For years, it has been assumed that the main source of β -ala in prokaryotes comes from the decarboxylation of aspartate in a single enzymatic step catalyzed by ADC. This reaction was discovered in *E. coli* and has been assumed to be the main source of β -ala in γ -proteobacteria.

This study in *R. etli* CFN42 and other alphaproteobacteria revealed a remarkable reduction of ADC orthologs in these bacteria.

The bioinformatics and experimental analyses performed with rhizobia indicate that in these alphaproteobacteria β -ala is synthesized through the reductive pyrimidine degradation pathway.

All these data highlight the metabolic plasticity for β -ala and pantothenate in bacteria.

ACKNOWLEDGMENTS

Mariana López-Sámano is a doctoral student from Programa de Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México (UNAM), and received fellowship 295956 from CONACYT. We are grateful to Susana Brom and Michael F. Dunn for their helpful discussion and for critically reviewing the manuscript. We gratefully acknowledge Laura Cervantes and Victor Antonio Becerra Rivera for their skillful technical assistance. This research was supported by the annual institutional budget that UNAM shares with scientists.

CONFLICT OF INTEREST

None declared.

AUTHOR CONTRIBUTIONS

Mariana López-Sámano; Formal analysis-Equal, investigation-Equal, methodology-Equal, Writing-original draft-Equal. Luis Fernando Lozano-Aguirre Beltàn; Data curation-Equal, software-supporting, supervision supporting, validation supporting. Rosina Sánchez-Thomas; Formal analysis -supporting, methodology-supporting, validation-supporting. Araceli Dávalos; Investigation-supporting, methodology-supporting, supervision-supporting. Tomás Villaseñor; Investigation-supporting, supervision-supporting. Jorge Donato García-García; Formal analysis -supporting, methodology-supporting, validation-supporting. Alejandro García- de los Santos; Conceptualization-Equal, formal analysis-Equal, funding acquisition-supporting, investigation -Equal, supervision-Equal, writing original draft-Equal.


ETHICS STATEMENT

None required.

DATA AVAILABILITY STATEMENT

All 204 data used in this study were download from RefSeq public NCBI database. The ID for each of the amino acid sequences is available through tools at different websites as described in Material and Methods, or upon request from the corresponding author. The supplementary material in the text referred to as Tables S1 and S2 was uploaded in a Zenodo data repository with DOI numbers <https://doi.org/10.5281/zenodo.3593441> and <https://doi.org/10.5281/zenodo.3593474>, respectively.

ORCID

Alejandro García-de los Santos  <https://orcid.org/0000-0002-3387-9280>

REFERENCES

- Brom, S., Garcia-De-los-Santos, A., Stepkowsky, T., Flores, M., Davila, G., Romero, D., & Palacios, R. (1992). Different plasmids of *Rhizobium leguminosarum* bv. phaseoli are required for optimal symbiotic performance. *Journal of Bacteriology*, *174*(16), 5183–5189. <https://doi.org/10.1128/jb.174.16.5183-5189.1992>
- Campbell, L. L. (1957). Reductive degradation of pyrimidines. III. Purification and properties of dihydrouracil dehydrogenase. *The Journal of Biological Chemistry*, *227*(2), 693–700.
- Cronan, J. E. Jr (1980). Beta-alanine synthesis in *Escherichia coli*. *Journal of Bacteriology*, *141*(3), 1291–1297. Retrieved from <http://jb.asm.org/content/141/3/1291.short>.
- Cronan, J. E., Littel, K. J., & Jackowski, S. (1982). Genetic and biochemical analysis of pantothenate biosynthesis in *Escherichia coli* and *Salmonella typhimurium*. *Journal of Bacteriology*, *149*(3), 916–922.
- Dalluge, J. J., Liao, H., Gokarn, R., & Jessen, H. (2005). Discovery of enzymatic activity using stable isotope metabolite labeling and liquid chromatography-mass spectrometry. *Analytical Chemistry*, *77*(20), 6737–6740. <https://doi.org/10.1021/ac051109y>
- David, W. E., & Lichstein, H. C. (1950). Aspartic acid decarboxylase in bacteria. *Experimental Biology and Medicine*, *73*(2), 216–218. <https://doi.org/10.3181/00379727-73-17632>
- García-García, J. D., Peña-Sanabria, K. A., Sánchez-Thomas, R., & Moreno-Sánchez, R. (2018). Nickel accumulation by the green algae-like *Euglena gracilis*. *Journal of Hazardous Materials*, *343*, 10–18. <https://doi.org/10.1016/j.jhazmat.2017.09.008>
- Genschel, U. (2004). Coenzyme A biosynthesis: Reconstruction of the pathway in archaea and an evolutionary scenario based on comparative genomics. *Molecular Biology and Evolution*, *21*(7), 1242–1251. <https://doi.org/10.1093/molbev/msh119>
- Gonzalez, V., Janga, S. C., Moreno-Hagelsieb, G., Jimenez-Jacinto, V., Bustos, P., Santamaria, R. I., ... Ramirez, M. A. (2006). The partitioned *Rhizobium etli* genome: Genetic and metabolic redundancy in seven interacting replicons. *Proceedings of the National Academy of Sciences*, *103*(10), 3834–3839. <https://doi.org/10.1073/pnas.0508502103>
- Hanington, P. C., Barreda, D. R., & Belosevic, M. (2006). A novel hematopoietic granulin induces proliferation of goldfish (*Carassius auratus* L.) macrophages. *Journal of Biological Chemistry*, *281*(15), 9963–9970.
- Hilton, M. G., Mead, G. C., & Elsdén, S. R. (1975). The metabolism of pyrimidines by proteolytic clostridia. *Archives of Microbiology*, *102*(1), 145–149. <https://doi.org/10.1007/BF00428359>
- Khan, S. R., Gaines, J., Roop, R. M., & Farrand, S. K. (2008). Broad-host-range expression vectors with tightly regulated promoters and their use to examine the influence of TraR and TraM expression on Ti plasmid quorum sensing. *Appl. Environ. Microbiol.*, *74*(16), 5053–5062.
- Lechner, M., Findeiß, S., Steiner, L., Marz, M., Stadler, P. F., & Prohaska, S. J. (2011). Proteinortho: Detection of (Co-) orthologs in large-scale analysis. *BMC Bioinformatics*, *12*(1), 124. <https://doi.org/10.1186/1471-2105-12-124>
- Leonardi, R., & Jackowski, S. (2007). Biosynthesis of pantothenic acid and coenzyme A. *EcoSal Plus*, *2*(2), 1–20. <https://doi.org/10.1128/ecosal.3.6.3.4>
- Lundgren, S., Gojković, Z., Piškur, J., & Dobritzsch, D. (2003). Yeast β -Alanine synthase shares a structural scaffold and origin with dizinc-dependent exopeptidases. *Journal of Biological Chemistry*, *278*(51), 51851–51862. <https://doi.org/10.1074/jbc.M308674200>
- Martínez-Gomez, A. I., Rodríguez-Vico, F., Clemente-Jimenez, J. M., Servi, S., Tessaro, D., Pozo-Dengra, J., ... Las Heras-Vazquez, F. J. (2008). Potential application of N-carbamoyl-L-alanine amidohydrolase from agrobacterium tumefaciens C58 for L-alanine acid production. *Applied and Environmental Microbiology*, *75*(2), 514–520. <https://doi.org/10.1128/aem.01128-08>
- Martínez-Rodríguez, S., Martínez-Gómez, A. I., Rodríguez-Vico, F., Clemente-Jiménez, J. M., & Las Heras-Vázquez, F. J. (2010). Carbamoylases: Characteristics and applications in biotechnological processes. *Applied Microbiology and Biotechnology*, *85*(3), 441–458. <https://doi.org/10.1007/s00253-009-2250-y>
- Martínez-Romero, E. (2003). Diversity of *Rhizobium phaseolus vulgaris* symbiosis: overview and perspectives. *Plant and Soil*, *252*(1), 11–23. <https://doi.org/10.1023/A:1024199013926>
- Martínez-Romero, E., & Caballero-Mellado, J. (1996). *Rhizobium* phylogenies and bacterial genetic diversity. *Critical Reviews in Plant Sciences*, *15*(2), 113–140. <https://doi.org/10.1080/07352689609701938>
- Muñoz-Gómez, S. A., Hess, S., Burger, G., Franz Lang, B., Susko, E., Slamovits, C. H., & Roger, A. J. (2019). An updated phylogeny of the Alphaproteobacteria reveals that the parasitic Rickettsiales and Holosporales have independent origins. *Elife*, *8*(2013), 1–23. <https://doi.org/10.7554/eLife.42535>
- Nanaya, T., Hidenori, O., Keiko, K., Ikeda, T., & Hama, T. (1982). Purification and properties of B-alanine aminotransferase from Rabbit Liver. *The Biochemical Journal*, *92*(4), 1009–1017.
- Naveed Riaz, N., Rehman M, F. U., Mahboob Ahmad, M., & (2017). B-amino acids: Role in human biology and medicinal chemistry - A review. *Medicinal Chemistry*, *7*(10), 302–307. <https://doi.org/10.4172/2161-0444.1000472>
- Niederwieser, A., Benson, J. V., Patterson, J. A., Pataki, G., Arbor, A., & Hann, C. S. (1971). *The popularity of thin-layer chromatography (TLC) for the analysis of amino acids and derivatives is astonishing*. There are other techniques which give, in general, better qualitative.
- Prell, J., & Poole, P. (2006). Metabolic changes of rhizobia in legume nodules. *Trends in Microbiology*, *14*(4), 161–168. <https://doi.org/10.1016/j.tim.2006.02.005>
- Sambrook, J., Fritsch, E. F., Maniatis, T. (1989). *Molecular cloning: A laboratory manual*, 2nd ed., Vols. 1-3. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Schäfer, A., Tauch, A., Jäger, W., Kalinowski, J., Thierbach, G., & Pühler, A. (1994). Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene*, *145*(1), 69–73.
- Segovia, L., Young, J. P. W., & Martínez, Romero E. (1993). Reclassification of American *Rhizobium leguminosarum* biovar phaseoli Type 1 strains as *Rhizobium etli* sp. nov. *International Journal of Systematic Bacteriology*, *43*, 372–374.
- Slotnick, I. J., & Weinfeld, H. (1956). Dihydrouracil as a growth factor for mutant strains of *Escherichia Coli*. *Journal of Bacteriology*, *74*, 122–125 Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/13475206>
- Stuecker, T. N., Bramhacharya, S., Hodge-Hanson, K. M., Suen, G., & Escalante-Semerena, J. C. (2015). Phylogenetic and amino acid conservation analyses of bacterial L-aspartate- α -decarboxylase and of its zymogen-maturation protein reveal a putative interaction

- domain. *BMC Research Notes*, 8(1), 1–11. <https://doi.org/10.1186/s13104-015-1314-6>
- Tiwari, S., & Lata, C. (2018). Heavy metal stress, signaling, and tolerance due to plant-associated microbes: An overview. *Frontiers in Plant Science*, 9, 1–12. <https://doi.org/10.3389/fpls.2018.00452>
- Tomita, H., Yokooji, Y., Ishibashi, T., Imanaka, T., & Atomia, H. (2014). An archaeal glutamate decarboxylase homolog functions as an aspartate decarboxylase and is involved in β -Alanine and coenzyme a biosynthesis. *Journal of Bacteriology*, 196(6), 1222–1230. <https://doi.org/10.1128/JB.01327-13>
- Villaseñor, T., Dávalos, A., Lozano, L., Brom, S., Romero, D., & Garcia-De-Los-Santos, A. (2011). Housekeeping genes essential for pantothenate biosynthesis are plasmid-encoded in *Rhizobium etli* and *Rhizobium leguminosarum*. *BMC Microbiology*, 11(1), 66. <https://doi.org/10.1186/1471-2180-11-66>
- Wang, Y., Xu, H., & White, R. H. (2014). β -alanine biosynthesis in *Methanocaldococcus jannaschii*. *Journal of Bacteriology*, 196(15), 2869–2875. <https://doi.org/10.1128/JB.01784-14>
- White, W. H., Gunyuzlu, P. L., & Toyn, J. H. (2001). *Saccharomyces cerevisiae* is capable of de Novo Pantothenic acid biosynthesis involving a novel pathway of β -alanine production from spermine. *Journal of Biological Chemistry*, 276(14), 10794–10800. <https://doi.org/10.1074/jbc.M009804200>
- Wilding, M., Peat, T. S., Newman, J., & Scott, C. (2016). A β -alanine catabolism pathway containing a highly promiscuous ω -transaminase in the 12-aminododecanate-degrading *Pseudomonas* sp. Strain AAC. *Applied and Environmental Microbiology*, 82(13), 3846–3856. <https://doi.org/10.1128/aem.00665-16>
- Williams, K. P., Sobral, B. W., & Dickerman, A. W. (2007). A robust species tree for the alphaproteobacteria. *Journal of Bacteriology*, 189(13), 4578–4586. <https://doi.org/10.1128/JB.00269-07>
- Yun, H., Lim, S., Cho, B., & Kim, B. (2004). B-amino acid: Pyruvate transaminase from. *Microbiology*, 70(4), 2529–2534. <https://doi.org/10.1128/AEM.70.4.2529>

How to cite this article: López-Sámamo M, Beltrán LFL-A, Sánchez-Thomas R, et al. A novel way to synthesize pantothenate in bacteria involves β -alanine synthase present in uracil degradation pathway. *MicrobiologyOpen*. 2020;9:e1006. <https://doi.org/10.1002/mbo3.1006>

APPENDIX 1

Mutant	Primer	
CFN42 RHE_CH02599-	5'ATC CTC GAA TTC AAG GCT CTA AGC TGC C 3'	Forward
	5' TGT GAA GGA TCC GCG CTT CAC AAA 3'	Reverse
CFN42 amaB-	5'- CGT GCA GGA TCC GGA CTT CGC CTA TAA C -3'	Forward
	5'- GAG CTT AAG CTT GTC GGG TGA GCG GAT G -3'	Reverse
CFN42 amaB-/amaB <i>R. etli</i>	5'- ATC CGC AAG CTT AAA ACC AAA GGC AAC T -3'	Forward
	5'- GAA GGT GGA TCC AAG GGT CGG ATG A -3'	Reverse
CFN42 amaB-/aam <i>R. etli</i>	5'- CAT GAT GGA TCC GTT TGC GTT GTC CAG A -3'	Forward
	5'- CCC ATC GAA TTC GTT TTG CCG CCG AAT A -3'	Reverse
CFN42 amaB-/amaB <i>A. fab</i>	5'- CGC CAT GGA TCC GCA ATG GCT GTT ATC T -3'	Forward
	5'- CTG CCG GAA TTC ATC CTG ATG TCT GCC T -3'	Reverse
CFN42 amaB-/panD <i>A. fab</i>	5'- CCG ATG AAG CTT CGA CAA AGA TCG GCA A -3'	Forward
	5'- GAT GTC GAA TTC GAA CCT CTG GTC GCC T -3'	Reverse
CFN42 amaB-/panD <i>E. coli</i>	5'- CAC CAG GAA TTC CAT CGT CTC CAG CGA A -3'	Forward
	5'- GGT GAG AAG CTT GCC GCA GGG ATA ACA A -3'	Reverse
BL21 (DE3) AmaB/pETSUMO	5'ATGGTGGCAGCACCAGGCGAGAATGC-3'	Forward
	5'- TCACACCACGATCTCCGCCGTCTCCACC-3'	Reverse

TABLE A1 Primer list

Abbreviations: *R. etli*, *Rhizobium etli*; *A. fab*, *Agrobacterium fabrum*; *E. coli*, *Escherichia coli*.

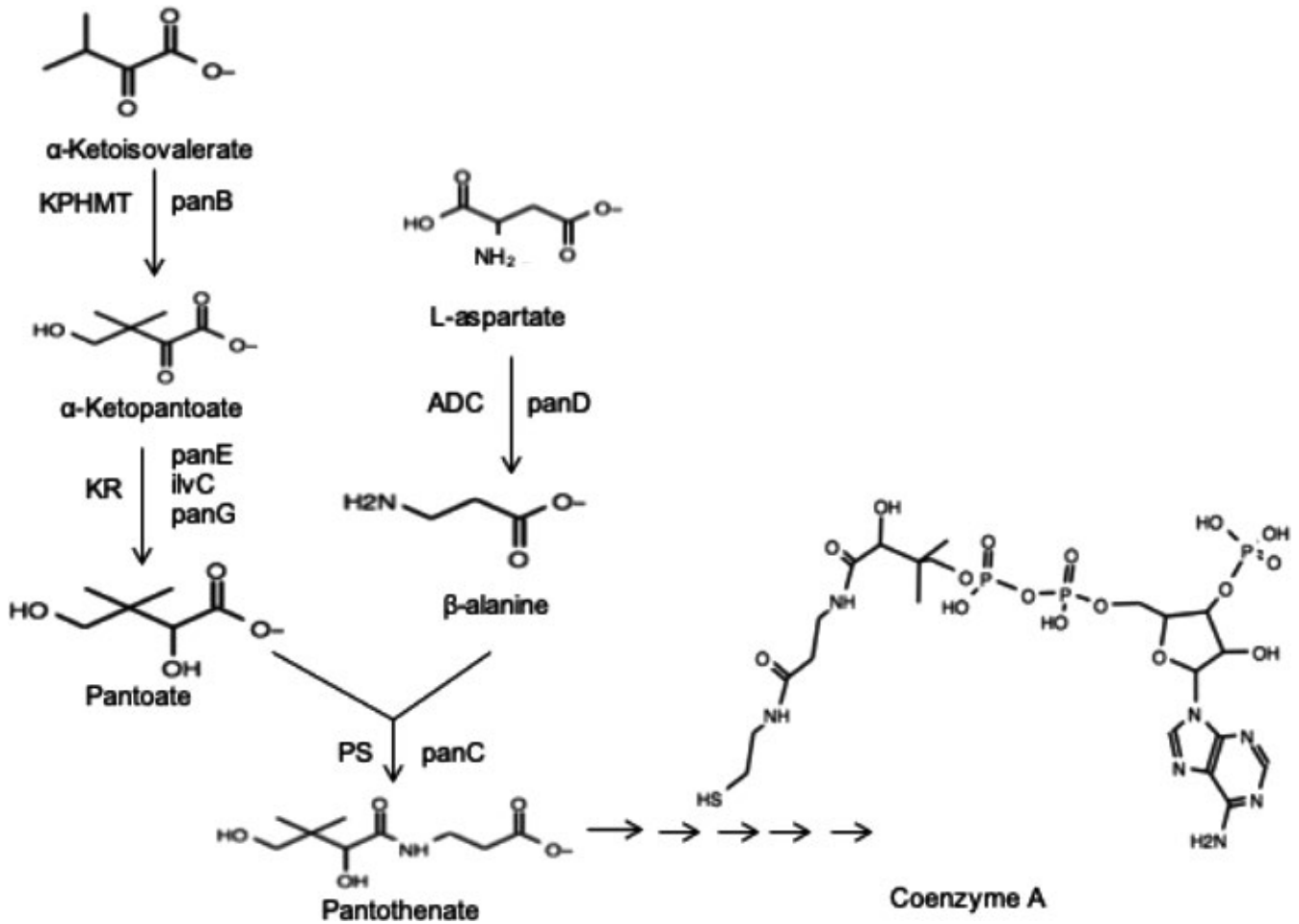


FIGURE A1 Panthotenate and CoA biosynthetic pathway

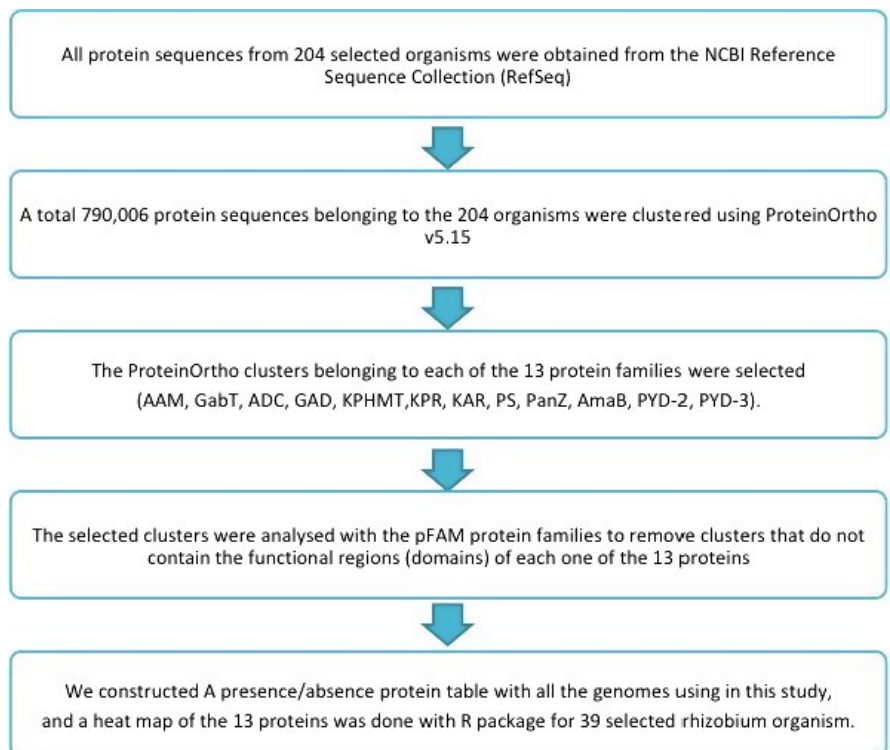


FIGURE A2 Flow diagram of the bioinformatics search methodology

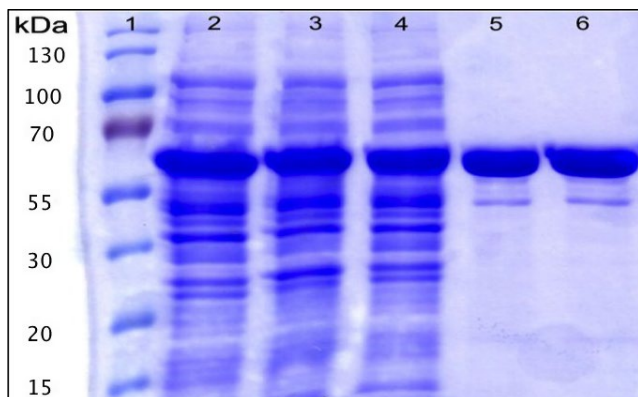


FIGURE A3 SDS-PAGE analysis of the His6-3-Ureidopropionase produced by *Escherichia coli* harboring pETSUMO. Lane 1: Protein size marker, lane 2: cell extract after 6h of induction at 30.C, lane 3 insoluble fraction, lane 4: soluble fraction after induction, lanes 5-6: elution fraction, lane: 7 elution fraction after filtration by 30 kDa amikon, lane8: concentrated elution fraction with glycerol

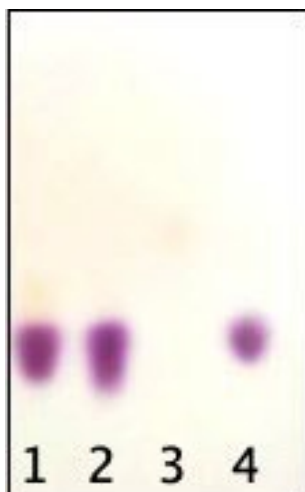


FIGURE A4 TLC enzymatic reactions plate revealed with ninhydrin. (1) Test reaction; (2) Reaction with β -ala; (3) Control reaction without enzyme; (4) β -ala standard

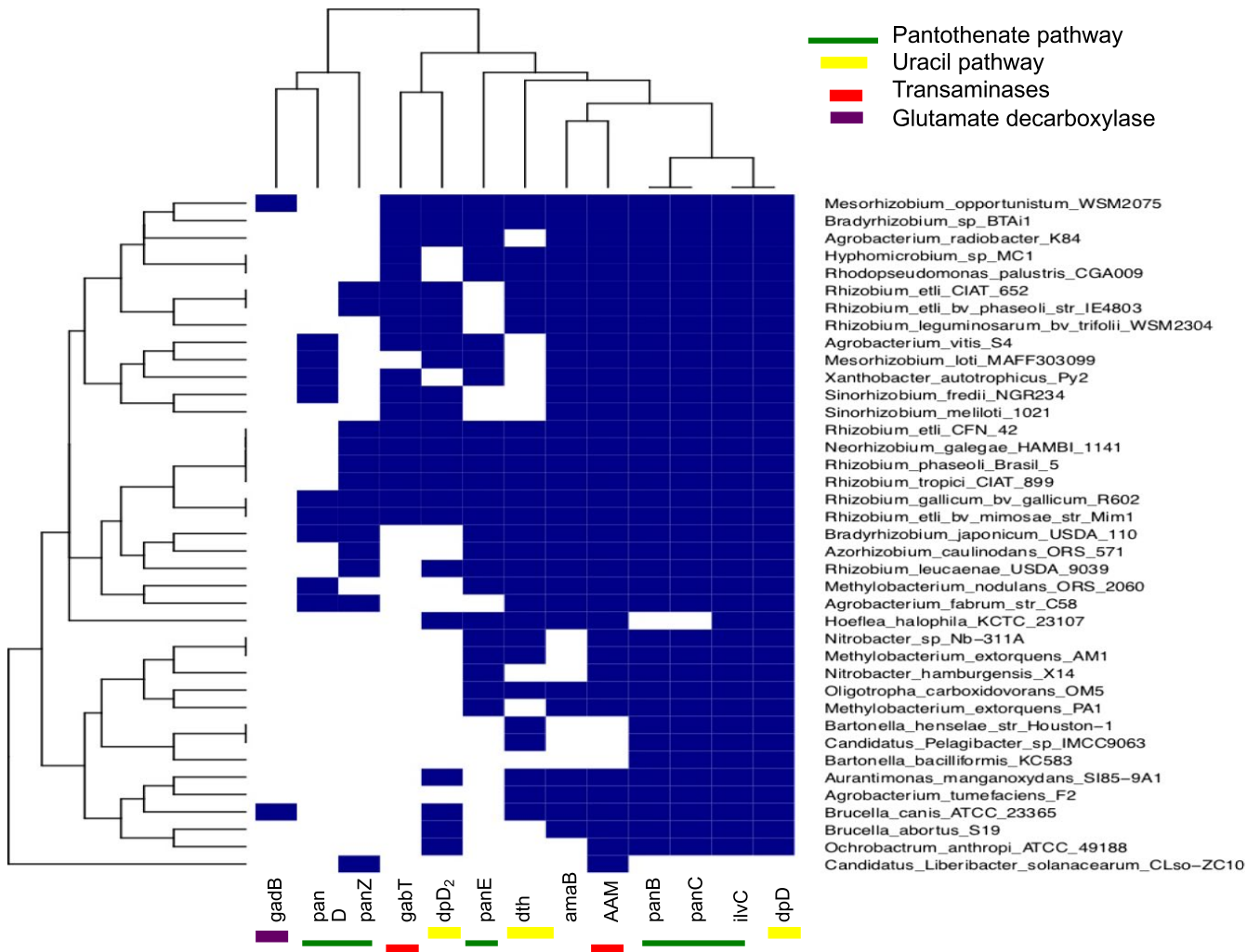


FIGURE A5 Alternatives pathways or enzymes that can act in the absent of ADC enzyme. Heat-map constructed with 40 representative Rhizobiales genomes

TABLE A2 Occurrence analysis in alpha-proteobacteria, represented by number of copies of each gene on strain genome.

Class	Order	Organism	Biosynthesis pathway										Gene				
			Pyrimidine degradation			Pantothenate pathway				Transaminases			Decarboxylase				
			PYD3	PYD2	PYD1	KAR	KPHMT	PS	ADC	KPR	panZ	AAM	GabT	GAD			
Alphaproteobacteria	Rhizobiales	<i>Afipia_sp_1NLS2</i>	1	1	1	1	1	1	1	1	1	1	3	0	1	0	0
		<i>Agrobacterium_fabrum_str_C58</i>	1	1	1	1	1	1	1	1	1	1	0	2	2	0	0
		<i>Agrobacterium_radiobacter_K84</i>	1	0	1	1	1	1	1	1	1	0	1	0	2	1	0
		<i>Agrobacterium_sp_H13-3</i>	1	1	1	1	1	1	1	1	1	0	1	0	2	0	0
		<i>Agrobacterium_tumefaciens_F2</i>	1	1	1	1	1	1	1	1	1	0	0	0	2	0	0
		<i>Agrobacterium_vitis_S4</i>	1	0	1	1	1	1	1	1	1	1	1	0	1	1	0
		<i>Aurantimonas_manganoxydans_S185-9A1</i>	1	3	1	1	1	1	1	1	1	0	0	0	1	0	0
		<i>Azorhizobium_caulinodans_OR5_571</i>	1	1	1	1	1	1	1	1	1	0	1	1	1	0	0
		<i>Bartonella_bacilliformis_KC583</i>	0	0	1	1	1	1	1	1	1	0	0	0	0	0	0
		<i>Bartonella_clarridgeiae_73</i>	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0
		<i>Bartonella_grahamii_as4aup</i>	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0
		<i>Bartonella_henselae_str_Houston-1</i>	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0
		<i>Bartonella_quintana_str_Toulouse</i>	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0
		<i>Bartonella_tribocorum_CIP_105476</i>	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0
		<i>Beijerinckia_indica_subsp_indica_ATCC_9039</i>	0	1	1	1	1	1	1	1	1	1	0	0	1	0	0
		<i>Bradyrhizobiaceae_bacterium_SG-6C (A. clevelandensis ATCC49720)</i>	1	1	1	1	1	1	1	1	1	1	3	0	2	0	0
		<i>Bradyrhizobium_diazoeficiens_USDA_110</i>	1	2	1	1	1	1	1	1	2	1	4	1	2	0	0
		<i>Bradyrhizobium_japonicum_USDA_6</i>	1	4	1	1	1	1	1	2	1	3	0	0	2	0	0
		<i>Bradyrhizobium_sp_BTA11</i>	1	2	1	1	1	1	1	1	1	0	1	0	2	1	0
		<i>Bradyrhizobium_sp_OR5_278</i>	1	4	1	1	1	1	1	1	1	0	1	2	1	1	0
		<i>Brucella_abortus_S19</i>	1	0	1	1	1	1	1	1	1	0	0	0	2	0	0
		<i>Brucella_canis_ATCC_23365</i>	1	1	1	1	1	1	1	1	1	0	0	0	2	0	1
		<i>Brucella_melitensis_bv_1_str_16M</i>	1	1	1	1	1	1	1	1	1	0	0	0	2	0	0
		<i>Brucella_ovis_ATCC_25840</i>	1	0	1	1	1	1	1	1	1	0	0	0	2	0	1
		<i>Brucella_Suis_1330</i>	1	2	1	1	1	1	1	1	1	0	0	0	2	0	1

(Continues)

TABLE A2 (Continued)

Class	Order	Organism	Biosynthesis pathway										Gene				
			Pyrimidine degradation					Pantothenate pathway					Transaminases		Decarboxylase		
			PYD3	PYD2	PYD1	KAR	KPHMT	PS	ADC	KPR	panZ	AAM	GabT	GAD			
		Candidatus_Liberibacter_asiaticus_str_psy62	0	0	1	0	0	0	0	0	0	0	0	1	0	0	
		Candidatus_Liberibacter_solanacearum_CLso-ZC1	0	0	1	0	0	0	0	0	0	0	0	1	0	0	
		Candidatus_Midichloria_mitochondrii_IricVA	0	0	0	0	0	0	0	0	0	0	0	1	0	0	
		Candidatus_Odyssella_thessalonicensis_L13	0	0	0	0	0	0	0	0	0	0	0	1	0	0	
		Candidatus_Puniceispirillum_marinum_IMCC1322	1	0	1	1	1	1	0	1	0	0	2	0	0	0	
		Chelatorans_sp_BNC1	1	0	1	1	1	1	1	1	0	0	0	1	0	0	
		Fulvimarina_pelagi_HTCC2506	0	2	1	1	1	1	1	1	0	0	1	0	0	0	
		Hoeflea_halophila_KCTC_23107	1	1	1	1	1	0	0	0	0	0	1	0	0	0	
		Hoeflea_phototrophica_DFL-43	1	1	1	1	1	0	0	0	0	0	1	0	0	0	
		Hyphomicrobium_denitrificans_ATCC_51888	0	1	1	1	1	1	1	1	1	0	1	2	1	0	
		Hyphomicrobium_sp_MC1	1	1	1	1	1	1	1	1	0	0	2	0	1	0	
		Mesorhizobium_ciceri_biovar_biserrulae_WSM1271	1	2	1	1	3	2	0	2	0	0	2	1	1	0	
		Mesorhizobium_japonicum_MAFF303099	1	0	1	1	2	1	2	1	0	2	2	0	3	1	0
		Mesorhizobium_lotii_NZP2037	1	3	1	1	3	1	2	1	0	2	3	0	3	1	0
		Mesorhizobium_opportunisticum_WSM2075	1	2	1	1	3	2	0	2	0	0	2	0	2	1	1
		Methylobacterium_extorquens_AM1	0	1	1	1	1	1	0	2	0	0	2	0	1	0	0
		Methylobacterium_extorquens_PA1	1	0	1	1	1	1	0	2	0	0	2	0	0	0	0
		Methylobacterium_nodulans_OR5_2060	1	2	1	1	1	1	1	1	1	1	3	0	3	0	0
		Methylobacterium_populii_BJ001	0	1	1	1	1	1	0	2	0	0	2	0	1	0	0
		Methylobacterium_radiotolerans_JCM_2831	2	0	1	1	1	1	0	2	0	0	2	0	2	1	0
		Methylobacterium_sp_4-46	0	1	1	1	1	1	0	1	0	0	1	4	0	0	0

(Continues)

TABLE A2 (Continued)

Class	Order	Organism	Biosynthesis pathway						Pantothenate pathway						Gene			
			PYD3	PYD2	PYD1	KAR	KPHMT	PS	ADC	KPR	panZ	AAM	GabT	GAD	Transaminases	Decarboxylase		
		<i>Methylocella silvestris</i> _BL2	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0	
		<i>Methylocystis</i> _sp._ATCC_49242	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0	
		<i>Methylosinus trichosporium</i> _OB3b	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0	
		<i>Neorhizobium galegae</i> _HAMB1_1141	1	1	1	1	1	2	1	1	1	1	1	1	1	0	0	
		<i>Nitrobacter hamburgensis</i> _X14	0	0	1	1	1	1	1	1	1	1	1	0	0	0	0	
		<i>Nitrobacter</i> _sp._Nb-311A	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0	
		<i>Nitrobacter winogradskyi</i> _Nb-255	0	0	1	1	1	1	1	1	1	1	1	0	0	0	0	
		<i>Ochrobactrum anthroproli</i> _ATCC_49188	0	0	1	1	1	1	1	1	1	1	1	0	0	0	0	
		<i>Oligotropha carboxidovorans</i> _OM5	1	1	1	1	1	1	1	1	1	1	0	3	0	0	0	
		<i>Parvibaculum lavamentivorans</i> _DS-1	0	0	1	1	1	1	1	1	1	1	0	0	0	0	1	
		<i>Pelagibacterium halotolerans</i> _B2	1	2	1	1	1	1	1	1	1	1	0	0	0	0	0	
		<i>Rhizobium etli</i> _bv_mimosae_str_Mim1	1	1	1	1	1	2	1	1	1	1	1	2	2	1	0	
		<i>Rhizobium etli</i> _bv_phaseoli_str_IE4803	1	1	1	1	1	2	1	1	1	1	0	0	1	1	0	
		<i>Rhizobium etli</i> _CFN_42	1	1	1	1	1	2	1	1	1	1	0	1	2	1	0	
		<i>Rhizobium etli</i> _CIAT_652	1	1	1	1	1	2	1	1	1	1	0	0	1	1	0	
		<i>Rhizobium gallicum</i> _bv_gallicum_R602	1	1	1	1	1	1	1	1	1	1	1	2	1	1	0	
		<i>Rhizobium leguminosarum</i> _bv_phaseoli_CCGM1	1	1	1	1	1	2	1	1	1	1	0	0	1	1	0	
		<i>Rhizobium leguminosarum</i> _bv_trifolii_WSM2304	1	1	1	1	1	2	1	1	1	1	0	0	0	1	0	
		<i>Rhizobium leguminosarum</i> _bv_viciae_3841	1	2	1	1	1	2	1	1	1	1	0	1	1	1	0	
		<i>Rhizobium leucaenae</i> _USDA_9039	1	1	1	1	1	3	1	1	1	1	0	1	1	0	0	
		<i>Rhizobium phaseoli</i> _Brasil_5	1	1	1	1	1	2	1	1	1	1	0	2	1	1	0	
		<i>Rhizobium tropici</i> _CIAT_899	1	1	1	1	1	1	1	1	1	1	0	2	1	1	0	
		<i>Rhodocyclidium vanniellii</i> _ATCC_17100	0	0	1	1	1	1	1	1	1	1	0	0	0	0	1	
		<i>Rhodospseudomonas palustris</i> _BisA53	0	0	1	1	1	1	1	1	1	1	0	2	0	0	0	

(Continues)

TABLE A2 (Continued)

Class	Order	Organism	Biosynthesis pathway										Gene			
			Pyrimidine degradation			Pantothenate pathway				Transaminases			Decarboxylase			
			PYD3	PYD2	PYD1	KAR	KPHMT	PS	ADC	KPR	panZ	AAM		GabT	GAD	
		Rhodopseudomonas_palustris_BisB18	1	0	1	1	1	1	0	2	0	0	1	1	0	
		Rhodopseudomonas_palustris_BisB5	1	0	1	1	1	1	0	3	0	0	2	0	0	
		Rhodopseudomonas_palustris_CGA009	1	1	1	1	1	1	0	3	0	0	2	1	0	
		Rhodopseudomonas_palustris_DX-1	1	1	1	1	1	1	0	2	0	0	2	1	0	
		Rhodopseudomonas_palustris_HaA2	1	0	1	1	1	1	0	3	0	0	2	1	0	
		Sinorhizobium_fredii_HH103	1	1	1	1	1	2	1	0	2	1	3	1	0	
		Sinorhizobium_fredii_NGR234	1	0	1	1	1	2	1	0	0	1	3	1	0	
		Sinorhizobium_medicae_WSM419	1	0	1	1	1	2	1	0	0	0	1	1	0	
		Sinorhizobium_melioloti_1021	1	0	1	1	1	2	1	0	0	0	2	1	0	
		Starkeya_novella_DSM_506	1	3	1	1	1	1	1	0	0	1	1	0	0	
		Xanthobacter_autotrophicus_Py2	1	0	1	1	1	1	1	1	1	1	2	1	0	
		Ahrensia_sp_R2A130	0	2	1	1	1	0	0	0	0	0	1	0	0	
		Citricella_sp_SE45	1	1	1	1	1	1	0	1	0	0	0	0	0	
		Dinoroseobacter_shibae_DFL_12	1	0	1	1	1	1	1	0	0	1	0	0	0	
		Hirschia_baltica_ATCC_49814	0	0	1	1	1	1	1	0	1	0	1	0	0	
		Hyphomonas_neptunium_ATCC_15444	0	0	1	1	1	1	1	1	1	0	1	0	1	
		Jannaschia_sp_CCS1	0	1	1	1	1	0	0	0	0	0	1	1	0	
		Ketogulonigenium_vulgarum_WSH-001	1	2	1	1	1	0	0	1	0	0	0	0	0	
		Labrenzia_aggregata_IAM_12614	1	2	1	1	1	1	0	1	0	1	3	1	1	
		Labrenzia_alexandrii_DFL-11	1	2	1	1	1	1	1	0	0	0	1	0	1	
		Maricaulis_maris_MCS10	0	0	1	1	1	1	1	0	0	0	0	0	0	
		Maritimibacter_alkaliphilus_HTCC2654	1	1	1	1	1	1	1	1	1	1	0	0	0	
		Oceanibulbus_indolifex_HEL-45	1	2	1	1	1	0	0	1	0	0	0	1	0	
		Oceanicaulis_sp_HTCC2633	0	0	1	1	1	1	1	0	1	0	0	0	0	
		Oceanicola_batsensis_HTCC2597	1	1	1	1	1	1	1	0	1	0	0	0	0	

(Continues)

TABLE A2 (Continued)

Class	Order	Organism	Biosynthesis pathway					Pantothenate pathway					Gene			
			PYD3	PYD2	PYD1	KAR	KPHMT	PS	ADC	KPR	panZ	AAM	GabT	GAD		
		Oceanicola_granulosus_HTCC2516	1	1	1	1	1	1	1	0	1	0	0	1	0	0
		Octadecabacter_antarcticus_238	1	1	1	1	1	0	0	0	0	0	0	0	1	0
		Octadecabacter_antarcticus_307	1	2	1	1	1	0	0	0	0	0	0	0	1	0
		Paracoccus_denitrificans_PD1222	2	2	1	1	1	1	1	0	1	0	0	1	1	0
		Paracoccus_sp_TRP	1	0	1	1	1	1	1	0	0	0	0	1	1	0
		Pelagibaca_bermudensis_HTCC2601	0	2	1	1	1	1	1	0	1	0	2	0	0	0
		Phaeobacter_gallaeciensis_DSM_17395	1	1	1	1	1	1	1	0	0	0	0	1	0	0
		Pseudovibrio_sp_FO-BEG1	1	1	1	1	1	1	1	0	0	0	0	2	0	1
		Rhodobacteraceae_bacterium_HTCC2083	1	2	1	1	1	0	0	0	0	0	0	0	0	0
		Rhodobacteraceae_bacterium_HTCC2150	0	1	1	1	1	1	1	0	0	0	0	0	0	0
		Rhodobacteraceae_bacterium_KLH11	2	1	1	1	1	1	1	0	0	0	0	0	1	0
		Rhodobacterales_bacterium_HTCC2255	1	1	1	1	1	1	1	0	0	0	0	0	0	0
		Rhodobacterales_bacterium_Y4I	1	2	1	1	1	1	1	0	0	0	0	2	0	0
		Rhodobacter_capsulatus_SB_1003	0	1	1	1	1	1	1	0	0	0	0	2	1	0
		Rhodobacter_sphaeroides_241	1	1	1	1	1	1	1	0	1	0	0	0	0	0
		Rhodobacter_sphaeroides_ATCC_17025	1	1	1	1	1	1	1	0	0	0	0	0	0	0
		Rhodobacter_sp_SW2	1	1	1	1	1	1	1	0	0	0	0	1	1	0
		Roseibium_sp_TrichSKD4	1	1	1	1	1	1	1	0	1	1	1	1	1	0
		Roseobacter_denitrificans_OCh_114	1	0	1	1	1	1	1	0	1	0	0	0	1	0
		Roseobacter_litoralis_Och_149	1	2	1	1	1	1	1	0	1	0	1	1	0	0
		Roseobacter_sp_AzWK-3b	1	1	1	1	1	1	1	0	0	0	0	0	0	0
		Roseobacter_sp_CCS2	1	1	1	1	1	1	1	0	1	0	0	0	0	0
		Roseobacter_sp_GAI101	1	1	1	1	1	0	0	0	1	0	0	0	1	0
		Roseobacter_sp_MED193	1	1	1	1	1	1	1	0	1	1	1	2	0	0

(Continues)

TABLE A2 (Continued)

Class	Order	Organism	Biosynthesis pathway										Gene			
			Pyrimidine degradation			Pantothenate pathway				Transaminases			Decarboxylase			
			PYD3	PYD2	PYD1	KAR	KPHMT	PS	ADC	KPR	panZ	AAM		GabT	GAD	
		Roseobacter_sp_SK209-2-6	2	1	1	1	2	1	0	0	1	0	0	1	0	0
		Roseovarius_nubinihibens_ISM	1	2	1	1	1	1	0	0	0	0	0	0	0	0
		Roseovarius_sp_217	1	1	1	1	1	1	1	0	1	0	0	1	0	0
		Roseovarius_sp_TM1035	1	1	1	1	1	1	0	0	0	0	0	1	0	0
		Ruegeria_pomeroyi_DSS-3	1	0	1	1	1	1	0	2	0	0	0	2	1	0
		Ruegeria_sp_R11	1	1	1	1	1	1	0	0	0	0	1	0	0	0
		Ruegeria_sp_TM1040	1	0	1	1	1	1	0	0	0	0	0	1	0	0
		Sagittula_stellata_E-37	1	3	1	1	1	1	0	3	0	0	1	1	0	0
		Silicibacter_sp_TrichCH4B	2	2	1	1	1	1	0	1	0	0	0	1	0	0
		Sulfitobacter_sp_EE-36	1	0	1	1	1	1	0	1	0	0	0	0	1	0
		Thalassibium_sp_R2A62	1	1	1	1	1	1	0	0	0	0	0	0	0	0
		Anaplasma_centrale_str_Israel	0	0	1	1	1	1	0	0	0	0	1	0	0	0
		Anaplasma_marginale_str_Florida	0	0	1	1	1	1	0	0	0	0	1	0	0	0
		Anaplasma_marginale_str_St_Maries	0	0	1	1	1	1	0	0	0	0	1	0	0	0
		Anaplasma_phagocytophilum_HZ	0	0	1	0	0	0	0	0	0	0	0	0	0	0
		Ehrlichia_canis_str_Jake	0	0	1	0	0	0	0	0	0	0	1	0	0	0
		Ehrlichia_chaffeensis_str_Arkansas	0	0	1	0	0	0	0	0	0	0	1	0	0	0
		Ehrlichia_ruminantium_str_Welgevonden	0	0	1	0	0	0	0	0	0	0	1	0	0	0
		Neorickettsia_risticii_str_Illinois	0	0	1	0	0	0	0	0	0	0	1	0	0	0
		Neorickettsia_sennetsu_str_Miyayama	0	0	1	0	0	0	0	0	0	0	1	0	0	0
		Orientia_tsutsugamushi_str_Boryong	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		Rickettsia_akari_str_Hartford	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		Rickettsia_bellii_OSU_85-389	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		Rickettsia_canadensis_str_McKiel	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		Rickettsia_conorii_Malish_7	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		Rickettsia_endosymbiont_of_Ixodes_scapularis	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		Rickettsia_felis_URRWXCal2	0	0	0	0	0	0	0	0	0	0	0	0	0	0

(Continues)

TABLE A2 (Continued)

Class	Order	Organism	Biosynthesis pathway						Gene						
			Pyrimidine degradation			Pantothenate pathway			Transaminases			Decarboxylase			
			PYD3	PYD2	PYD1	KAR	KPHMT	PS	ADC	KPR	panZ	AAM	GabT	GAD	
		<i>Rickettsia prowazekii_str_Madrid_E</i>	0	0	0	0	0	0	0	0	0	0	0	0	0
		<i>Rickettsia sibirica_246</i>	0	0	0	0	0	0	0	0	0	0	0	0	0
		<i>Rickettsia typhi_str_Wilmington</i>	0	0	0	0	0	0	0	0	0	0	0	0	0
		<i>Wolbachia_sp_wRi</i>	0	0	1	0	0	0	0	0	0	0	0	0	0
	<i>Sphingomonadales</i>	<i>Blastomonas_sp_RAC04</i>	0	1	1	1	1	1	1	0	0	1	0	0	0
		<i>Citromicrobium_bathymarinum_JL354</i>	0	0	1	1	1	1	1	0	0	0	0	0	0
		<i>Citromicrobium_sp_JLT1363</i>	0	0	1	1	1	1	1	0	0	0	0	1	0
		<i>Novosphingobium_aromaticivorans_DSM_12444</i>	0	0	1	1	1	1	1	0	2	0	0	1	0
		<i>Novosphingobium_capsulatum_NBRC_12533</i>	0	0	1	1	1	1	1	0	2	0	0	2	0
		<i>Novosphingobium_nitrogenifigans_DSM_19370</i>	0	0	1	1	1	1	1	0	0	0	0	2	0
		<i>Novosphingobium_sp_PP1Y</i>	1	1	1	1	1	1	1	0	0	1	0	1	0
		<i>Sphingobium_chlorophenolicum_L-1</i>	1	1	1	1	1	1	1	0	0	0	0	1	0
		<i>Sphingobium_japonicum_UT26S</i>	0	1	1	1	1	1	1	0	0	0	0	1	0
		<i>Sphingobium_sp_SYK-6</i>	0	1	1	1	1	1	1	0	0	1	0	0	0
		<i>Sphingobium_yanoikuyae_XLDN2-5</i>	0	1	1	1	1	1	1	0	0	0	0	0	0
		<i>Sphingomonas_paucimobilis_NBRC_13935</i>	0	1	1	1	1	1	1	0	0	0	0	1	0
		<i>Sphingomonas_sp_S17</i>	0	1	1	1	1	1	1	0	0	0	0	1	0
		<i>Sphingomonas_sp_SKA58</i>	0	1	1	1	1	1	1	0	0	0	0	1	0
		<i>Sphingomonas_wittichii_RW1</i>	0	0	1	1	1	1	1	0	2	0	0	1	0
		<i>Sphingopyxis_alaskensis_RB2256</i>	0	0	1	1	1	1	1	0	0	0	0	2	0
		<i>Sphingopyxis_macroglabida_strain_203</i>	0	0	1	1	1	1	1	0	0	0	1	2	0
		<i>Zymomonas_mobilis_subsp_mobilis_ATCC_10988</i>	0	1	0	1	1	1	1	0	0	0	0	0	0
		<i>Zymomonas_mobilis_subsp_pomaceae_ATCC_29192</i>	0	1	0	1	0	0	0	0	0	0	0	0	0

(Continues)

TABLE A2 (Continued)

Class	Order	Organism	Biosynthesis pathway										Gene					
			Pyrimidine degradation			Pantothenate pathway				Transaminases			Decarboxylase					
			PYD3	PYD2	PYD1	KAR	KPHMT	PS	ADC	KPR	panZ	AAM		GabT	GAD			
Rhodospirillales		Acetobacter_pasteurianus_	1	0	1	1	1	1	1	1	1	0	2	0	2	1	0	
		IFO_3283-01																
		Acetobacter_pomorum_DM001	0	1	1	1	1	1	1	1	1	0	1	0	1	0	0	0
		Acidiphilium_cryptum_JF-5	1	0	1	1	1	1	1	1	1	0	2	0	1	1	1	0
		Azospirillum_brasiliense_Sp245	1	1	1	1	1	1	1	1	1	1	1	2	0	2	1	0
		Azospirillum_lipoferum_4B	2	2	1	1	1	1	1	1	1	0	2	0	1	1	1	0
		Azospirillum_sp_B510	2	0	1	1	1	1	1	1	1	0	2	0	0	1	1	0
		Enhydrobacter_aerosaccus_SK60	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0
		Erythrobacter_litoralis_HTCC2594	0	0	1	1	1	1	1	1	1	0	0	0	0	0	0	1
		Erythrobacter_sp_NAP1	0	0	1	1	1	1	1	1	1	0	0	0	0	0	0	1
Caulobacteriales		Erythrobacter_sp_SD-21	0	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0
		Gluconacetobacter_diazotrophicus_	1	2	1	1	1	1	1	1	1	0	1	0	2	1	0	0
		PAI_5																
		Gluconacetobacter_hansenii_	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	0
		ATCC_23769																
		Gluconacetobacter_sp_SXCC-1	1	2	1	1	1	1	1	1	1	0	1	0	2	1	1	0
		Gluconacetobacter_xylinus_	1	2	1	1	1	1	1	1	0	1	0	0	2	0	0	0
		NBRC_3288																
		Gluconobacter_oxydans_621H	0	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0
		Granulibacter_bethesdensis_	1	0	1	1	1	1	1	1	1	0	0	1	0	1	0	1
CGDNIH1																		
Magnetospirillum_magneticum_	0	0	1	1	1	1	1	1	1	1	1	1	0	1	0	0		
AMB-1																		
Rhodospirillum_rubrum_	0	0	1	1	1	1	1	1	1	0	0	1	0	1	0	0		
ATCC_11170																		
Asticcacaulis_biprosthecum_C19	0	1	1	2	1	1	1	1	1	1	1	1	0	1	1	0		
Asticcacaulis_excentricus_CB_48	0	1	1	2	1	1	1	1	1	1	1	1	0	1	0	0		
Brevundimonas_diminuta_	0	1	1	1	1	0	0	0	0	0	0	0	0	1	0	0		
ATCC_11568																		
Brevundimonas_sp_BAL3	0	1	1	2	1	1	1	1	1	1	1	0	0	0	0	0		
Brevundimonas_subvibrioides_	0	1	1	2	1	1	1	1	1	1	1	1	0	1	0	0		
ATCC_15264																		

(Continues)

TABLE A2 (Continued)

Class	Order	Organism	Biosynthesis pathway										Gene			
			Pyrimidine degradation			Pantothenate pathway				Transaminases			Decarboxylase			
			PYD3	PYD2	PYD1	KAR	KPHMT	PS	ADC	KPR	panZ	AAM	GabT	GAD		
		Brevundimonas_vesicularis_FDAARGOS_289	0	1	1	2	1	1	1	1	1	1	0	0	0	0
		Caulobacter_crescentus_CB15	0	0	1	1	1	1	1	1	1	1	0	0	0	0
		Caulobacter_mirabilis_FWC_38	0	1	1	1	1	1	1	1	1	1	0	0	0	0
		Caulobacter_segnis_ATCC_21756	0	1	1	1	1	1	1	1	1	1	0	0	0	0
		Caulobacter_sp_K31	0	0	1	1	1	1	1	1	1	1	0	0	0	0
		Phenyllobacterium_zucineum_HLK1	0	0	1	1	1	1	1	1	1	0	1	0	1	1
	Magnetococcales	Loktanelia_vestfoldensis_SKA53	1	2	1	1	1	0	0	0	0	1	0	0	0	0
		Magnetococcus_marinus_MC-1	0	0	1	1	1	1	1	1	1	0	0	1	0	0
	Parvularculales	Parvularcula_bermudensis_HTCC2503	0	1	1	1	1	1	1	1	1	0	0	0	0	0
	Pelagibacterales	Candidatus_Pelagibacter_sp_HTCC7211	0	1	1	1	1	1	1	1	0	0	0	0	0	0
		Candidatus_Pelagibacter_sp_IMCC9063	0	1	1	1	1	1	1	1	0	0	0	0	0	0
		Candidatus_Pelagibacter_ubique_HTCC1062	0	0	1	1	1	1	1	1	0	0	0	0	0	0
			108	123	187	183	167	170	35	44	28	137	68	14		
			PYD3	PYD2	PYD1	KAR	KPHMT	PS	ADC	KPR	panZ	AAM	GabT	GAD		

TABLE A3 ADC phylogeny data set

Accession number	Organism	Order	Class
WP_0093402	<i>Afipia</i> sp. 1NLS2	Rhizobiales	α-proteobacteria
NP_356949	<i>Agrobacterium fabrum</i> str. C58	Rhizobiales	α-proteobacteria
WP_0416991	<i>Agrobacterium vitis</i>	Rhizobiales	α-proteobacteria
WP_0027130	<i>Afipia clevelandensis</i>	Rhizobiales	α-proteobacteria
WP_0123835	<i>Beijerinckia indica</i>	Rhizobiales	α-proteobacteria
NP_768736	<i>Bradyrhizobium diazoefficiens</i> USDA 110	Rhizobiales	α-proteobacteria
WP_0281441	<i>Bradyrhizobium japonicum</i> USDA 6	Rhizobiales	α-proteobacteria
WP_0109135	(1) <i>Mesorhizobium japonicum</i> MAFF 303099	Rhizobiales	α-proteobacteria
WP_0109160	<i>Mesorhizobium japonicum</i> MAFF 303099	Rhizobiales	α-proteobacteria
WP_0198632	(1) <i>Mesorhizobium loti</i> NZP2037	Rhizobiales	α-proteobacteria
WP_0198633	<i>Mesorhizobium loti</i> NZP2037	Rhizobiales	α-proteobacteria
WP_0159320	<i>Methylobacterium nodulans</i>	Rhizobiales	α-proteobacteria
WP_0125900	<i>Methylocella silvestris</i>	Rhizobiales	α-proteobacteria
WP_0209190	<i>Rhizobium etli</i> bv. Mimosae str. Mim1	Rhizobiales	α-proteobacteria
WP_0401142	<i>Rhizobium gallicum</i>	Rhizobiales	α-proteobacteria
WP_0143319	<i>Sinorhizobium fredii</i> HH103	Rhizobiales	α-proteobacteria
YP_0028234	<i>Sinorhizobium fredii</i> NGR234	Rhizobiales	α-proteobacteria
WP_0121143	<i>Xanthobacter autotrophicus</i>	Rhizobiales	α-proteobacteria
WP_0062710	<i>Asticcacaulis biprosthecium</i>	Caulobacterales	α-proteobacteria
WP_0134778	<i>Asticcacaulis excentricus</i>	Caulobacterales	α-proteobacteria
WP_0082631	<i>Brevundimonas</i> sp. BAL3	Caulobacterales	α-proteobacteria
WP_0132701	<i>Brevundimonas subvibrioides</i>	Caulobacterales	α-proteobacteria
WP_0666264	<i>Brevundimonas vesicularis</i>	Caulobacterales	α-proteobacteria
NP_421098	<i>Caulobacter crescentus</i> CB15	Caulobacterales	α-proteobacteria
WP_0996228	<i>Caulobacter mirabilis</i>	Caulobacterales	α-proteobacteria
WP_0109201	<i>Caulobacter segnis</i> ATCC 21756	Caulobacterales	α-proteobacteria
WP_0122875	<i>Caulobacter</i> sp. K31	Caulobacterales	α-proteobacteria
WP_0125231	<i>Phenylobacterium zucineum</i>	Caulobacterales	α-proteobacteria
WP_0133008	<i>Parvularcula bermudensis</i>	Parvularculales	α-proteobacteria
EEV21831.1	<i>Enhydrobacter aerosaccus</i> SK60	Rhodospirillales	α-proteobacteria
WP_0141979	<i>Azospirillum brasilense</i>	Rhodospirillales	α-proteobacteria
WP_0113863	<i>Magnetospirillum magneticum</i>	Rhodospirillales	α-proteobacteria
WP_0158280	<i>Hirschia baltica</i>	Rhodobacterales	α-proteobacteria
WP_0116463	<i>Hyphomonas neptunium</i> ATCC 15444	Rhodobacterales	α-proteobacteria
WP_0116440	<i>Maricaulis maris</i>	Rhodobacterales	α-proteobacteria
WP_0098031	<i>Oceanicaulis</i> sp. HTCC2633	Rhodobacterales	α-proteobacteria
WP_0083357	<i>Maritimibacter alkaliphilus</i>	Rhodobacterales	α-proteobacteria
YP_224431	<i>Corynebacterium glutamicum</i> ATCC 13032	Corynebacteriales	ε-proteobacteria
WP_0108978	<i>Bacillus halodurans</i> c-125	Bacillales	ε-proteobacteria
NP_414673	<i>Escherichia coli</i> str. K-12 substr. MG1655	Enterobacterales	γ-proteobacteria
ABG68179.1	<i>Escherichia coli</i> 536	Enterobacterales	γ-proteobacteria
NP_459185	<i>Salmonella enterica</i> subsp. enterica serovar Typhimurium str. LT2	Enterobacterales	γ-proteobacteria
F6FYI9 F6F	<i>Ralstonia solanacearum</i> Po82	Burkholderiales	β-proteobacteria
NP_880521	<i>Bordetella pertussis</i> Tohama I	Burkholderiales	β-proteobacteria
NP_253419	<i>Pseudomonas aeruginosa</i> PAO1	Pseudomonadales	β-proteobacteria