BMC Research Notes

RESEARCH ARTICLE

Phylogenetic and amino acid conservation analyses of bacterial ʟ-aspartate-α-decarboxylase and of its zymogen-maturation protein reveal a putative interaction domain

Tara N Stuecker[1,4], Shanti Bramhacharya[1,2], Kelsey M Hodge-Hanson[1,3], Garret Suen[1] and Jorge C Escalante-Semerena[5*]

Abstract

Background: All organisms must synthesize the enzymatic cofactor coenzyme A (CoA) from the precursor pantothenate. Most bacteria can synthesize pantothenate de novo by the condensation of pantoate and β-alanine. The synthesis of β-alanine is catalyzed by ʟ-aspartate-α-decarboxylase (PanD), a pyruvoyl enzyme that is initially synthesized as a zymogen (pro-PanD). Active PanD is generated by self-cleavage of pro-PanD at Gly24-Ser25 creating the active-site pyruvoyl moiety. In *Salmonella enterica,* this cleavage requires PanM, an acetyl-CoA sensor related to the Gcn5-like *N*-acetyltransferases. PanM does not acetylate pro-PanD, but the recent publication of the three-dimensional crystal structure of the PanM homologue PanZ in complex with the PanD zymogen of *Escherichia coli* provides validation to our predictions and provides a framework in which to further examine the cleavage mechanism. In contrast, PanD from bacteria lacking PanM efficiently cleaved in the absence of PanM in vivo.

Results: Using phylogenetic analyses combined with in vivo phenotypic investigations, we showed that two classes of bacterial ʟ-aspartate-α-decarboxylases exist. This classification is based on their posttranslational activation by self-cleavage of its zymogen. Class I ʟ-aspartate-α-decarboxylase zymogens require the acetyl-CoA sensor PanM to be cleaved into active PanD. This class is found exclusively in the Gammaproteobacteria. Class II ʟ-aspartate-α-decarboxylase zymogens self cleave efficiently in the absence of PanM, and are found in a wide number of bacterial phyla. Several members of the Euryarchaeota and Crenarchaeota also contain Class II ʟ-aspartate-α-decarboxylases. Phylogenetic and amino acid conservation analyses of PanM revealed a conserved region of PanM distinct from conserved regions found in related Gcn5-related acetyltransferase enzymes (Pfam00583). This conserved region represents a putative domain for interactions with ʟ-aspartate-α-decarboxylase zymogens. This work may inform future biochemical and structural studies of pro-PanD-PanM interactions.

Conclusions: Experimental results indicate that *S. enterica* and *C. glutamicum* ʟ-aspartate-α-decarboxylases represent two different classes of homologues of these enzymes. Class I homologues require PanM for activation, while Class II self cleave in the absence of PanM. Computer modeling of conserved amino acids using structure coordinates of PanM and ʟ-aspartate-α-decarboxylase available in the protein data bank (RCSB PDB) revealed a putative site of interactions, which may help generate models to help understand the molecular details of the self-cleavage mechanism of ʟ-aspartate-α-decarboxylases.

*Correspondence: jcescala@uga.edu
[5] Department of Microbiology, University of Georgia, 212C, Biological Sciences Building, 120 Cedar Street, Athens, GA 30602, USA
Full list of author information is available at the end of the article

Stuecker *et al. BMC Res Notes* (2015) 8:354

Page 2 of 11

Background

Coenzyme A (CoA) is essential to all forms of life. Prokaryotes use CoA for diverse purposes. In some cases CoA is used as a carrier of naturally occurring or xenobiotic, weak organic acids of varied lengths, while in others CoA is critical to the maintenance of thiol-based redox homeostasis [1–5]. CoA is synthesized from the precursor pantothenate, which can be generated de novo by bacteria, plants and fungi [6]. In bacteria, pantothenate synthesis is a branched pathway in which the intermediates β-alanine and pantoate are generated and then condensed to form pantothenate [7]. L-aspartate-α-decarboxylase (PanD; EC 4.1.1.11) catalyzes the conversion of L-aspartate to β-alanine [8]. L-aspartate-α-decarboxylase is a pyruvoyl enzyme that is synthesized as a 14-kDa zymogen, which undergoes self-catalyzed proteolysis to yield active enzyme [9]. The cleavage reaction yields an 11-kDa α-subunit and 3-kDa β-subunit with the concomitant formation of a pyruvoyl moiety at the *N*-terminus of the α-subunit; the pyruvoyl moiety is necessary for L-aspartate decarboxylation [9].

The *Salmonella enterica* L-aspartate-α-decarboxylase zymogen contains all the determinants needed for maturation. However, at 37 °C, the optimal growth temperature of this bacterium, the ancillary protein PanM (formerly YhhK) is required in vitro and in vivo for cleavage of the L-aspartate-α-decarboxylase zymogen [10, 11]. At present, the mechanism by which PanM stimulates L-aspartate-α-decarboxylase cleavage is not known. What is known is that although PanM is a homologue of Gcn5-like *N*-acetyltransferases (GNATs), it lacks acetyltransferase activity [11]. Interestingly, PanM activity is stimulated by acetyl-CoA, a result that led us to hypothesize that PanM functions as an acetyl-CoA sensor to regulate L-aspartate-α-decarboxylase zymogen maturation [11].

Previously, we showed that the L-aspartate-α-decarboxylase zymogen from *Corynebacterium glutamicum* did not require PanM to process its own maturation [10]. This suggested that two classes of bacterial L-aspartate-α-decarboxylases might exist in nature, one class that would require PanM for processing, and a second class that could mature in the absence of PanM. To determine the distribution of these two forms of PanD in the prokaryotes, we performed a phylogenetic analysis of PanM and L-aspartate-α-decarboxylase. By comparing PanM sequences amongst all homologues, we identified a putative domain of interactions between the L-aspartate-α-decarboxylase zymogen and PanM.
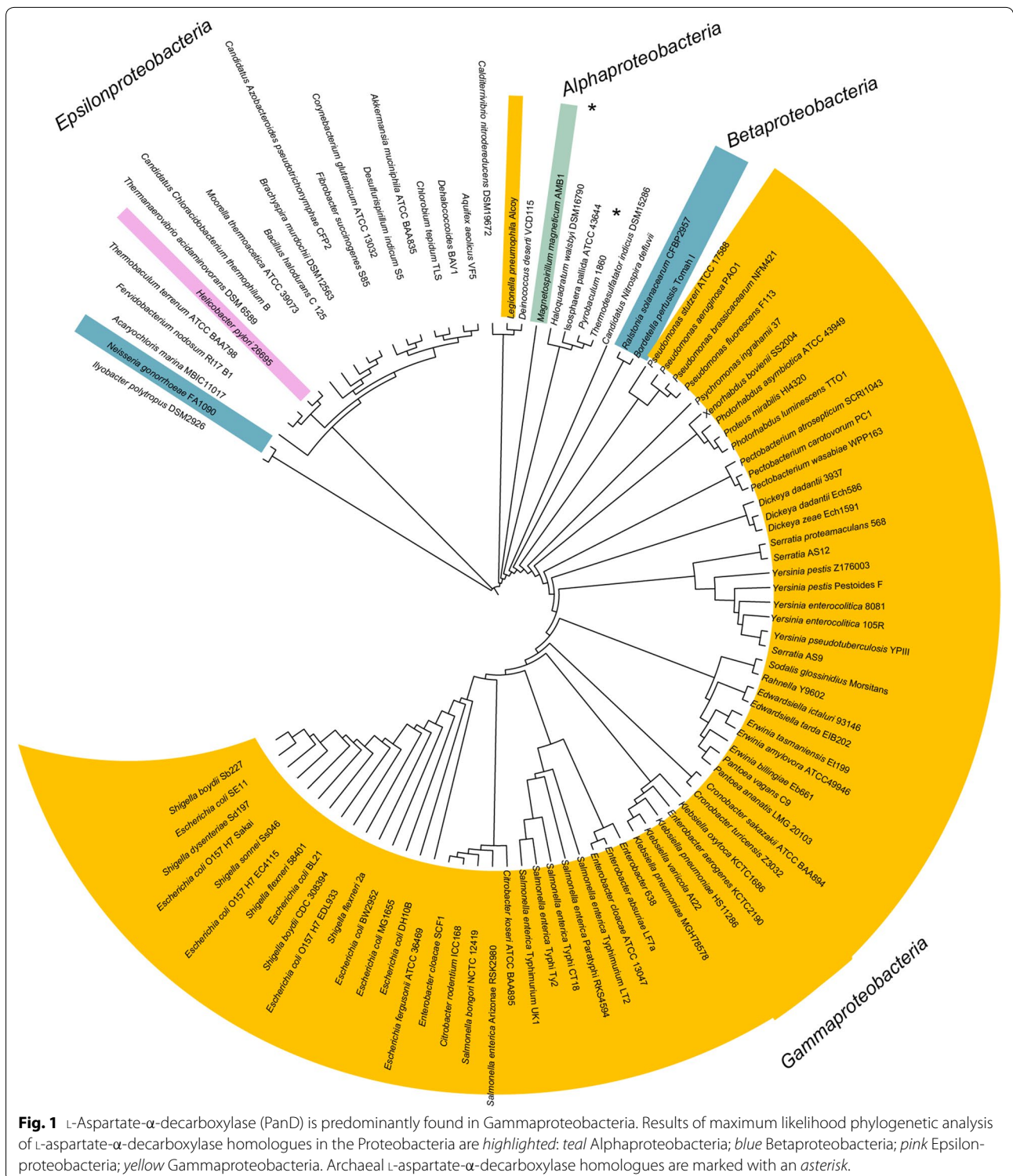
Results and discussion

Phylogenetic distribution of prokaryotic L-aspartate-α-decarboxylase and PanM proteins

We determined the distribution of PanM and L-aspartate-α-decarboxylase proteins by searching the National Center for Biotechnology Information (NCBI) database of completed prokaryotic genomes for homologues of the *S. enterica* PanM and L-aspartate-α-decarboxylase proteins. Homologues of the latter were found in 829 genomes (~52 %) (Additional file 1: Dataset S1). Notably, seven L-aspartate-α-decarboxylase homologues were found in members of the Archaea, a finding that differs from those by Genschel who did not find archaeal L-aspartate-α-decarboxylase homologues [12]. Most likely this difference is due to the increase in sequenced genomes since 2004. In contrast, PanM homologues were far less abundant, with only 128 genomes (~8 %) containing homologues of this protein (Additional file 2: Dataset S2). Importantly, all the PanM-containing genomes belonged to the domain Bacteria.

To gain insights into the evolutionary history of PanM and L-aspartate-α-decarboxylase, phylogenetic trees were constructed using alignments of all homologues for each protein. L-aspartate-α-decarboxylase was widely distributed amongst numerous bacterial phyla and two archaeal phyla (Fig. 1). Interestingly, the phylogeny of L-aspartate-α-decarboxylase did not follow standard 16S phylogenetic relationships. This was observed in the phylum Proteobacteria (Fig. 1), where L-aspartate-α-decarboxylase homologues from Gamma, Alpha, Beta, and Epsilonproteobacteria did not cluster as expected. In light of this information, we posited that horizontal gene transfer might have played a role in L-aspartate-α-decarboxylase evolution. In contrast, homologues of PanM were only found in genomes belonging to the Gammaproteobacteria (Fig. 2). However, not all gammaproteobacterial genomes contained a *panM* gene. Of the 216 gammaproteobacterial genomes encoding an L-aspartate-α-decarboxylase homologue, only 129 (~60 %) also encoded a PanM homologue. All bacterial genomes encoding a PanM homologue also encoded L-aspartate-α-decarboxylase, which supported the physiological role of PanM as a

Stuecker *et al. BMC Res Notes* (2015) 8:354

Page 3 of 11

Fig. 1 L-Aspartate-α-decarboxylase (PanD) is predominantly found in Gammaproteobacteria. Results of maximum likelihood phylogenetic analysis of L-aspartate-α-decarboxylase homologues in the Proteobacteria are *highlighted*: *teal* Alphaproteobacteria; *blue* Betaproteobacteria; *pink* Epsilon-proteobacteria; *yellow* Gammaproteobacteria. Archaeal L-aspartate-α-decarboxylase homologues are marked with an *asterisk*.

maturation factor of the L-aspartate-α-decarboxylase zymogen.

We predict that these L-aspartate-α-decarboxylase homologues require PanM for maturation. Notably, ~77 % of genomes encoding L-aspartate-α-decarboxylase lacked a PanM homologue, suggestive of L-aspartate-α-decarboxylase proteins that self-cleave in the absence of PanM.

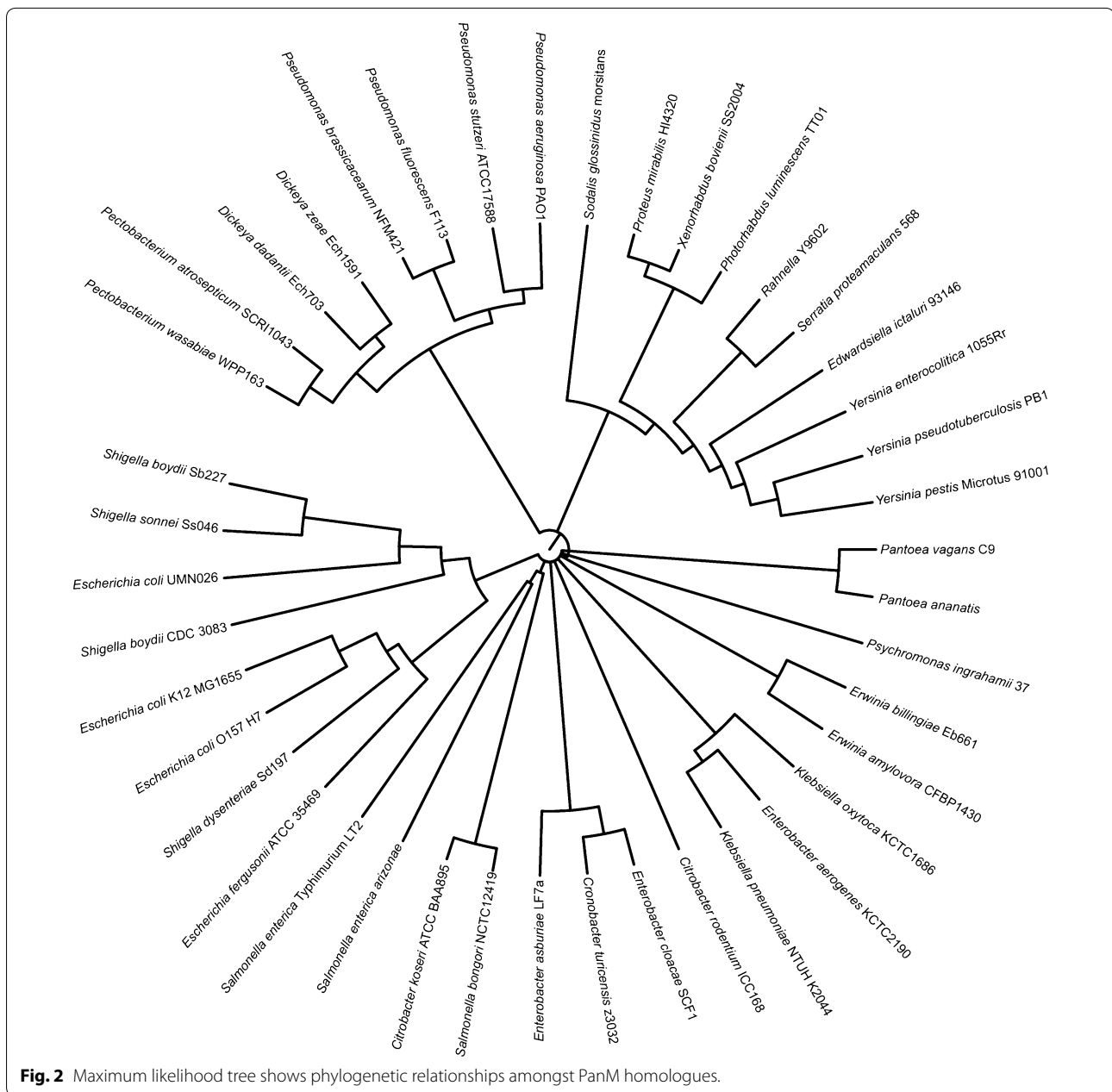Stuecker *et al. BMC Res Notes* (2015) 8:354

Page 4 of 11

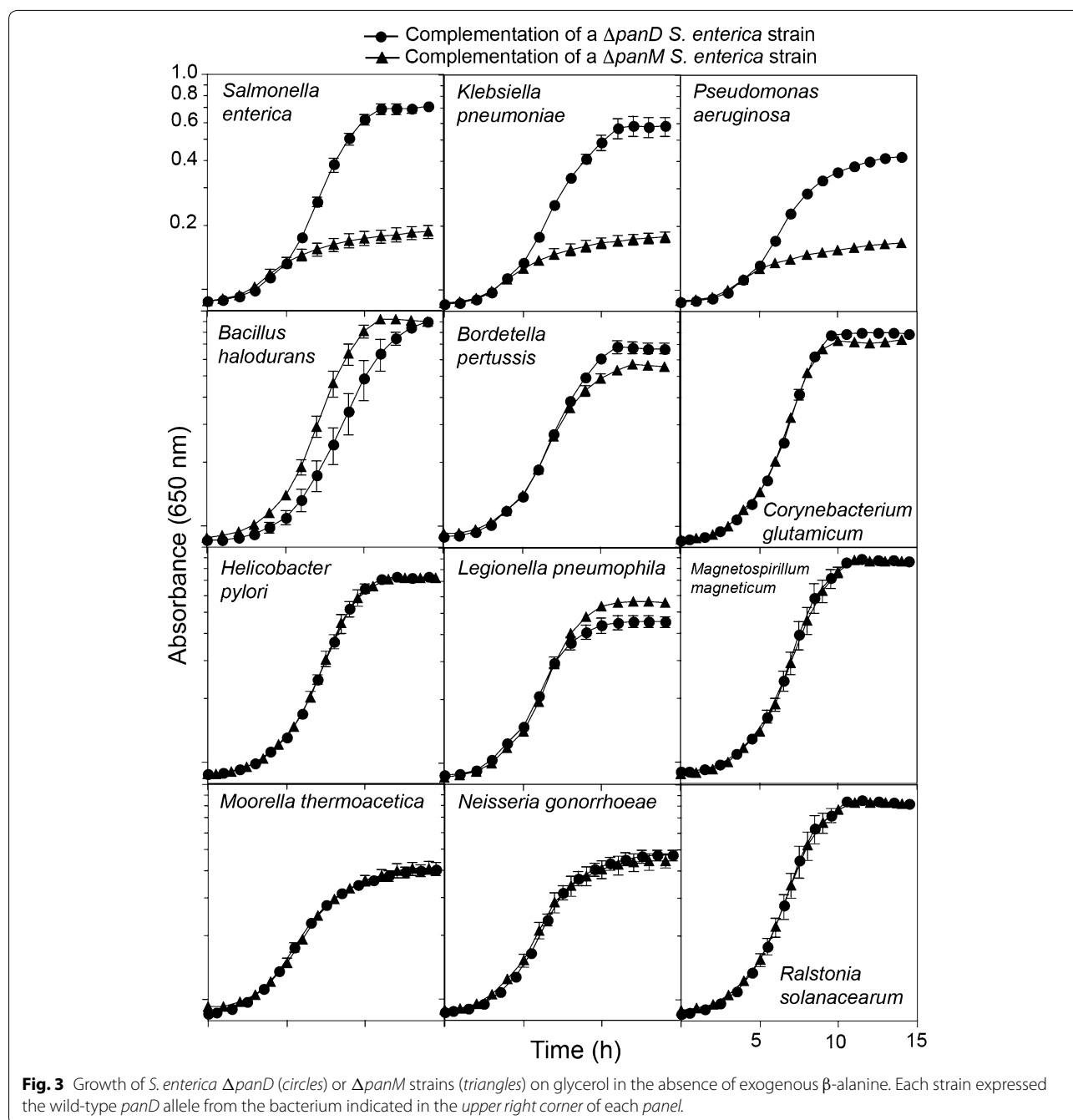**Fig. 2** Maximum likelihood tree shows phylogenetic relationships amongst PanM homologues.

On the basis of these phylogenetic data we predicted the existence of two classes of ʟ-aspartate-α-decarboxylase enzymes, one class that required PanM for maturation, (Class I), and a second class that did not require PanM (Class II).

In vivo validation
To validate the phylogenetic results, *panD* genes were cloned from several bacteria that contained *panM* and several that lacked *panM.* Each *panD* gene was expressed ectopically in a Δ*panD S. enterica* strain grown on minimal medium devoid of β-alanine to verify

ʟ-aspartate-α-decarboxylase function in vivo. To determine whether each ʟ-aspartate-α-decarboxylase could mature in the absence of PanM, the *panD* genes were also expressed in a Δ*panM S. enterica* strain. The *S. enterica panD*⁺ allele was used as control for a *bona fide* ʟ-aspartate-α-decarboxylase that required PanM for maturation [10]. The *C. glutamicum panD*⁺ was included as a control of a gene encoding a ʟ-aspartate-α-decarboxylase that did not require PanM for processing [10] (Fig. 3).

All *panD* homologues restored growth of the *S. enterica* Δ*panD panM*⁺ strain in the absence of pantothenate, indicating that all PanD proteins had

Stuecker *et al. BMC Res Notes* (2015) 8:354

Page 5 of 11

Fig. 3 Growth of *S. enterica* Δ*panD* (*circles*) or Δ*panM* strains (*triangles*) on glycerol in the absence of exogenous β-alanine. Each strain expressed the wild-type *panD* allele from the bacterium indicated in the *upper right corner* of each *panel*.

L-aspartate-α-decarboxylase activity in vivo (Fig. 3). When expressed in the *S. enterica* Δ*panM* strain, all *panD* homologues from bacteria that also contained a *panM* gene (*S. enterica*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*) failed to restore growth on minimal medium (Fig. 3, row 1). These data supported the phylogenetic analysis in assigning these L-aspartate-α-decarboxylase homologues to Class I, which we predicted would require PanM for maturation. In contrast,

panD genes from bacteria that lacked *panM* (e.g., *Bacillus halodurans*, *Bordetella pertussis*, *C. glutamicum*, *Helicobacter pylori*, *Legionella pneumophila*, *Magnetospirillum magneticum*, *Moorella thermoacetica*, *Neisseria gonorrhoeae* and *Ralstonia solanacearum*) restored growth of an *S. enterica* strain carrying a Δ*panM* deletion in the absence of β-alanine or pantothenate (Fig. 3, rows 2–4). These in vivo results combined with the phylogenetic data supported the existence of two
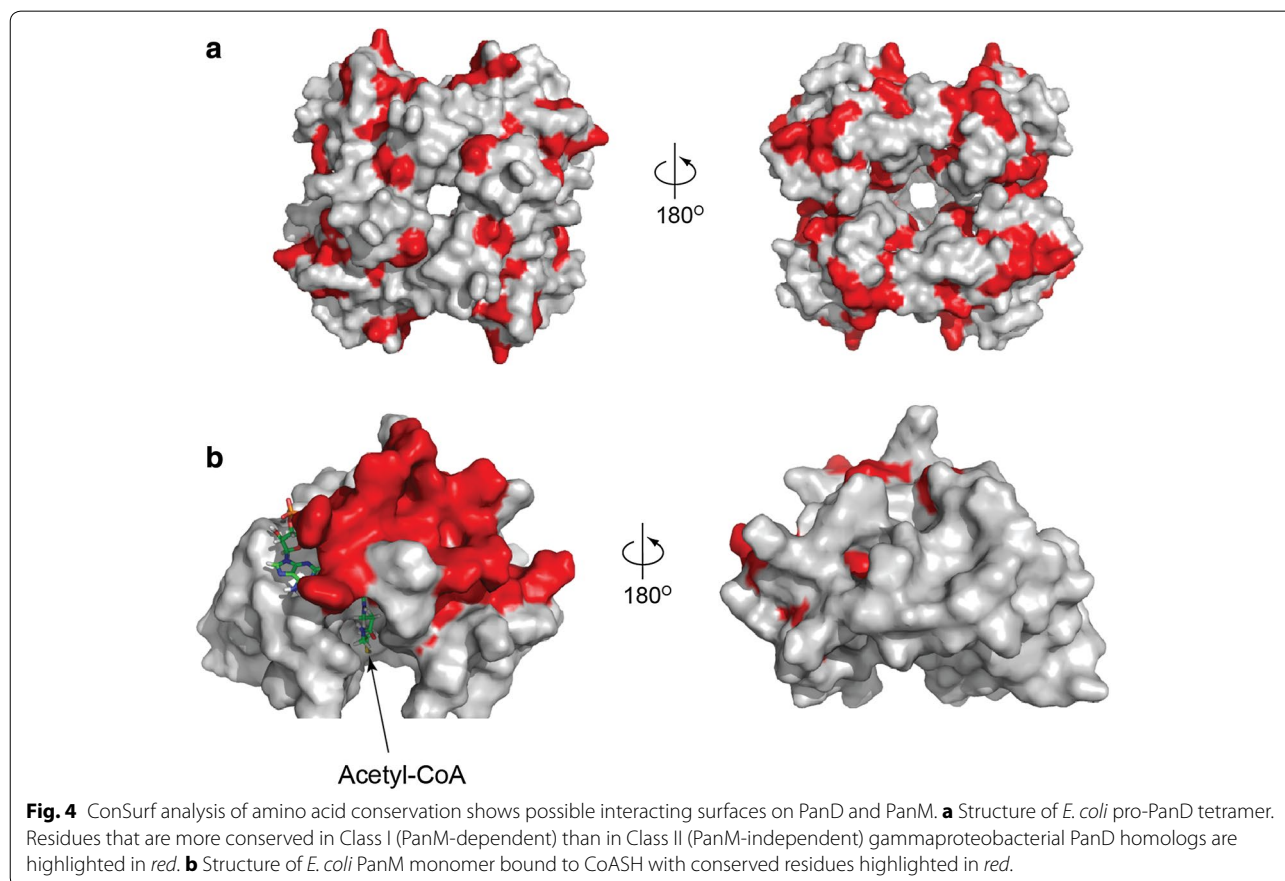
Stuecker *et al. BMC Res Notes* (2015) 8:354

Page 6 of 11

classes of L-aspartate-α-decarboxylase enzymes. Class I L-aspartate-α-decarboxylases that require PanM for activation were present only in the Gammaproteobacteria. Class II L-aspartate-α-decarboxylases that did not require PanM to be active were found in a number of bacterial phyla along with a handful of archaeal species.

Conserved regions of PanM form a domain where putative interactions with L-aspartate-α-decarboxylases may interact

It is known that L-aspartate-α-decarboxylases and PanM interact [10], but the interaction domain has not been identified. We posited that the domain of PanM responsible for interaction with L-aspartate-α-decarboxylase might contain conserved amino acids responsible for the interaction. Given the above data, it followed that L-aspartate-α-decarboxylase residues interacting with PanM should be conserved in the class of homologues requiring PanM for activation (Class I), but not conserved in the class of homologues that did not require PanM (Class II). The ConSurf server [13] was used to calculate evolutionary conservation of amino acids in PanM, and Class I and Class II L-aspartate-α-decarboxylase homologues. Since Class I enzymes were only found in

Gammaproteobacteria, we limited our conservation analysis of Class II L-aspartate-α-decarboxylases to the Gammaproteobacteria. Residues that were more conserved in Class I than Class II L-aspartate-α-decarboxylases were highlighted on the crystal structure of *E. coli* L-aspartate-α-decarboxylase zymogen (PDB 1PPY) [14] (Fig. 4a). Residues conserved in all PanM homologues were highlighted on the nuclear magnetic resonance (NMR) solution structure of *E. coli* YhhK (Cort, J. R., Yee, A., Arrowsmith, C. H. and Kennedy, M. A.; unpublished; PDB 2K5T) (Fig. 4b); the *E. coli* YhhK is now known as PanZ [15].

Several conserved regions were present in L-aspartate-α-decarboxylase (Figs. 4a, 5), so further in vitro and in vivo analyses are necessary to determine which conserved region interacts with PanM. However, only one conserved region was observed for PanM (Fig. 4b). Notably, this region was in a unique location when compared to other Gcn5-like *N*-acetyltransferases (GNATs), the family of proteins to which PanM belongs. In GNATs, the protein substrate binds the acetyltransferase in the same cleft to which acetyl-CoA binds [16]. This allows both substrates to be positioned for transfer of the acetyl moiety from acetyl-CoA to the protein substrate. In PanM,

Fig. 4 ConSurf analysis of amino acid conservation shows possible interacting surfaces on PanD and PanM. **a** Structure of *E. coli* pro-PanD tetramer. Residues that are more conserved in Class I (PanM-dependent) than in Class II (PanM-independent) gammaproteobacterial PanD homologs are highlighted in *red*. **b** Structure of *E. coli* PanM monomer bound to CoASH with conserved residues highlighted in *red*.

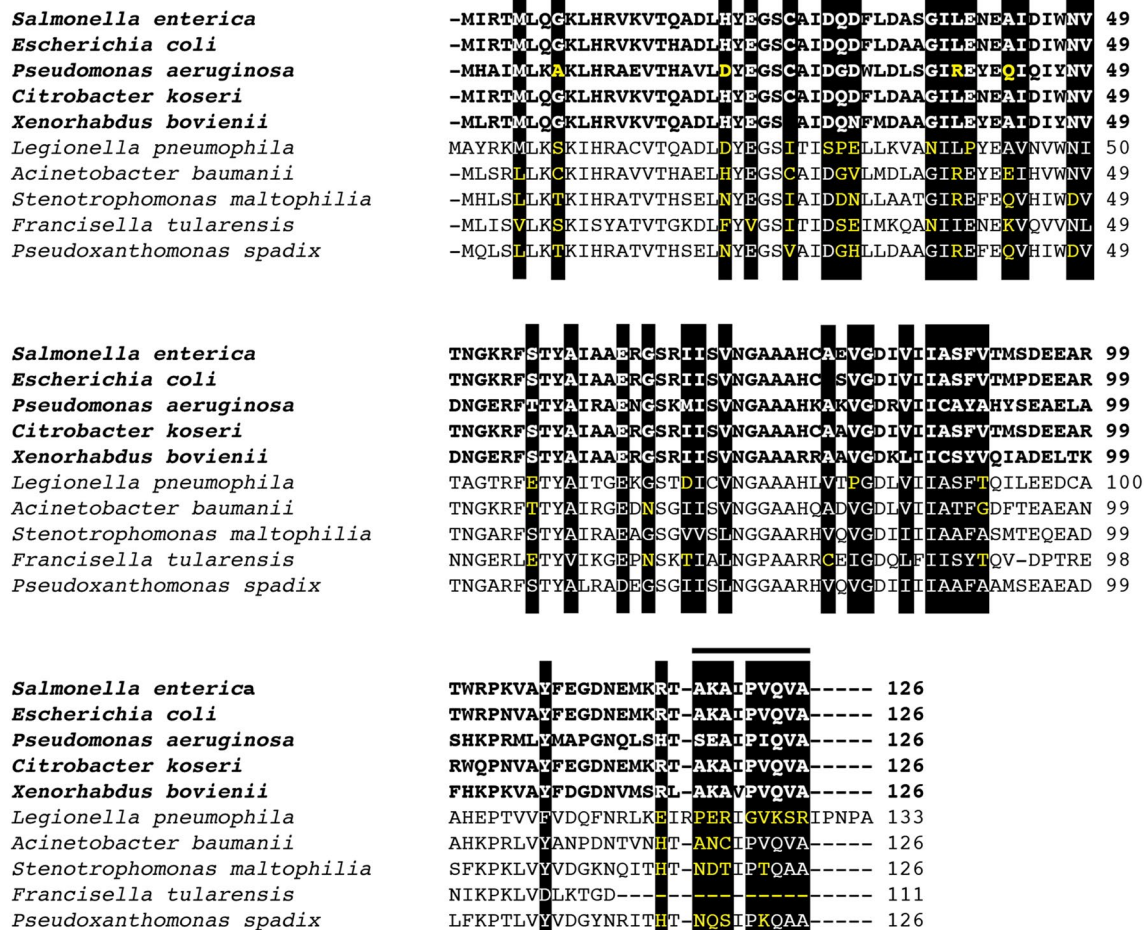Stuecker *et al. BMC Res Notes* (2015) 8:354

Page 7 of 11

Fig. 5 Comparison of PanD protein sequences with residues more conserved in bacteria that synthesize PanM than in bacteria that do not synthesize PanM. Highlighted residues represent potential PanM binding regions. The PanD sequences are shown for five representative bacteria that synthesize PanM (*bold*), and five representative species that do not (*regular text*). Residues that are not conserved are shown in *yellow font*.

the conserved region was located adjacent to the acetyl-CoA binding cleft rather than in the cleft itself (Figs. 4b, 6). If the conserved region of PanM were the PanD interaction domain, PanD would not be positioned for acetyl-transfer, supporting previous studies classifying PanM as an acetyl-CoA sensor [11].

During the review of the work reported herein, Monteiro et al. published the structure of *E. coli* PanD in complex with PanM (PanZ in *E. coli*). This elegant and thorough set of studies revealed, among other things, the interaction domain between the PanD zymogen and PanM [17]. Their analysis established that PanM binds to the C-terminus of PanD. In our study, we found several possible PanM binding regions with higher amino acid conservation among organisms that also contained a PanM homologue when compared with those lacking PanM (Fig. 5). Interestingly, the C-terminal portion

of PanD identified by Monteiro et al. [17] was one of the possible PanM binding regions identified in our study. Specifically, variations in the motif Ala118 to Ala126 identified by the bar in Fig. 5, may be used to predict which PanD zymogens require PanM to expedite self processing.

For PanM (*E.c.* PanZ), Monteiro et al. also demonstrated that many of the conserved residues found in our study (Fig. 6) were important for PanD binding. Specifically, they found the loop formed by residues Leu66-Gly76 to be stabilized upon Ac-CoA binding, and conserved residue Asn45 on PanM to be critical for PanD binding. Both the Leu66-Gly76 loop and Asn45 were predicted to be involved in PanD binding in our study (Fig. 6). This shows the power of using conservation analysis for the prediction of protein–protein interactions.

There are many intriguing questions regarding the evolution of PanM in pro-PanD maturation. For example,
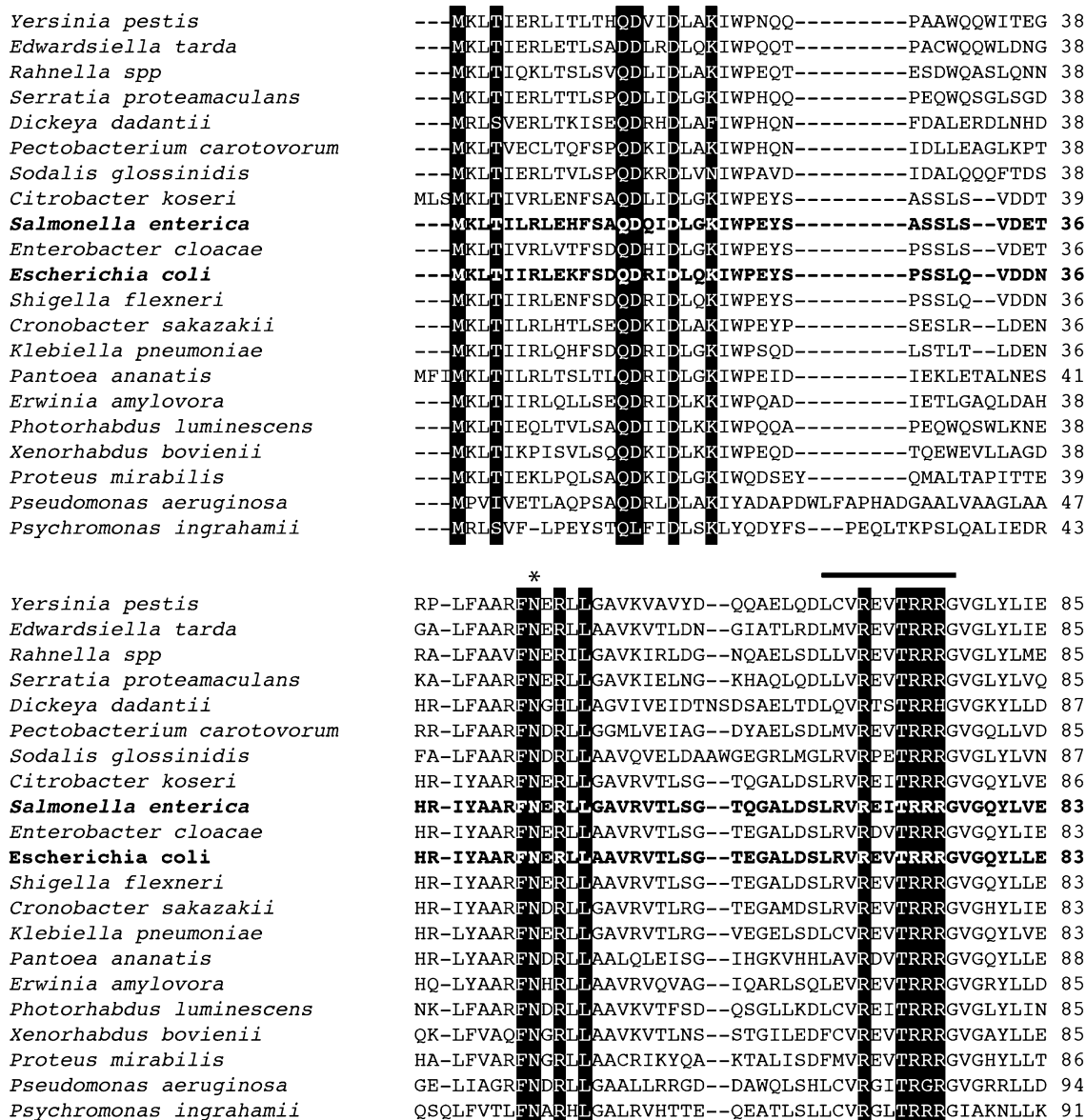
Stuecker *et al. BMC Res Notes* (2015) 8:354

Page 8 of 11

```
Yersinia pestis              ---MKLTIERLITLTHQDVIDLAKIWPNQQ---------PAAWQQWITEG 38
Edwardsiella tarda           ---MKLTIERLETLSADDLRDLQKIWPQQT---------PACWQQWLDNG 38
Rahnella spp                 ---MKLTIQKLTSLSVQDLIDLAKIWPEQT---------ESDWQASLQNN 38
Serratia proteamaculans      ---MKLTIERLTTLSPQDLIDLGKIWPHQQ---------PEQWQSGLSGD 38
Dickeya dadantii             ---MRLSVERLTKISEQDRHDLAFIWPHQN---------FDALERDLNHD 38
Pectobacterium carotovorum   ---MKLTVECLTQFSPQDKIDLAKIWPHQN---------IDLLEAGLKPT 38
Sodalis glossinidis          ---MKLTIERLTVLSPQDKRDLVNIWPAVD---------IDALQQQFTDS 38
Citrobacter koseri           MLSMKLTIVRLENFSAQDLIDLGKIWPEYS---------ASSLS--VDDT 39
Salmonella enterica          ---MKLTILRLEHFSAQDQIDLGKIWPEYS---------ASSLS--VDET 36
Enterobacter cloacae         ---MKLTIVRLVTFSDQDHIDLGKIWPEYS---------PSSLS--VDET 36
Escherichia coli             ---MKLTIIRLEKFSDQDRIDLQKIWPEYS---------PSSLQ--VDDN 36
Shigella flexneri            ---MKLTIIRLENFSDQDRIDLQKIWPEYS---------PSSLQ--VDDN 36
Cronobacter sakazakii        ---MKLTILRLHTLSEQDKIDLAKIWPEYP---------SESLR--LDEN 36
Klebiella pneumoniae         ---MKLTIIRLQHFSDQDRIDLGKIWPSQD---------LSTLT--LDEN 36
Pantoea ananatis             MFIMKLTILRLTSLTLQDRIDLGKIWPEID---------IEKLETALNES 41
Erwinia amylovora            ---MKLTIIRLQLLSEQDRIDLKKIWPQAD---------IETLGAQLDAH 38
Photorhabdus luminescens     ---MKLTIEQLTVLSAQDIIDLKKIWPQQA---------PEQWQSWLKNE 38
Xenorhabdus bovienii         ---MKLTIKPISVLSQQDKIDLKKIWPEQD---------TQEWEVLLAGD 38
Proteus mirabilis            ---MKLTIEKLPQLSAQDKIDLGKIWQDSEY--------QMALTAPITTE 39
Pseudomonas aeruginosa       ---MPVIVETLAQPSAQDRLDLAKIYADAPDWLFAPHADGAALVAAGLAA 47
Psychromonas ingrahamii      ---MRLSVF-LPEYSTQLFIDLSKLYQDYFS---PEQLTKPSLQALIEDR 43

                                          *                      _____
Yersinia pestis              RP-LFAARFNERLLGAVKVAVYD--QQAELQDLCVREVTRRRGVGLYLIE 85
Edwardsiella tarda           GA-LFAARFNERLLAAVKVTLDN--GIATLRDLMVREVTRRRGVGLYLIE 85
Rahnella spp                 RA-LFAAVFNERILGAVKIRLDG--NQAELSDLLVREVTRRRGVGLYLME 85
Serratia proteamaculans      KA-LFAARFNERLLGAVKIELNG--KHAQLQDLLVREVTRRRGVGLYLVQ 85
Dickeya dadantii             HR-LFAARFNGHLLAGVIVEIDTNSDSAELTDLQVRTSTRRHGVGKYLLD 87
Pectobacterium carotovorum   RR-LFAARFNDRLLGGMLVEIAG--DYAELSDLMVREVTRRRGVGQLLVD 85
Sodalis glossinidis          FA-LFAARFNDRLLAAVQVELDAAWGEGRLMGLRVRPETRRRGVGLYLVN 87
Citrobacter koseri           HR-IYAARFNERLLGAVRVTLSG--TQGALDSLRVREITRRRGVGQYLVE 86
Salmonella enterica          HR-IYAARFNERLLGAVRVTLSG--TQGALDSLRVREITRRRGVGQYLVE 83
Enterobacter cloacae         HR-IYAARFNERLLAAVRVTLSG--TEGALDSLRVRDVTRRRGVGQYLIE 83
Escherichia coli             HR-IYAARFNERLLAAVRVTLSG--TEGALDSLRVREVTRRRGVGQYLLE 83
Shigella flexneri            HR-IYAARFNERLLAAVRVTLSG--TEGALDSLRVREVTRRRGVGQYLLE 83
Cronobacter sakazakii        HR-IYAARFNDRLLGAVRVTLRG--TEGAMDSLRVREVTRRRGVGHYLIE 83
Klebiella pneumoniae         HR-LYAARFNERLLGAVRVTLRG--VEGELSDLCVREVTRRRGVGQYLVE 83
Pantoea ananatis             HR-LYAARFNDRLLAALQLEISG--IHGKVHHLAVRDVTRRRGVGQYLLE 88
Erwinia amylovora            HQ-LYAARFNHRLLAAVRVQVAG--IQARLSQLEVREVTRRRGVGRYLLD 85
Photorhabdus luminescens     NK-LFAARFNDRLLAAVKVTFSD--QSGLLKDLCVREITRRRGVGLYLIN 85
Xenorhabdus bovienii         QK-LFVAQFNGRLLAAVKVTLNS--STGILEDFCVREVTRRRGVGAYLLE 85
Proteus mirabilis            HA-LFVARFNGRLLAACRIKYQA--KTALISDFMVREVTRRRGVGHYLLT 86
Pseudomonas aeruginosa       GE-LIAGRFNDRLLGAALLRRGD--DAWQLSHLCVRGITRGRGVGRRLLD 94
Psychromonas ingrahamii      QSQLFVTLFNARHLGALRVHTTE--QEATLSLLCVRGLTRRRGIAKNLLK 91
```

Fig. 6 Alignment of PanM proteins used in ConSuf analyses. Surface exposed residues with high conservation highlighted in *black*. These residues form the predicted PanD binding site. The *asterisk* indicates identifies residue Asn45, which was found to be required for PanM binding to pro-PanD [17]. The *bar* atop the sequence on the *lower panel* spans residues Leu66-Gly76, which was also identified by structural studies to be required for PanZ/PanM:PanD interactions [17].

what was the selective pressure that led to the evolution of PanM? This question is interesting in light of the existence of PanM-independent lineages of self-processing pro-PanD proteins.

In *E. coli*, it is known that the PanD zymogen can self-cleave in the absence of PanM. However, maturation, can only occur at high temperatures, optimally at 50 °C [9]. This dependence on high-temperature is not physiological for neither *E. coli* nor *S. enterica*, since neither bacterium can grow at such a temperature.

In *E. coli* and *S. enterica*, PanD zymogen maturation in the absence of PanM does not occur, or if it does, the amount of mature PanD generated is insufficient to support growth as indicated by the clear pantothenate phenotype of a *S. enterica panM* or PanZ mutant strain [10, 18].

The availability of the three-dimensional structure of the PanZ/PanM:pro-PanD complex will serve as a critical framework within which we can analyze results from experiments aimed at furthering our understanding

Stuecker *et al. BMC Res Notes* (2015) 8:354

Page 9 of 11

of the mechanism of self-cleavage of L-aspartate α-decarboxylases.

Conclusions

In nature there are two classes of self-cleaving zymogens of L-aspartate-α-decarboxylase (PanD) enzymes. In both classes of zymogens, the cleaving event yields an *N*-terminal pyruvoyl moiety that is critical for substrate binding and catalysis. The putative L-aspartate-α-decarboxylase zymogen:PanM interaction region generated by amino acid conservation analysis using available structural models should facilitate the testing and analysis of the mechanism of L-aspartate-α-decarboxylase zymogen self-cleavage in vivo and in vitro. Structures of L-aspartate-α-decarboxylase zymogens that do not require PanM for maturation would be valuable to better our understanding of the differences between both zymogens and to understand what PanM does to the structure of the zymogen to trigger cleavage.

Methods

Phylogenetic analysis

The protein sequences for L-aspartate-α-decarboxylase and PanM were obtained from the complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2 (Accession: NC_003197, accessed: 04/01/2012). Specifically, this corresponded to the GenBank Protein IDs 16763570 (PanD) and 16766851 (PanM). We obtained all protein sequences associated with the complete prokaryotic genome collection found in GenBank (ftp://ncbi.nih.gov/genomes/Bacteria/all.faa.tar.gz, accessed: 04/01/2012) and used the stand alone version of BLAST [19] to format these sequences into a searchable BLAST database. This iteration of the complete prokaryotic genome collection contained a total of 1,606 genome sequences. BLAST was then used to query the L-aspartate-α-decarboxylase and PanM proteins against the complete prokaryotic proteome database and only those alignments with an e value $<1e^{-03}$ were retained.

Phylogenetic trees were then generated as follows. The alignment program MUSCLE [20] was first used to generate an amino acid alignment of all L-aspartate-α-decarboxylase homologues obtained from the complete prokaryotic proteome. This alignment was then imported into the phylogenetic analysis program MEGA [21], and a maximum likelihood (ML) tree was generated. Visualization of this tree was performed using the Interactive Tree of Life web server [22]. For each protein, the originating prokaryotic genus and species was retained throughout the phylogenetic tree construction process. A second tree was also generated using this approach for all homologues of PanM. In both cases, a phylogenetic tree using

maximum parsimony was also constructed, but showed no difference in overall topology from the ML tree.

Plasmid construction

panD genes were amplified from genomic DNA listed in Additional file 3: Table S1 using GeneAmp High Fidelity PCR kit (Applied Biosystems) and cloned into the pBAD24 [23] expression plasmid using EcoRI and HindIII restriction sites. Insert sequences were verified using BigDye® sequencing (Applied Biosystems) at the University of Wisconsin Biotechnology Center (Madison, WI, USA). The resulting plasmids are listed in Additional file 4: Table S2.

Bacterial growth conditions

The Δ*panD* (JE13233) and Δ*panM* (JE12555) *S. enterica* strains [10] were transformed with pBAD24 plasmids expressing *panD* genes (Additional file 4: Table S2). Strains were grown overnight on lysogeny broth (LB) [24, 25] containing 100 μg ml^{-1} ampicillin, then sub-cultured at 0.5 % (v/v) into no-carbon E medium (NCE) [26] containing 20 mM glycerol as the sole source of carbon and energy, and ampicillin (100 μg ml^{-1}). Growth was monitored as an increase in optical density at 650 nm using an ELx808 plate reader (BioTek). All growth curves were performed in triplicate and data presented are averaged from at least two independent experiments. Data were graphed using Prism v4.0 (GraphPad).

Determination of conserved regions in PanD and PanM

PanM protein sequences from bacteria listed in Dataset S2 were aligned using ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalW2). For species where multiple strains were listed, only one representative strain was used in the alignment. Conservation of PanM residues was determined from the alignment using ConSurf (consurf.tau.ac.il) [13]. All residues with a ConSurf conservation score above 7 were highlighted on the NMR solution structure of *Escherichia coli* PanM (Cort, J. R., Yee, A., Arrowsmith, C. H. and Kennedy, M. A.; PDB 2K5T) using Pymol [27]. To predict residues of L-aspartate-α-decarboxylase that may participate in interactions with PanM, two alignments of L-aspartate-α-decarboxylase homologues were created. One alignment contained L-aspartate-α-decarboxylase homologues from Gammaproteobacteria that also contained a *panM* gene (Class I). The second alignment contained L-aspartate-α-decarboxylase homologues from Gammaproteobacteria that lacked *panM* (Class II). Both alignments were generated and conserved residues determined as described above. Alignments with conservation scores were manually compared and residues with higher ConSurf scores in Class I compared to Class II L-aspartate-α-decarboxylases were

Stuecker *et al. BMC Res Notes* (2015) 8:354

Page 10 of 11

highlighted on the crystal structure of *E. coli* L-aspartate-α-decarboxylase zymogen (PDB 1PPY) [14].

Availability of supporting data

All data associated with the phylogenetic trees generated in this study are publicly available in the Dryad Digital Repository at http://dx.doi.org/10.5061/dryad.j9d6q.

Additional files

Additional file 1: Dataset S1. PanD list.

Additional file 2: Dataset S2. PanM list.

Additional file 3: Table S1. Bacteria from which *panD* genes were amplified and cloned.

Additional file 4: Table S2. Strains and plasmids used in this study.

Authors' contributions

TNS, designed and performed the in vivo and in vitro molecular genetics experiments, performed the conservation analysis of pro-PanD and PanM and generated the models; wrote the paper; SB, performed the bioinformatics and phylogenetics analyses; KMH-H, participated in the performance of the in vivo and in vitro molecular genetics experiments; GS, designed the bioinformatics and phylogenetics analyses, helped write the paper; and JCE-S, designed the in vivo experiments; wrote the paper. The funding agencies supporting this work did not contribute to the conception or performance of this work. All authors read and approved the final manuscript.

Authors' information

Members of the Escalante-Semerena research group (TNS, KMH-H, JCE-S) are cell physiologists that were first to describe the protein-mediated maturation of the L-aspartate-α-decarboxylase zymogen in any organism. Their report identified PanM (formerly YhhK) as the protein required for this process. The work reported here are the initial steps taken towards the elucidation of the molecular details of the interactions between PanM and the zymogen that lead to enzyme maturation. Members of the She Suen research group (SB, GS) are expert bioinformaticists interested in broad analysis of distribution of proteins of interest throughout the prokaryotic domains of life.

Author details

[1] Department of Bacteriology, University of Wisconsin-Madison, Madison, WI 53706, USA. [2] Present Address: Microsoft Corporation, 7000 State Highway 161, Irving, TX 75039, USA. [3] Present Address: Department of Microbiology, University of Georgia, Biological Sciences Building, 120 Cedar Street, Athens, GA 30602, USA. [4] Present Address: Department of Biological Sciences, University of Arkansas, 850 W. Dickson St., SCEN 601, Fayettevile, AR 72701, USA. [5] Department of Microbiology, University of Georgia, 212C, Biological Sciences Building, 120 Cedar Street, Athens, GA 30602, USA.

Compliance with ethical guidelines

Competing interests

The authors declare that they have no competing interests.

References

1. Abiko Y (1975) Metabolism of coenzyme A. In: Greenburg DM (ed) Metabolic pathways: metabolism of sulfur compounds. Academic, New York, pp 1–25
2. Lee C-Y, Chen A (1982) Immobilized coenzymes and derivatives. In: You E (ed) The pyridine nucleotide coenzymes. Academic, New York, pp 189–224
3. Wallace BD, Edwards JS, Wallen JR, Moolman WJ, van der Westhuyzen R, Strauss E et al (2012) Turnover-dependent covalent inactivation of Staphylococcus aureus coenzyme A-disulfide reductase by coenzyme A-mimetics: mechanistic and structural insights. Biochemistry 51(39):7699–7711
4. Darnell M, Weidolf L (2013) Metabolism of xenobiotic carboxylic acids: focus on coenzyme A conjugation, reactivity, and interference with lipid metabolism. Chem Res Toxicol 26(8):1139–1155
5. Hunt MC, Tillander V, Alexson SE (2014) Regulation of peroxisomal lipid metabolism: the role of acyl-CoA and coenzyme A metabolizing enzymes. Biochimie 98:45–55
6. Smith CM, Song WO (1996) Comparative nutrition of pantothenic acid. J Nutrition Biochem 7:312–321
7. Cronan JE Jr, Littel KJ, Jackowski S (1982) Genetic and biochemical analyses of pantothenate biosynthesis in *Escherichia coli* and *Salmonella typhimurium*. J Bacteriol 149:916–922
8. Cronan J Jr (1980) Beta-alanine synthesis in *Escherichia coli*. J Bacteriol 141:1291–1297
9. Ramjee MK, Genschel U, Abell C, Smith AG (1997) *Escherichia coli* L-aspartate-alpha-decarboxylase: preprotein processing and observation of reaction intermediates by electrospray mass spectrometry. Biochem J 323:661–669
10. Stuecker TN, Hodge KM, Escalante-Semerena JC (2012) The missing link in coenzyme A biosynthesis: PanM (formerly YhhK), a yeast GCN5 acetyltransferase homologue triggers aspartate decarboxylase (PanD) maturation in *Salmonella enterica*. Mol Microbiol 84:608–619
11. Stuecker TN, Tucker AC, Escalante-Semerena JC (2012) PanM, an acetyl-coenzyme A sensor required for maturation of L-aspartate decarboxylase (PanD). MBio 3:e00158–e00112
12. Genschel U (2004) Coenzyme A biosynthesis: reconstruction of the pathway in archaea and an evolutionary scenario based on comparative genomics. Mol Biol Evol 21:1242–1251
13. Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N (2010) ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic Acids Res 38(Web Server issue):W529–W533
14. Schmitzberger F, Kilkenny ML, Lobley CM, Webb ME, Vinkovic M, Matak-Vinkovic D et al (2003) Structural constraints on protein self-processing in L-aspartate-alpha-decarboxylase. EMBO J 22:6193–6204
15. Nozaki S, Webb ME, Niki H (2012) An activator for pyruvoyl-dependent L-aspartate alpha-decarboxylase is conserved in a small group of the gamma-proteobacteria including *Escherichia coli*. MicrobiologyOpen 1:298–310
16. Vetting MW, Carvalho LPSD, Yu M, Hegde SS, Magnet S, Roderick SL et al (2005) Structure and functions of the GNAT superfamily of acetyltransferases. Arch Biochem Biophys 433:212–226
17. Monteiro DC, Patel V, Bartlett CP, Nozaki S, Grant TD, Gowdy JA et al (2015) The structure of the PanD/PanZ protein complex reveals negative feedback regulation of pantothenate biosynthesis by coenzyme A. Chem Biol 22:492–503
18. Adams MD, Wagner LM, Graddis TJ, Landick R, Antonucci TK, Gibson AL et al (1990) Nucleotide sequence and genetic characterization reveal six essential genes for the LIV-I and LS transport systems of *Escherichia coli*. J Biol Chem 265:11436–11443
19. Altschul SF, Gish W, Miller W, Myers EW (1990) Basic local alignment search tool. J Mol Biol 215:403–410
20. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797
21. Kumar S, Nei M, Dudley J, Tamura K (2008) MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief Bioinform 9:299–306
22. Letunic I, Bork P (2007) Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. Bioinformatics 23:127–128

Stuecker *et al. BMC Res Notes* (2015) 8:354

Page 11 of 11

23. Guzman LM, Belin D, Carson MJ, Beckwith J (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J Bacteriol 177:4121–4130
24. Bertani G (1951) Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. J Bacteriol 62:293–300
25. Bertani G (2004) Lysogeny at mid-twentieth century: P1, P2, and other experimental systems. J Bacteriol 186:595–600
26. Berkowitz D, Hushon JM, Whitfield HJ Jr, Roth J, Ames BN (1968) Procedure for identifying nonsense mutations. J Bacteriol 96:215–220
27. DeLano WL (2002) The pymol molecular graphics system. Schrodinger, LLC