

Structure and evolution of Alanine/Serine Decarboxylases and the engineering of theanine production

Reviewed Preprint

Published from the original preprint after peer review and assessment by eLife.

[About eLife's process](#)

Reviewed preprint posted

November 13, 2023 (this version)

Sent for peer review

September 8, 2023

Posted to bioRxiv

September 4, 2023

Hao Wang, Biying Zhu, Siming Qiao, Chunxia Dong, Xiaochun Wan, Weimin Gong , Zhaoliang Zhang 

Department of Life Sciences and Medicine, University of Science and Technology of China, Hefei, China • State Key Laboratory of Tea Plant Biology and Utilization, Anhui Agricultural University, Hefei, Anhui 230036, China

 https://en.wikipedia.org/wiki/Open_access

 Copyright information

Abstract

Ethylamine (EA), the precursor of theanine biosynthesis, is synthesized from alanine decarboxylation by Alanine Decarboxylase (AlaDC) in tea plants. AlaDC evolves from Serine Decarboxylase (SerDC) through neofunctionalization and has lower catalytic activity. However, lacking structure information hinders the understanding of the evolution of substrate specificity and catalytic activity. In this study, we solved the X-ray crystal structures of AlaDC from *Camellia sinensis* (CsAlaDC) and SerDC from *Arabidopsis thaliana* (AtSerDC). Tyr³⁴¹ of AtSerDC or the corresponding Tyr³³⁶ of CsAlaDC is essential for their enzymatic activity. Tyr¹¹¹ of AtSerDC and the corresponding Phe¹⁰⁶ of CsAlaDC determine their substrate specificity. Both CsAlaDC and AtSerDC have a distinctive zinc finger that contributes to the activity and has not been identified in any other Group II PLP-dependent amino acid decarboxylases. Based on the structure comparison and evolution analysis AlaDC or SerDC led us to perform mutation screen of CsAlaDC. The results indicated that the mutation of L110F or P114A in the CsAlaDC dimerization interface significantly improved the catalytic activity by 110% and 59%, respectively. Combining double mutant CsAlaDC^{L110F/P114A} with theanine synthetase increased 672% of theanine production in an *in vitro* system. This study provides the structural basis for the substrate selectivity and catalytic activity of CsAlaDC and AtSerDC and sheds light on a new direction for the efficient industrial synthesis of theanine.

eLife assessment

This study reports biochemical and structural analysis of two PLP decarboxylase enzymes from plants. The findings are **useful** due to the utility of these enzymes in industrial theanine production. While certain aspects of the study are **solid**, other components elucidating the role of a Zn(II)-binding motif are **incomplete**. In addition, some of the finding could be presented more clearly, including the connections between the structural findings and the reaction mechanism. The work will be of interest to enzymologists studying PLP enzymes and those interested in enzyme engineering in plants.

Introduction

Tea is one of the most popular natural non-alcoholic beverages consumed worldwide. About 2 billion cups of tea are consumed worldwide daily (1). The popularity of tea is determined by its favorable flavor and numerous health benefits (2–4). The flavor and health-beneficial effects of tea are conferred by the abundant secondary metabolites, including catechins, caffeine, theanine, volatiles, etc (5). Theanine (γ -glutamylethylamide) is a unique non-protein amino acid and the most abundant free amino acid in tea plants (*Camellia sinensis*). It accounts for more than 50% of the total free amino acids and approximately 1–2% of the dry weight of the new shoots of tea plants (6). Theanine is the secondary metabolite conferring the umami taste of tea infusion and also balances the astringency and bitterness of tea infusion caused by catechins and caffeine (7). It has also many health-promoting functions, such as reducing stress and anxiety, improving mood and cognition, protecting the nervous system, etc (6, 8–13). Therefore, theanine content is highly correlated with green tea quality (14).

Theanine is synthesized from EA and glutamate (Glu) by theanine synthetase (15, 16). Importantly, the large amount of theanine biosynthesis is determined by the high availability of EA in tea plants (17). Therefore, the evolution of EA biosynthesis in tea plants largely determined the quality formation of tea. EA is synthesized from alanine decarboxylation by alanine decarboxylase (CsAlaDC) (18) (**Figure 1A**). Indeed, CsAlaDC expression level and catalytic activity largely determine the theanine accumulation level in tea plants (19). As a novel gene, CsAlaDC had not been reported in other plant species before it was identified in tea plants. Previous studies indicated that AlaDC originated from the serine decarboxylase gene (*SerDC*) by gene duplication in tea plants (20). However, CsAlaDC and CsSerDC specifically catalyze alanine and serine decarboxylation, respectively, despite they share highly conserved amino acid sequences (20) (**Figure 1A**, 1C). In addition, CsAlaDC exhibited a much lower enzymatic activity compared with CsSerDC. However, the structural basis and key amino acids underlying the evolution of the substrate specificity and enzymatic activity of CsAlaDC are unknown.

SerDC belongs to the Group II pyridoxal-5-phosphate (PLP)-dependent decarboxylase superfamily (21). These Group II amino acid decarboxylases produce many bioactive secondary metabolites and signaling molecules (22–24). In plants, SerDC catalyzes the biosynthesis of ethanolamine from serine, which was first functionally characterized in *Arabidopsis thaliana* (25) (**Figure 1A**). Ethanolamine is an important metabolite in the synthesis of phosphatidylethanolamine and phosphatidylcholine, two main phospholipids involved in maintaining eukaryotic membrane structures in eukaryotic cell membranes (26–28). Furthermore, the study has shown that SerDC is essential in the embryogenesis of *Arabidopsis* (29). SerDC is of vital importance for plant growth and development, but the mechanism of substrate recognition and catalytic activity of SerDC also remains unclear.

In this work, we attempt to understand the mechanism of functional divergence of plant AlaDC and SerDC by analyzing their structural characteristics. Here, we obtained the X-ray crystal structure of CsAlaDC and AtSerDC. According to the crystal structure, we found a distinctive zinc finger structure located at both CsAlaDC and AtSerDC, which has not been identified in any other Group II PLP-dependent amino acid decarboxylases that have been previously characterized. By comparing the substrate binding pockets, we identified Phe¹⁰⁶ of CsAlaDC and Tyr¹¹¹ of AtSerDC as the crucial sites for substrate specificity. By conducting mutation screening based on the protein structures, we identified the amino acids repressing the catalytic activity and discovered that CsAlaDC^{L110F/P114A} exhibited a 2.3-fold increase in catalytic activity compared to that of CsAlaDC and enhanced the engineering of theanine production *in vitro*.

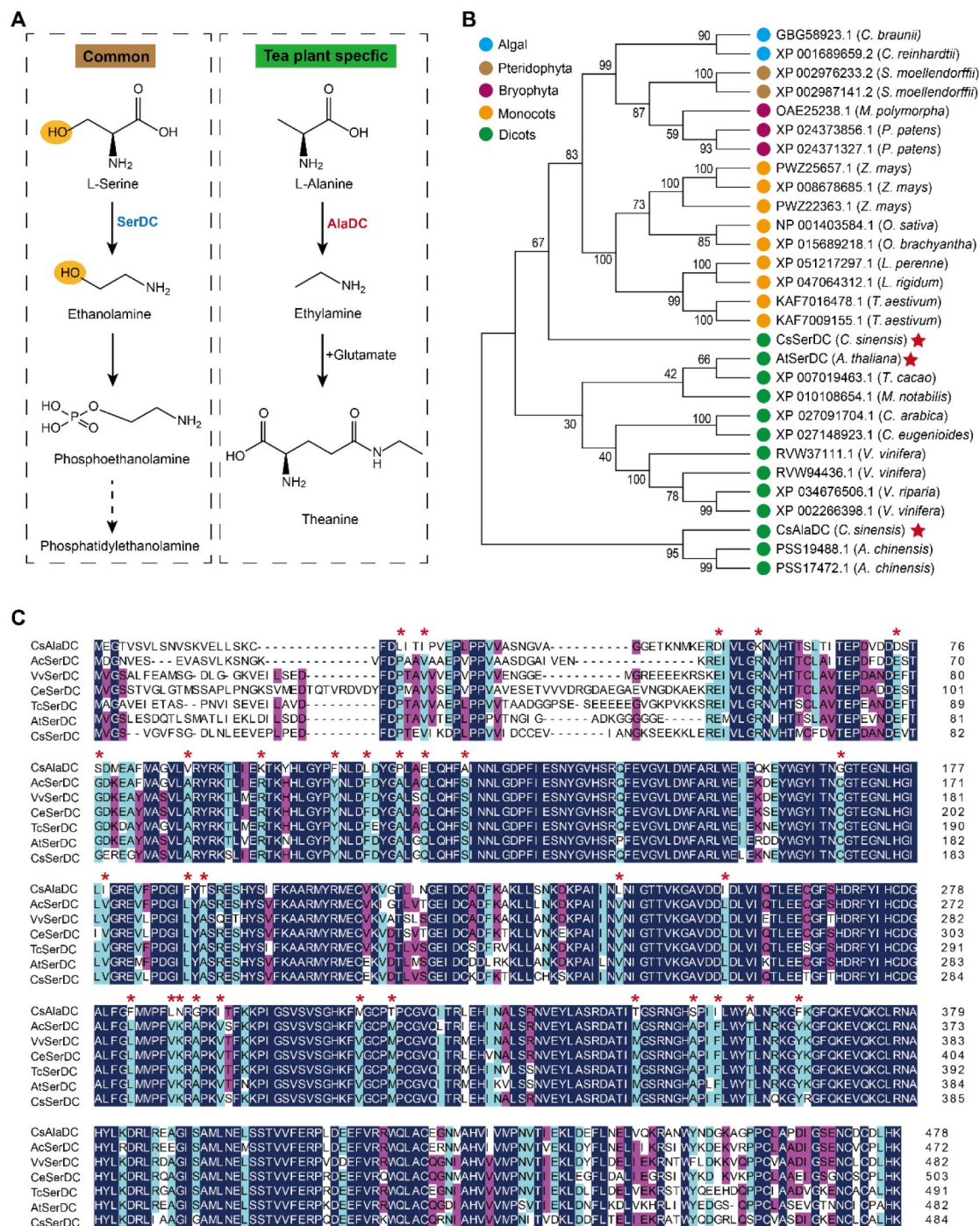


Figure 1.

Metabolic pathways and phylogeny analysis of SerDC and AlaDC in plants.

(A) The decarboxylation of serine and alanine in plants. SerDC, serine decarboxylase; AlaDC, alanine decarboxylase. (B) Phylogenetic analysis of the SerDC homologous proteins in Algal, Pteridophyta, Bryophyta, Dicots and Monocots. The phylogenetic tree was performed using the neighbor-joining method in MEGA 7.0 software and bootstrap values based on 1000 replicates were displayed at branch points. These proteins were represented by their GenBank accession number. (C) Multiple alignment of the amino acid sequences of AlaDC and SerDCs. Primary (100%), secondary (80%), and tertiary (60%) conserved percent of similar amino acid residues were shaded in deep blue, light blue and cheer red, respectively. The multiple alignment was constructed by MEGA 7.0 software by ClustalW method, and GeneDoc software was used to visualize multiple sequence alignments. "*" indicated amino acid residues mutated only in CsAlaDC.

Enzymatic properties of CsAlaDC, AtSerDC, and CsSerDC

CsAlaDC originates from CsSerDC by gene duplication and neofunctionalization in tea plants, but they catalyze different metabolic processes (**Figure 1A** [↗](#)). A phylogenetic tree was constructed to further reveal the evolutionary relationship between CsAlaDC and its homologs in *Algal*, *Pteridophyta*, *Bryophyta*, *Monocots* and *Dicots* (**Figure 1B** [↗](#)). Phylogenetic analysis indicated that CsAlaDC is homologous with SerDCs, and its closer relationship was Dicots plants. As expected, CsSerDC was most closed to AtSerDC, which implies that they shared similar functions. However, CsAlaDC is relatively distant from CsSerDC. Next, we performed a multiple sequence alignment of the amino acid sequences of the 6 SerDCs and CsAlaDC (**Figure 1C** [↗](#)). The results showed that the amino acid sequences of CsAlaDC and SerDCs were highly conserved, but CsAlaDC has amino acid mutations in conserved motifs compared with other SerDCs.

Next, to verify the substrate specificity and enzyme activity, we conducted enzyme activity detection and enzyme kinetics analysis for CsAlaDC, CsSerDC and AtSerDC. The 5' truncated *CsAlaDC* and *AtSerDC* were constructed into the pET22b and pET28a expression vectors to generate recombinant plasmids pET22b-*CsAlaDC* and pET28a-*AtSerDC*, correspondingly. Additionally, the full-length protein of CsSerDC was inserted into the pET28a expression vector to generate recombinant plasmid pET28a-*CsSerDC*. Subsequently, the recombinant proteins were expressed in *Escherichia coli* and purified via nickel affinity chromatography. The purified proteins were examined through SDS-PAGE analysis (**Figure 2A** [↗](#)). To investigate the substrate specificity of CsAlaDC, AtSerDC, and CsSerDC, enzyme activity assays were conducted using Ala and Ser as substrates, followed by product identification via UPLC analysis. The results showed that CsAlaDC can effectively participate in the decarboxylation of alanine to generate EA, whereas its catalytic efficacy towards serine is relatively inferior. Conversely, AtSerDC and CsSerDC selectively catalyzed serine decarboxylation to yield ethanolamine but did not exhibit activity towards alanine (**Figure 2B** [↗](#)). These findings confirmed that CsAlaDC and SerDCs do not have promiscuous decarboxylase activity on Ala and Ser.

The kinetic properties of CsAlaDC, AtSerDC, and CsSerDC were determined through the use of corresponding substrates (**Figure 2C** [↗](#)), and the results are presented in **Table 1** [↗](#). The kinetic parameter K_m of CsAlaDC was determined to be 1.215 mmol/L, which is similar to that of AtSerDC and CsSerDC at 1.522 mmol/L and 2.364 mmol/L, respectively. However, AtSerDC has a V_{max} of 7.053 $\mu\text{mol/L/s}$, which is 4-fold that of CsAlaDC's V_{max} of 1.709 $\mu\text{mol/L/s}$; while CsSerDC has a V_{max} of 9.031 $\mu\text{mol/L/s}$, which is 5.3-fold that of CsAlaDC. These findings indicated that the catalytic efficiency of CsAlaDC is considerably lower than that of both CsSerDC and AtSerDC. Due to the similarity in catalytic activity between CsSerDC and AtSerDC, we chose the more representative AtSerDC for further analysis.

The overall structures of CsAlaDC and AtSerDC

To enhance our comprehension of CsAlaDC and AtSerDC, we conducted structural analyses of these two proteins. Through optimization of crystallization conditions, we successfully determined the crystal structures of CsAlaDC, as well as the CsAlaDC-EA, and AtSerDC at resolutions of 2.50 Å, 2.60 Å and 2.85 Å, respectively (**Supplementary table 1** [↗](#)).

The overall architecture of CsAlaDC and AtSerDC comprises homodimeric structures with two subunits exhibiting an asymmetrical arrangement (**Figure 3A** [↗](#) and **3D** [↗](#)). The monomer of CsAlaDC and AtSerDC is divided into three distinct structural domains: an N-terminal domain (N-terminal -104 aa in CsAlaDC, N-terminal -109 aa in AtSerDC), a large domain (105-364 aa in CsAlaDC, 110-369 aa in AtSerDC), and a small C-terminal domain (365 aa-C-terminus in CsAlaDC, 370-C-terminus as in AtSerDC), colored light pink, khaki, and sky blue, respectively. Compared to

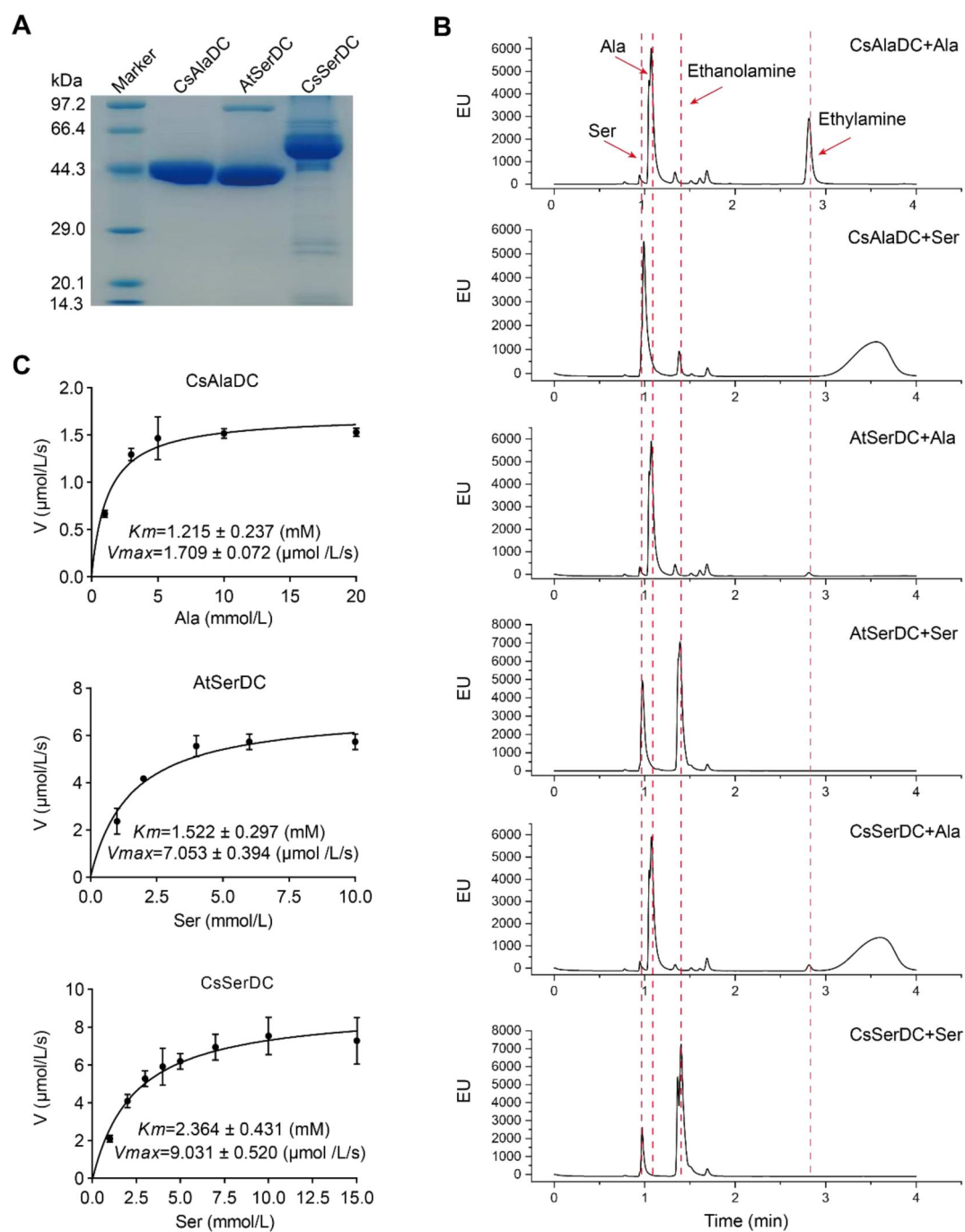


Figure 2.

Purification and characterization of CsAlaDC, AtSerDC, and CsSerDC.

(A) Identification of CsAlaDC, AtSerDC, and CsSerDC using SDS-PAGE. (B) Detection of enzyme activities of CsAlaDC, AtSerDC, and CsSerDC by UPLC. (C) Reaction rates of substrates with different concentrations catalyzed by CsAlaDC, AtSerDC, and CsSerDC.

Enzyme	Substrate	K_m (mmol/L)	V_{max} ($\mu\text{mol/L/s}$)
CsAlaDC	Alanine	1.215 (0.802, 1.766)	1.709 (1.568, 1.865)
AtSerDC	Serine	1.522 (1.010, 2.234)	7.053 (6.311, 7.945)
CsSerDC	Serine	2.364 (1.634, 3.351)	9.031 (8.087, 10.160)

The values enclosed in parentheses denote the 95% confidence intervals for the corresponding parameters.

Table 1.

Kinetic parameters of CsAlaDC, AtSerDC and CsSerDC

the full-length protein, the N-terminal structures of CsAlaDC and AtSerDC are truncated by 60 and 65 amino acid residues, respectively. The truncated protein N-terminus contains a long α -helix, which is connected to the large domain through a long loop. The large domain contains a seven-stranded mixed β -sheet surrounded by eight α -helices, while the C-terminal domain composes of a three-stranded antiparallel β -sheet and three α -helices. The catalytic site of the enzyme is positioned within a superficial crevice at the junction of two subunits forming a dimeric arrangement. The cofactor-binding process involves amino acid residues derived from both subunits. One monomer accommodates both PLP and substrate, while the other monomer also contributes to cofactor binding and enzymatic function (**Figure 3A** and **3D**).

Notably, our investigation has revealed the presence of a distinctive zinc finger structure (as depicted in **Figure 3A** and **3D**) located at the C-terminus of CsAlaDC and AtSerDC. This structure is composed of a ring structure spanning 17 amino acid residues, wherein coordination of Zn^{2+} is facilitated by three Cys residues and one His residue. Importantly, this particular configuration is exclusive to the two proteins under examination and has not been identified in any other Group II PLP-dependent amino acid decarboxylases that have been previously characterized.

The crystal structures of CsAlaDC-EA complex and AtSerDC were obtained and further analyzed. In the former complex, PLP bound to Lys³⁰⁹ via a Schiff base linkage (internal aldimine form), and the pyridine moiety of PLP is positioned between the imidazole ring of His¹⁹⁶ and the methyl group of Ala²⁷⁹ in a parallel orientation to the imidazole ring. The N1 of PLP establishes a salt bridge interaction with Asp²⁷⁷, while the O3 forms hydrogen bonds with Thr²⁴⁷ and Lys³⁰⁹. Additionally, the phosphate group of PLP participates in hydrogen bonding interactions with His³⁰⁸, Gly¹⁶⁹, Thr¹⁷⁰, Lys³⁰⁹, and Ser^{347*} (*** denotes the amino acids on adjacent subunits.) (**Figure 3C**). Similarly, in the latter structure, the pyridine ring of PLP was sandwiched between the imidazole ring of His²⁰¹ and the methyl group of Ala²⁸⁴. The N1 of PLP formed a salt bridge with Asp²⁸², while the O3 formed hydrogen bonds with Thr²⁵². The phosphate group of PLP established hydrogen bonding interactions with Lys³¹⁴, His³¹³, Gly¹⁷⁴, Thr¹⁷⁵, and Ser^{352*} (**Figure 3E**).

Factors affecting CsAlaDC and AtSerDC activity

The Group II PLP-dependent amino acid decarboxylases are dimeric enzymes with each monomer containing an active site located within a shallow cavity (30). A flexible loop of amino acids originating from one monomer extends into the active site of the other monomer and plays a crucial role in catalysis. The catalytic loop in CsAlaDC, consisting of amino acid residues 328-341, is well-structured and exhibits clear electron density. However, the corresponding loop in AtSerDC is disordered (**Figure 3F**). The aforementioned loop harbors a conserved Tyr residue that is believed to function as a proton donor of the carbanion in the catalytic process (31). In particular, CsAlaDC Tyr³³⁶ and AtSerDC Tyr³⁴¹ correspond to that Tyr residue. Upon superimposing the active site residues of CsAlaDC apo and CsAlaDC-EA complex, it was observed that the hydroxyl group of Tyr^{336*} in the CsAlaDC-EA complex exhibits a substantial 60-degree deviation along the C α -C β bond (**Figure 3G**). To assess the function of the Tyr, we substituted the corresponding Tyr residues in CsAlaDC and AtSerDC with Phe. The resulting CsAlaDC^{Y336F} and AtSerDC^{Y341F} mutants were then exposed to L-alanine and L-serine, respectively. We measured the production of EA or ethanolamine in the reaction mixtures. Our findings indicate that these mutants catalyze abortive decarboxylation, as opposed to wild-type CsAlaDC and AtSerDC, as evidenced by the absence of detectable EA and only a small amount of ethanolamine observed in the reaction mixture (**Figure 3H**). This result suggested this Tyr is required for the catalytic activity of CsAlaDC and AtSerDC.

CsAlaDC and AtSerDC both possess a C-terminal zinc finger motif whose removal led to protein insoluble when expressed in *E. coli*. Therefore we used dithiothreitol (DTT) to block the zinc finger structure to evaluate the role of this zinc finger structure in the catalytic activity. The results

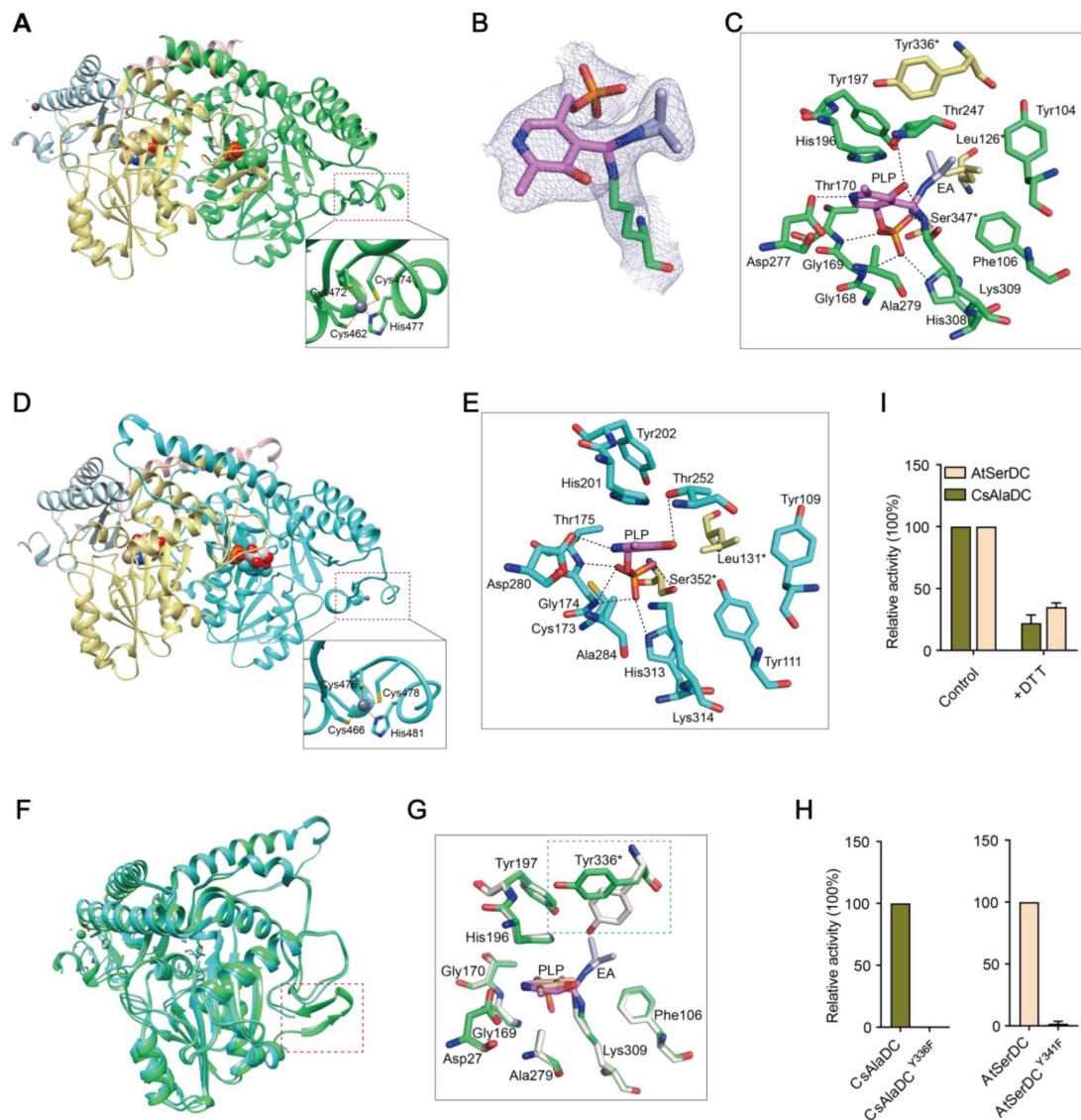


Figure 3.

Crystal structures of CsAlaDC and AtSerDC.

(A) Dimer structure of CsAlaDC. The color display of the N-terminal domain, large domain, and C-terminal domains of chain A is shown in light pink, khaki and sky blue, respectively. Chain B is shown in spring green. The PLP molecule is shown as a sphere model. The zinc finger structure at the C-terminus of CsAlaDC is indicated by the red box. The gray spheres represent zinc ions, while the red dotted line depicts the coordination bonds formed by zinc ions with cysteine and histidine. (B) The 2Fo-Fc electron density maps of K309-PLP-EA (contoured at 1σ level). The PLP is shown in violet, the K309 is shown in spring green, and the EA is shown in lightblue. (C) Active center of the CsAlaDC-EA complex, with hydrogen bonds denoted by black dotted lines. "*" denotes the amino acids on adjacent subunits. (D) Dimer structure of AtSerDC. The color display of the N-terminal domain, large domain, and C-terminal domains of chain A is shown in light pink, khaki and sky blue, respectively. Chain B is shown in cyan. The PLP molecule is shown as a sphere model. The zinc finger structure at the C-terminus of AtSerDC is indicated by the red box. The gray spheres represent zinc ions, while the red dotted line depicts the coordination bonds formed by zinc ions with cysteine and histidine. (E) Active center of the AtSerDC, with hydrogen bonds denoted by black dotted lines. "*" denotes the amino acids on adjacent subunits. (F) The monomer of CsAlaDC and AtSerDC are superimposed. CsAlaDC is depicted in spring green, while AtSerDC is shown in cyan. The conserved amino acid catalytic ring is indicated by the red box (G) Amino acid residues of the active center in CsAlaDC apo and CsAlaDC-EA complex are superimposed. CsAlaDC apo is shown in floral white, while CsAlaDC-EA complex is shown in spring green. (H) The relative activity of wild-type CsAlaDC and its Y336F mutant (left), as well as wild-type AtSerDC and its Y341F mutant (right) is shown. (I) The effects of DTT on the activity of CsAlaDC and AtSerDC are depicted. Three independent experiments were conducted.

showed that 5 mM L-DTT reduced the relative activity of CsAlaDC and AtSerDC to 22.0% and 35.2%, respectively, compared to the control (**Figure 3I**). These results suggested that this characteristic zinc finger motif is probably critical for the protein folding of CsAlaDC and AtSerDC.

Identification of key amino acids for the substrate specificity

Through the superposition of the amino acid residues present in the substrate binding pocket of CsAlaDC and AtSerDC, it was observed that except for Phe at position 106 in CsAlaDC (**Figure 4A**), the residues were identical. In AtSerDC, this position is occupied by Tyr at 111, suggesting that this specific amino acid residue is related to substrate specificity.

To obtain more insights into the role of specific amino acid residue in substrate specificity, we utilized amino acid sequences of AtSerDC to identify all potential serine decarboxylases in *Embryophyta* and constructed phylogenetic trees (**Supplementary figure 1**). We further identified conserved motifs in these SerDC homologs in *Embryophyta* (**Figure 4B**). Importantly, the Phe¹⁰⁶ of CsAlaDC or Tyr¹¹¹ of AtSerDC locates in a conserved motif (**Figure 4B**). In this conserved motif, the first two amino acids Y and P are 100% conserved, but the third amino acid is variable and can be Y, T, F, V, A, I, or other amino acids (**Figure 4B and C**); interestingly, the Phe¹⁰⁶ of CsAlaDC and Tyr¹¹¹ of AtSerDC locates at this variable amino acid position. For these first three amino acids, the YPY was identified in 83.7% of the SerDC homologs, while the YPF motif was found in 2.3% of these SerDC homologs (**Figure 4C**). These analyses provided more insights into the role of Phe¹⁰⁶ of CsAlaDC and Tyr¹¹¹ of AtSerDC in the substrate specificity.

To verify the role of Phe¹⁰⁶ of CsAlaDC or Tyr¹¹¹ of AtSerDC in substrate specificity, we mutated Phe¹⁰⁶ of CsAlaDC into Tyr (CsAlaDC^{F106Y}) and Tyr¹¹¹ of AtSerDC into Phe (AtSerDC^{Y111F}) and tested enzymatic activity *in vitro*. The results showed that CsAlaDC^{F106Y} totally lost the alanine decarboxylase activity, and AtSerDC^{Y111F} gained alanine decarboxylase activity (**Figure 4E**). We also performed other mutations of Tyr¹¹¹ of AtSerDC, including AtSerDC^{Y111A}, AtSerDC^{Y111I}, AtSerDC^{Y111L}, AtSerDC^{Y111V} and AtSerDC^{Y111W}. Within these mutations, only AtSerDC^{Y111I} exhibited a low level of alanine decarboxylase activity (**Figure 4E**). These results verified that the Phe¹⁰⁶ of CsAlaDC is required for the alanine decarboxylase activity from serine decarboxylase activity.

On the other side, both CsAlaDC and CsAlaDC^{F106Y} exhibit a very low level of serine decarboxylase activity (**Figure 4F**). Moreover, compared with AtSerDC, the serine decarboxylase activity of AtSerDC^{Y111F} was about 30% of AtSerDC; AtSerDC^{Y111I} retained lower than 5% of serine decarboxylase activity of AtSerDC; while other mutations, including AtSerDC^{Y111A}, AtSerDC^{Y111L}, AtSerDC^{Y111V} and AtSerDC^{Y111W} abolished the serine decarboxylase activity (**Figure 4F**). These results suggested the Tyr¹¹¹ of AtSerDC is also important for the substrate specificity of AtSerDC.

To further verify that Phe¹⁰⁶ of CsAlaDC and Tyr¹¹¹ of AtSerDC were key amino acid residues determining its substrate recognition *in planta*, *Nicotiana benthamiana* transient expression system was carried out. To this end, *A. tumefaciens* strain GV3101 (pSoup-p19), carrying recombinant plasmid, was infiltrated into leaves of 5-week-old *N. benthamiana* plants, and the pCambia1305 empty vector was used as the control (EV). Relative mRNA levels were detected (**Supplementary figure 2**), indicating that they had been successfully overexpressed in tobacco. Here, we found that a high level of EA was produced in the CsAlaDC-expressing tobacco leaves; while no EA product was detected in tobacco leaves infiltrated with the mutant CsAlaDC^{F106Y}. In addition, we did not detect EA products in the AtSerDC-expressing tobacco leaves, while a high level of EA was detected in tobacco leaves infiltrated with mutant AtSerDC^{Y111F}. As anticipated, EA was not detected in tobacco leaves infiltrated with EV (**Figure 4G and H**). These results further verified the critical role of Phe¹⁰⁶ in the substrate specificity of CsAlaDC.

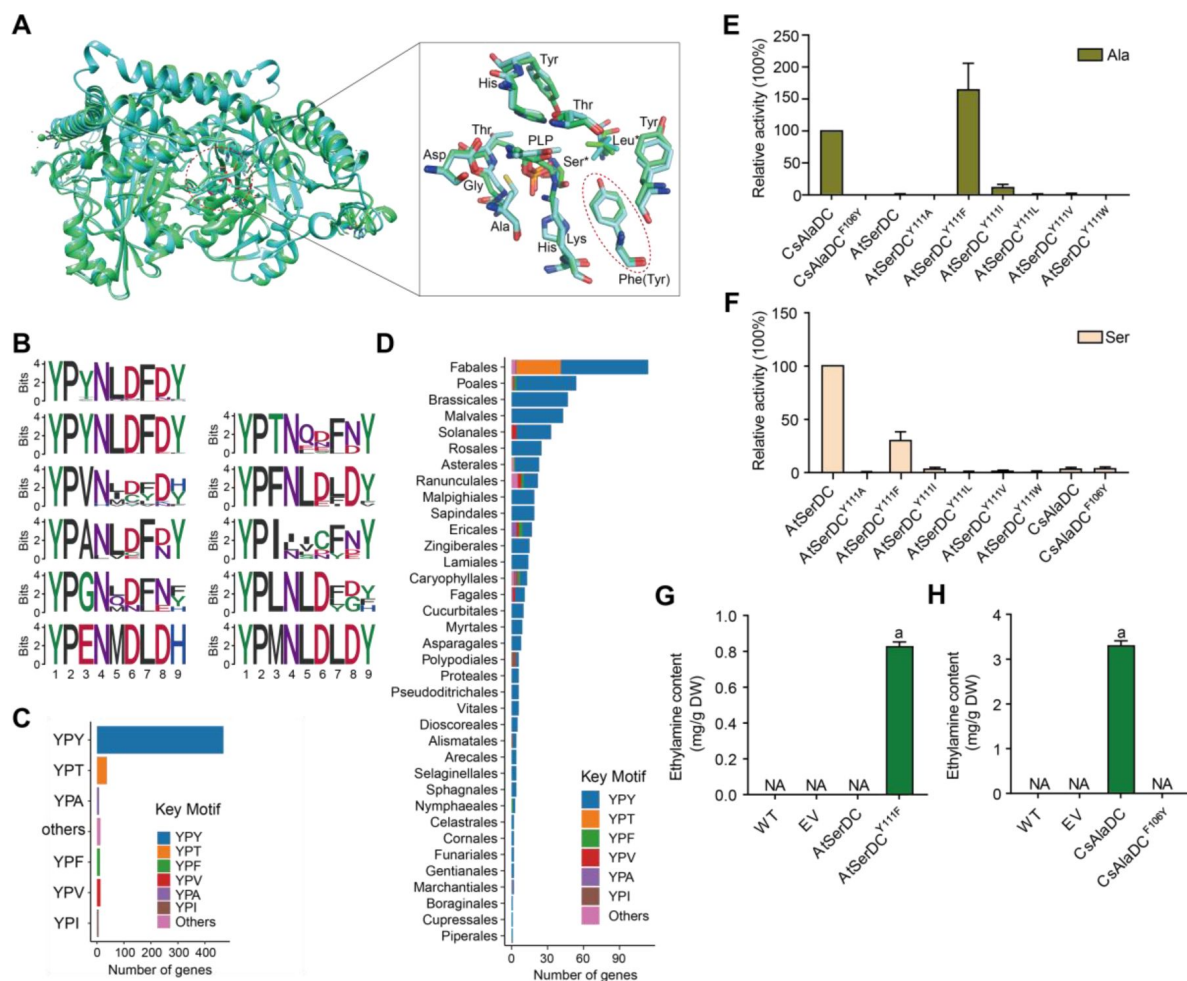


Figure 4.

Key amino acid residues for substrate recognition.

(A) Superposition of substrate binding pocket amino acid residues in CsAlaDC and AtSerDC. The amino acid residues of CsAlaDC are shown in spring green, the amino acid residues of AtSerDC are shown in cyan, with the substrate specificity-related amino acid residue highlighted in a red ellipse. (B) Active-site-lining residues from *Embryophyta* AlaDC were identified. The height of the residue label displays the relative amino acid frequency. (C) Histogram showing the distribution of the number of key motifs. (D) Histogram showing the number of key motifs in different plant orders. (E) Relative enzyme activities of wild-type CsAlaDC and mutant protein CsAlaDC^{F106Y} against Ala substrate (columns 1 and 2), and enzyme activities of wild-type AtSerDC and various AtSerDC mutant proteins against Ala substrate (columns 3-9) are presented. The percentage graph shows the relative activity of each protein compared to wild-type CsAlaDC activity (taken as a 100% benchmark). (F) Relative enzyme activities of wild-type AtSerDC and AtSerDC mutant proteins (columns 1-7) against Ser substrates, and enzyme activities of wild-type CsAlaDC and mutant protein CsAlaDC^{F106Y} against Ser substrates (columns 8, 9) were measured. The percentage graph shows the relative activity of each protein compared to the wild-type AtSerDC (taken as a 100% benchmark). Three independent experiments were conducted. (G) The EA contents of AtSerDC and its mutant AtSerDC^{Y111F} in *N. benthamiana*. (H) The EA contents of CsAlaDC and its mutant CsAlaDC^{F106Y} in *N. benthamiana*. The significance of the difference ($P < 0.05$) was labeled with different letters according to Duncan's multiple range test.

Key amino acids for the evolution of CsAlaDC enzymatic activity

CsAlaDC and AtSerDC have a high sequence similarity of 74.5% and a nearly identical structure with an RMSD (root mean square deviation) of only 0.77Å for monomer structures. However, the two enzymes catalyze different amino acid decarboxylation reactions, and AtSerDC exhibits significantly higher activity than CsAlaDC (**Figure 2B and C**). To elucidate this observation, we proposed a prediction test: EA, generated via Ala decarboxylation by CsAlaDC, can be toxic and harmful to plants if accumulated excessively. Thus, during the evolution of plant serine decarboxylase into alanine decarboxylase, the enzyme has evolved not just to alter the substrate preference but also to reduce catalytic activity to control EA production within a suitable range. Based on this hypothesis, we suggest that mutating specific amino acids in CsAlaDC to those corresponding amino acids in SerDCs could enhance its activity, and the results could provide insights into the evolution of CsAlaDC enzymatic activity.

Dimerization is essential to AlaDC/SerDC activities because the active site is composed of residues from two monomers. To identify the amino acids repressing the enzymatic activity of CsAlaDC during evolution from SerDC, we analyzed the crystal structures and the amino acids at the dimer interface between CsAlaDC and AtSerDC. This analysis revealed that the amino acid residues at positions 66, 97, 110, 114, 116, 117, 122, 315 and 345 are different in CsAlaDC and SerDCs (**Figure 1C**). Therefore, we mutated these amino acids of CsAlaDC into those in the corresponding positions of AtSerDC and CsSerDC, and performed enzyme activity assays (**Figure 5A**). The results demonstrated that CsAlaDC^{L110F} and CsAlaDC^{P114A} exhibited significantly enhanced enzyme activity compared to the wild-type CsAlaDC, with a 2.1-fold and 1.59-fold increase, respectively. Furthermore, the catalytic activity of the CsAlaDC^{L110F/P114A} double mutant exhibits a remarkable 2.3-fold enhancement compared to that of the wild-type protein (**Figure 5B**). These findings suggested a critical role of Leu¹¹⁰ and Pro¹¹⁴ of CsAlaDC in the evolution of enzymatic activity and provided a basis to improve CsAlaDC activity. It is possible that these amino acid residues could potentially augment the hydrophobic nature of the protein dimer interface.

In Vitro synthesis of theanine

The biosynthetic pathway of theanine in tea plants comprises two sequential enzymatic steps: alanine decarboxylase mediates alanine decarboxylation to yield EA, which is subsequently utilized by theanine synthetase to effectuate the condensation reaction with glutamate to synthesize theanine. By conducting mutation screening, we discovered a CsAlaDC mutant protein (L110F/P114A) that exhibited a 2.3-fold increase in catalytic activity compared to the wild-type protein. Subsequently, we employed an *in vitro* theanine synthesis system utilizing CsAlaDC and either glutamate synthetase (PtGS [*Pseudomonas syringae* pv. *syringae*]) or gamma-glutamate methylamine ligase (GMAS [*Methylovorus mays*]) which have the activity to synthesize theanine from EA and Glu (32, 33).

The results illustrated the successful synthesis of theanine using CsAlaDC in conjunction with two theanine synthetases, employing Ala and glutamate as substrates (**Figure 5C**). The theanine content generated by the combination of CsAlaDC^{L110F} and GMAS, as well as CsAlaDC^{L110F/P114A} and GMAS, are 4.57-fold and 6.72-fold higher than the content produced in the combination of wild-type CsAlaDC and GMAS, respectively. Similarly, when combined with the PsGS, comparable outcomes are observed as well. The theanine content resulting from the combination of CsAlaDC^{L110F} and PsGS, as well as CsAlaDC^{L110F/P114A} and PsGS, exhibit enhancements of 1.62-fold and 4.33-fold compared to the wild-type protein combination, respectively. (**Figure 5C**). Hence, the utilization of CsAlaDC^{L110F/P114A} could effectively enhance the theanine production yield and thus, holds potential for large-scale engineering production of theanine.

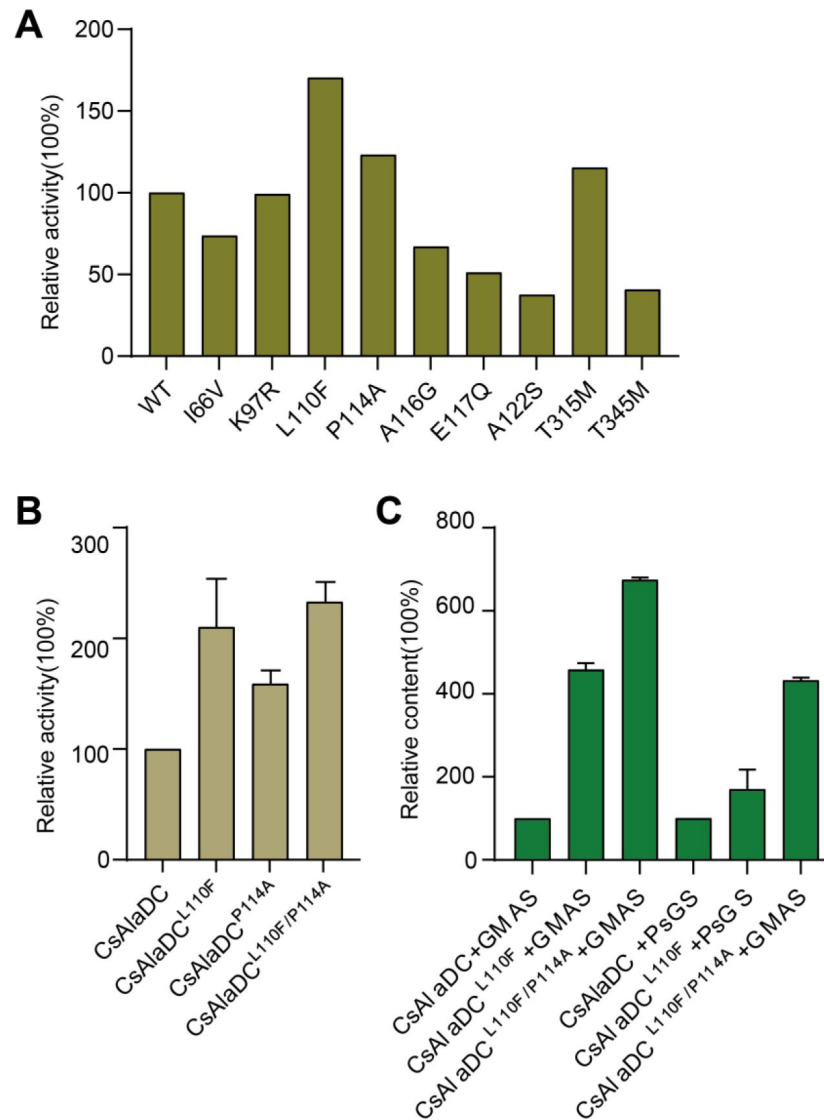


Figure 5.

Mutations enhance CsAlaDC enzyme activity and theanine synthesis *in vitro*.

(A) Relative enzyme activities of CsAlaDC mutant proteins against Ala substrate. (B) Relative enzyme activities of CsAlaDC^{L110F}, CsAlaDC^{P114A}, and CsAlaDC^{L110F/P114A} against Ala substrate. (C) Histogram showing the relative content of theanine resulting from different combinations of alanine decarboxylase and theanine synthetase. Three independent experiments were conducted.

Comparison of CsAlaDC and AtSerDC with other decarboxylases

The Dali search (34) revealed that AtSerDC and CsAlaDC exhibit structural similarity to several other decarboxylases, including 7CIG (MetDC [*Streptomyces* sp. 590]), 7ERV (HisDC1 [*Photobacterium phosphoreum*]), 6KHO (TrpDC [*Oryza sativa Japonica* Group]), 4E1O (HisDC2 [*Homo sapiens*]), 6JY1 (AspDC [*Methanocaldococcus jannaschii*]), and 5GP4 (GluDC [*Levilactobacillus brevis*]). The monomers of these amino acid decarboxylases all consist of three characteristic domains (Figure 3 and Supplementary figure 4). The C-terminal domain, also known as the small domain, is composed of three antiparallel β -sheets and three α -helices, facing toward the opposite monomers. The large domain contains seven stranded β -sheets surrounded by eight α -helices, exhibiting characteristics typical of Type I enzymes. This domain includes a conserved PLP-binding active center and amino acid residues that are conserved across species. The N-terminal domain consists of elongated α -helices and is connected to the large domain via a loop. The helix of one subunit is antiparallel to the corresponding helix in the neighboring subunit, forming a clamp. Therefore, the N-terminal domain may not be an independent folding unit and is likely only stable in dimers.

The monomeric structure of CsAlaDC and AtSerDC exhibits numerous structural features common to other Group II PLP-dependent amino acid decarboxylases, and the amino acid residues that bind to PLP cofactors are conserved across multiple enzymes (Supplementary figure 3 and 4). Out of the eight proteins analyzed, a total of thirteen amino acid residues were found to be conserved across all sequences (Glu¹⁴³, Glu¹⁷¹, Lys²⁰¹, Gly²⁴⁵, Thr²⁴⁷, Asp²⁵³, His²⁷⁵, Asp²⁷⁷, Ala²⁷⁹, Pro²⁸⁶, Ser³⁰², Lys³⁰⁹, and Tyr³³⁶ in CsAlaDC). Some of them are situated within the active center and play a role in stabilizing PLP or facilitating catalytic reactions. Conversely, other residues reside outside the active center and their function remains unclear. (Supplementary figure 3). PLP is bound to conserved lysine residues in the active center through the Schiff base. The carboxylic group of conserved Asp residues stabilizes protonated pyridine nitrogen of PLP by electrostatic action, providing the latter with strong electrophilicity to stabilize carbanion intermediates during enzymatic catalysis. Another conserved His residue is situated laterally to the PLP, while on the opposite side, there is a conserved Ala residue. The oxygen atom at position 3 of PLP is stabilized by hydrogen bonding with the Thr residue. The phosphate moiety of PLP is additionally secured to the protein through an elongated system of hydrogen bonding. The α 5 helix located within the large domain exhibits evolutionary conservation in Group II enzymes, facilitating interaction with the PLP phosphate moiety and conferring appropriate stabilization thereof.

The superimposition of the crystal structures of CsAlaDC and AtSerDC with those of related enzymes suggests that residue Phe106 in CsAlaDC (corresponding to Tyr in AtSerDC and Glu in MetDC) may play a role in determining substrate specificity. This hypothesis was tested via mutagenesis experiments, which revealed that substitution of Phe¹⁰⁶ for Tyr rendered CsAlaDC inactive, while substitution of Tyr¹¹¹ for Phe in AtSerDC enabled alanine decarboxylase activity. Likewise, the substitution of Glu⁶⁴ for Ala in MetDC reduced enzyme activity by approximately 50-fold compared to the wild-type (35), supporting this conclusion.

Group II amino acid decarboxylase is characterized by the existence of highly flexible catalytic rings, which are of great significance for the catalytic mechanism of decarboxylase (31). The catalytic rings are located at the dimer interface and extend to the active sites of other monomers in a closed conformation. In the structure of CsAlaDC, the catalytic ring is orderly and allocated, while in the structure of AtSerDC, it is disordered and lacks electron density (Figure 3F). Prior research has established that the conserved amino acid residue Tyr within the ring plays a crucial role in catalysis by donating protons to carbanions of quinone intermediates that arise following decarboxylation (36, 37). Our experiments demonstrate that the substitution of the

corresponding Tyr with Phe in the loop renders the protein inactive. Furthermore, alterations in the structure of CsAlaDC when combined with EA provide further evidence supporting this conclusion. Specifically, the structure of CsAlaDC-EA complex displayed a significant conformational change in Tyr^{336*} located in the catalytic site (as depicted in **Figure 3G**). Comparing the structure of CsAlaDC apo, the CsAlaDC-EA complex exhibited a notable 60-degree shift in the p-hydroxyl group of Tyr^{336*} along the Ca-C β bond.

The monomeric configuration of CsAlaDC and AtSerDC is akin to that of other Group II amino acid decarboxylases, except for a conspicuous dissimilarity. Unlike other Group II amino acid decarboxylases, the C-terminal of CsAlaDC and AtSerDC harbors an evident zinc finger structure. We propose that this structure could potentially influence enzyme stability. To test this hypothesis, we truncated the zinc finger structure from both proteins and expressed them. Our findings indicate that, following the excision of the zinc finger structure, CsAlaDC became insoluble, while AtSerDC displayed a similar trait to a certain extent.

Evolution of AlaDC and SerDC

The sequence identity between CsAlaDC and AtSerDC is 74.5%, indicating a significant degree of homology. Moreover, the two proteins exhibit a high level of monomer structural similarity. Notably, they differ in substrate specificity. Using the crystal structure of CsAlaDC and AtSerDC as a basis, we identified pivotal amino acid residues that govern the substrate specificity of serine or alanine decarboxylase. We observed that sequences proximal to these sites exhibit a high degree of conservation in both proteins and possess characteristic motifs (**Figure 4B**). Moreover, we analyzed serine decarboxylase-like proteins in *Embryophyta* that are highly homologous to CsAlaDC and constructed phylogenetic trees to gain insight into their evolutionary relationships. Based on our experimental findings and evolutionary evidence, we have identified a conserved YPX motif at the substrate binding pocket. For most plants, the residue X in this motif is predominantly Tyr, which corresponds to the serine decarboxylase protein. However, other residues such as Thr, Phe, Val, Ala, and Ile can also occupy this position (**Figure 4B and C**). Interestingly, serine decarboxylase-like proteins containing YPF motifs, which have the potential to function as alanine decarboxylases, were found to be distributed throughout the phylogenetic tree, including *Asteroidae*, *Ericales*, *Chenopodiaceae*, *Poaceae*, and *Ranunculales* (**Supplementary figure 1B**). However, these proteins were absent in some more recent species. This observation implies that the emergence of alanine decarboxylase in these particular species could be attributed to convergent evolution.

Moreover, we have observed the enrichment of serine decarboxylase-like proteins that possess a YPT motif within the *Fabales*. Prior research has demonstrated that XP_004496485-1, which possesses a YPT motif, displays catalytic efficacy toward the processes of decarboxylation and oxidative deamination of Phe, Met, Leu, and Trp (38). Given that the YPT motif is highly conserved and widely distributed in *Fabales*, serine decarboxylase-like proteins bearing the YPT motif may have developed a unique substrate specificity in *Fabales*, beyond their conventional decarboxylation functions, as exemplified by XP_004496485-1 protein mentioned above. We speculate that these proteins may be capable of catalyzing other reactions, potentially involving non-protein amino acids or other substances as substrates.

Applied to improve the synthesis of theanine

Theanine is an important indicator of green tea quality. Therefore, improving the synthesis of theanine in tea plants is the focus of research. In this study, through crystal structure analysis and mutation verification, we found that the catalytic activity of CsAlaDC^{L110F/P114A} is 2.3 times higher than that of the wild-type protein, resulting in a more abundant synthesis of theanine *in vitro*. This gives us great inspiration to improve the synthesis of EA in tea plants by gene editing, thus increasing the content of theanine in tea plants (**Figure 5**).

Theanine is highly demanded, by the market, due to its health effects and medicinal value, and as a constituent in food, cosmetics and in other fields. To meet the market demand, a variety of methods have been used to acquire theanine, with the main methods including direct extraction, chemical synthesis, biotransformation (microbial fermentation) and plant cell culture. Based on this study, we can use mutation technology to modify engineering bacteria and improve the possibility of synthesis of theanine in bacteria, yeast, model plants, etc, promoting its use for industrial theanine production.

Conclusions

In summary, we determined the crystal structures of the CsAlaDC-EA complex and AtSerDC. We employed structural biology to reveal the substrate selection mechanism of these two decarboxylases. By structural analysis and enzyme activity experiments, we found that Tyr¹¹¹ of AtSerDC and Phe¹⁰⁶ of CsAlaDC as crucial sites for substrate specificity. Next, through mutation screening, we identified a CsAlaDC mutant protein (L110F/P114A) with catalytic activity 2.3 times higher than that of the wild-type protein. In an *in vitro* L-theanine biosynthesis system with added PsGS or GMAS, we demonstrated that the L-theanine yield of the CsAlaDC^{L110F/P114A} group was 4.3 and 6.7 times higher than that of the wild-type control group, respectively. In addition, by utilizing the substrate recognition key motif of CsAlaDC and AtSerDC, evolutionary analysis was conducted on protein sequences exhibiting high homologous to CsAlaDC in *Embryophyta*, which resulted in the discovery of 13 potential alanine decarboxylases. Furthermore, we discovered serine protease-like proteins containing special motifs that have evolved in *Fabale*, which may possess unique substrate specificity and catalytic functions. On the other hand, our study has effectively enhanced the efficiency of the initial stage of theanine synthesis and is anticipated to offer a promising avenue for developing a novel approach to producing L-theanine, thereby facilitating the progress of the tea industry.

Materials and Methods

Plants materials

Tobacco (*Nicotiana benthamiana*) plants were grown in a controlled chamber, under a 16-h light and 8-h dark photoperiod at 25 °C. Leaves of 5-week-old tobacco plants were used for transient transformation, mediated by *Agrobacterium tumefaciens* strain GV3101.

Gene cloning and protein expression

The coding sequences of AtSerDC and CsAlaDC, amino acid residues 66-482 and 61-478, respectively, were amplified using PCR and then ligated into the *Nde* I and *Xho* I restriction sites of the pET-28a and pET-22b vector, respectively, providing the recombinant vector pET-28a-AtSerDC and pET-22b-CsAlaDC. The cDNAs encoding CsSerDC were amplified using PCR and then ligated into the *Nde* I and *Xho* I restriction sites of the pET-28a vector. Gene-specific primers were listed in **Supplementary table 2** [↗](#). The recombinant plasmid was transformed into *E. coli* BL21 (DE3) competent cells. Positive transformants were grown in a 5 mL LB medium containing 30 µg/mL kanamycin or 50 µg/mL ampicillin at 37 °C overnight and then subcultured into an 800 mL LB medium containing the corresponding antibiotic. Protein expression was induced by the addition of 0.2 mM isopropyl-β-d-thiogalactoside (IPTG) for 20 h at 16 °C when the optical density (OD) at 600 nm reached 0.6-0.8, harvested by centrifugation at 4 °C and 4,000 rpm for 30 min.

Protein production and crystallization

The cell pellet was suspended in 30 mL of lysis buffer (20 mM HEPES, pH 7.5, 200 mM NaCl, 0.1 mM PLP), disrupted by High-Pressure Homogenizer, and then centrifuged at 16,000 rpm for 30 min at 4 °C to remove the cell debris. The supernatants were purified with a Ni-Agarose resin column followed by size-exclusion chromatography. Before crystallization, purified proteins were concentrated at 10 mg/mL. Crystallization conditions were screened by the sitting-drop vapor diffusion method using the reservoir solutions supplied in commercially available screening kits (Crystal Screen, Crystal Screen 2, PEGRx 1, 2, and SaltRx 1, 2). A droplet made by mixing 1.0 μ L of purified AtSerDC or CsAlaDC (10 mg/mL) with an equal volume of a reservoir solution was equilibrated against 100 μ L of the reservoir solution at 16°C. The crystal of AtSerDC was obtained using buffer pH 8.0 containing 20% (w/v) PEG400 as a precipitate and 0.2 M CaCl_2 . The crystal of the CsAlaDC-EA complex was obtained at pH 7.5 containing 2.6 M sodium acetate. The crystal of CsAlaDC was obtained at pH 6.0 containing 3.5 M sodium formate.

Data collection and processing

Crystals were grown by sitting-drop vapor diffusion method at 16 °C. The volume of the reservoir solution was 100 μ L and the drop volume was 2 μ L, containing 1 μ L of protein sample and 1 μ L of reservoir solution. The reservoir solution of AtSerDC contained 0.1 M Tris-HCl (pH 8.0), 0.2 M CaCl_2 , and 20% PEG 400. The reservoir solution of CsAlaDC contained 0.1 M HEPES (pH 7.5), 2.6 M sodium acetate or 0.1 M Bis-Tris (pH 6.0), and 3.5 M sodium formate. Crystals grew in 3-5 days using a protein concentration of 10 mg/mL. Diffraction data were collected at Shanghai Synchrotron Radiation Facility (China). The collected data sets were indexed, integrated, and scaled using the HKL3000 software package. The structure of AtSerDC was solved by molecular replacement using the structure of HisDC (PDB code: 7ERV [*Photobacterium phosphoreum*]) as the model, and utilizing AtSerDC as a molecular displacement model for both CsAlaDC and the CsAlaDC-EA complex. The AtSerDC, CsAlaDC, and CsAlaDC-EA complex exhibit resolutions of 2.85 Å, 2.50 Å, and 2.60 Å, respectively. The statistics for data collection and processing are summarized in **Supplementary table 1** [↗](#).

Enzyme activity assays

Decarboxylase activity was measured by detecting products (EA or ethanolamine) in Waters Acquity ultraperformance liquid chromatography (UPLC) system ([20](#) [↗](#), [39](#) [↗](#)). The 100 μ L reaction mixture, containing 20 mM substrate (Ala or Ser), 100 mM potassium phosphate, 0.1 mM PLP, and 0.025 mM purified enzyme, was prepared and incubated at standard conditions (45 °C and pH 8.0 for CsAlaDC, 40 °C and pH 8.0 for AtSerDC for 30 min). Then, the reaction was stopped with 20 μ L of 10% trichloroacetic acid.

The detection methodology for theanine production remains consistent with the aforementioned approach. The 100 μ L reaction mixture, containing 20 mM Ala, 45 mM Glu, 50 mM HEPES (pH 7.5), 0.1 mM PLP, 30 mM MgCl_2 , 10 mM ATP, 0.03 mM PsGS/GMAS, and 0.025 mM CsAlaDC/CsAlaDC L110F/P114A, was prepared and incubated at standard conditions for 1 h.

Subsequently, the reaction was terminated via immersion of the reaction vessel in a metal bath at 96 °C for 3 min.

The product was derivatized with 6-aminoquinolyl-N-hydroxy-succinimidyl carbamate (AQC) and subjected to analysis by UPLC. All enzymatic assays were performed in triplicate.

Site-directed mutagenesis

Site-directed mutagenesis experiment was conducted using a PCR method from the wild-type construct pET-28a-AtSerDC and pET-22b-CsAlaDC, respectively. *Dpn* I endonucleases were used to digest the parental DNA template. The reaction mixture was used to transform *E. coli* DH5 α

competent cells and the plasmids from positive strains were extracted to *E. coli* BL21 (DE3) for protein expression, purification, and analyzing the enzymatic activity.

In vivo* enzyme activity assay in *N. benthamiana

The amplified PCR products were fused to the plant expression vector, pCambia1305. Linearization was conducted by restriction digest with *Spe* I and *Bam*H I. The recombinant colonies were selected for PCR validation on the appropriate antibiotics plate. After validation, the plasmids were electroporated into *Agrobacterium tumefaciens* strain GV3101(pSoup-p19). Empty pCambia1305 vector, with an intron containing the GFP gene, was treated as the control.

Agrobacterium transient expression assays were performed on 5-week-old *N. benthamiana* plants. *Agrobacterium tumefaciens* strain GV3101 (pSoup-p19), carrying the above-described vectors, were cultured in Luria Bertani (LB) medium, containing appropriate antibiotics, at 28 °C. When the absorbance of bacteria colonies reached OD₆₀₀ = 0.6-0.8, bacterial cells were collected and resuspended in MMA solution (10 mM MgCl₂, 10 mM 2-(N-morpholino) ethane sulfonic acid (MES), pH 5.6). After the OD₆₀₀ of the resuspended bacterial solution was adjusted to approx. 1.0, acetosyringone (AS) was added with a final concentration of 200 µM, and this solution was then incubated, at room temperature, for at least 3 h in darkness. Next, cell suspensions were infiltrated into *N. benthamiana* leaves with a needle-free syringe. These *N. benthamiana* leaves were then collected 3 days post infiltration, frozen in liquid nitrogen and stored at -80 °C. Internal EA in *N. benthamiana* leaves was extracted as previously described ([22](#)), and then the solvent was subjected to gas chromatography-mass spectrometry GS-MS system.

Transcript level analysis in *N. benthamiana*

Total RNA was isolated from samples using the RNeasy Pure Plant Kit (Qiagen, Beijing, China), according to the manufacturer's protocol. The cDNAs were synthesized using TransScript One-Step gDNA Removal and cDNA Synthesis SuperMix Kit (TransGen Biotech, Beijing, China). The qRT-PCR assays were performed, as previously described ([40](#)). Primers used for qRT-PCR assays were listed in **Supplementary Table 3** and the qRT-PCR was run on a Bio-Rad CFX96TM RT PCR detection system and CFX Manager Software. Each reaction reagent (20 µL) contained 0.4 µL forward and reverse primers (10 µM), 2 µL cDNA (200±5 ng/µL), 10 µL SYBR Green Supermix (Vazyme, Nanjing, China) and 6.2 µL double-distilled water. Reaction was performed by a two-step method: 95 °C for 5 min; 40 cycles of 95 °C for 10 s; and 60 °C for 30 s. The glyceraldehyde-3-phosphate dehydrogenase (GAPDH) gene was used for internal normalization in each RT-PCR, and the 2^{-ΔΔCT} method ([41](#)) was used to calculate the relative gene expression. All samples were performed with three replicates.

Multiple sequence alignment

MUSCLE was used to generate the protein multiple sequence alignment of amino acid decarboxylases with default settings ([42](#)). ESPript 3.x was used to display the multiple sequence alignment ([43](#)).

Phylogenetic analysis

The CsAlaDC was used as a query for searching its homologs in *Embryophyta* using BLASTP ([44](#)) (version 2.13.0). To fetch homologs of CsAlaDC in distant species, we performed a two-round blast. We first search for the best hit in the given order under *Embryophyta*. Then we blast each species under a provided order using the order best hit as a query and as many as 10 hits were kept. The sequences with e values ≤ 0.001 were removed. We further dropped sequences with lengths less than 200 or larger than 800 amino acids. Then multiple sequence alignment was performed via clustalo (version 1.2.4) with default settings on these filtered sequences ([45](#)). The NJ algorithm

implemented in clustalw (version 2.1) was adopted to build the phylogenetic tree (46 [↗](#)). The R package ggtree (version 3.2.1) and universalmotif (version 1.12.4) were used for tree and motif visualization (47 [↗](#)).

Acknowledgements

We thank the team of beamline BL18U1 in the Shanghai Synchrotron Radiation Facility for diffraction data collection.

Funding

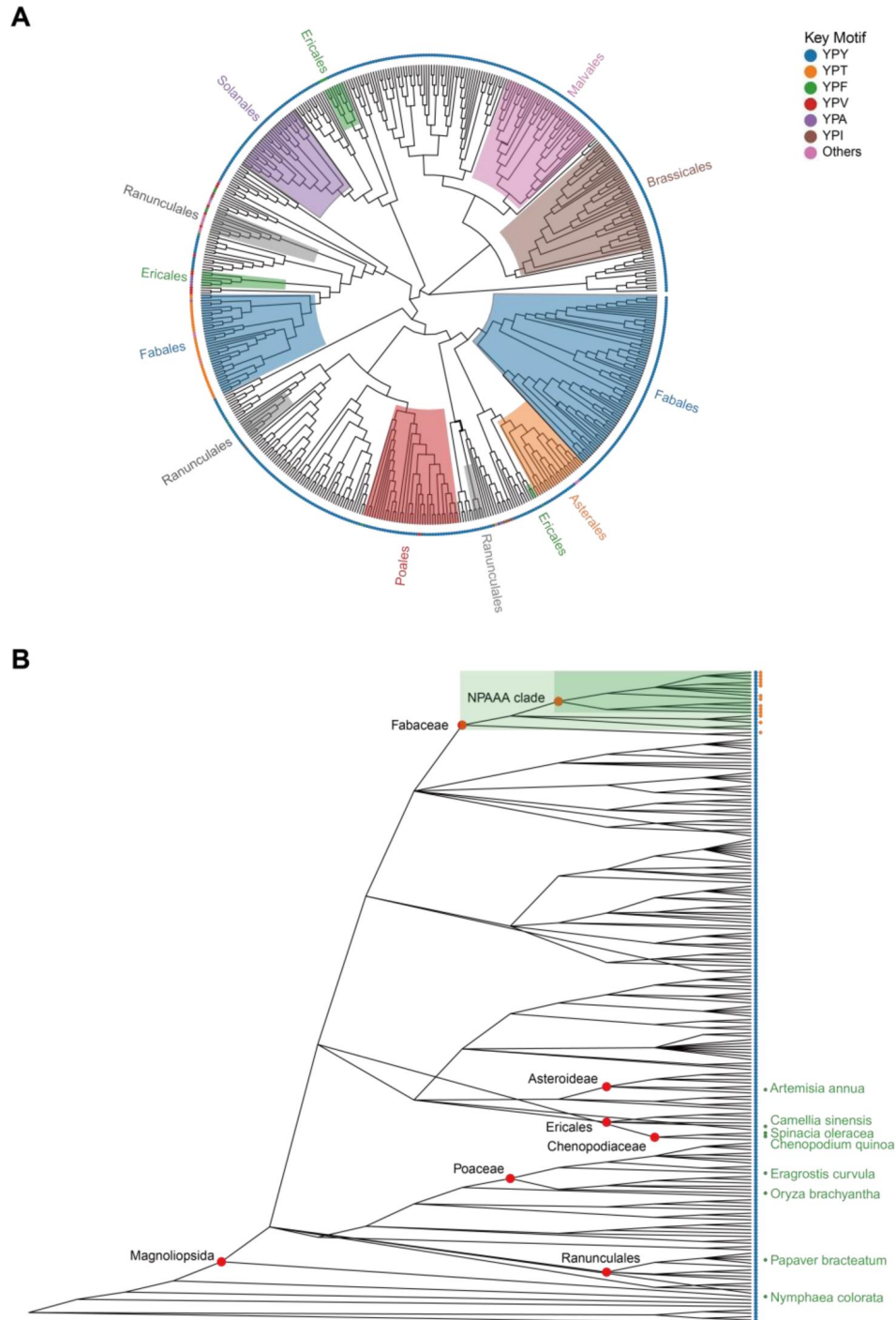
This study is supported by grants from the Ministry of Science and Technology of China (2019YFA0904100 to W.G, 2022YFF1003103 to X.W) and the Natural Science Foundation of China (T2221005 to W.G, 32072624 to Z.Z).

Data availability

The structures of AtSerDC, CsAlaDC-EA complex and CsAlaDC have been deposited in the Protein Data Bank (PDB) under accession numbers 8JG7, 8JIK and 8JIJ respectively.

Supplementary figures

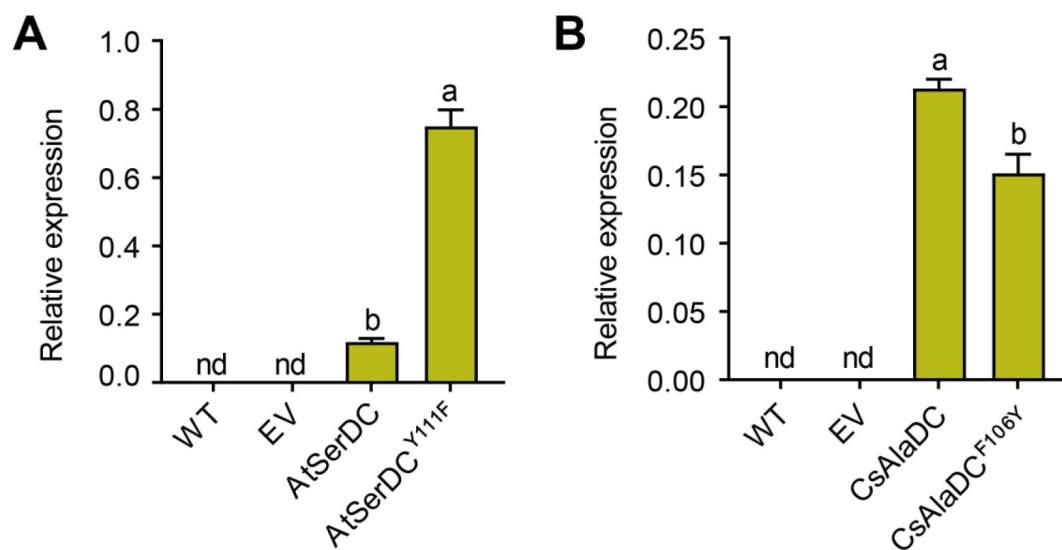
Supplementary tables



Supplementary figure 1.

Evolutionary analysis of CsAlaDC in *Embryophyta*.

(A) The presented diagram depicts an evolutionary tree of CsAlaDC, which is devoid of a root and solely portrays the topological structure of the tree without including distance information. The color of the inner ring corresponds to various orders, while the outer ring's leaf nodes are colored based on the motif types that the sequence exhibits. (B) Diversity of serine decarboxylase-like proteins in *Embryophyta* (196 species). Colored scatter spots on the right side of leaf nodes correspond to the respective motifs shown in Figure A.



Supplementary figure 2.

The relative mRNA levels of AtSerDC and its mutant Y111F (A), CsAlaDC and its mutant F106Y (B) in *N. benthamiana* leaves were measured by two primers. WT, wild type of *N. benthamiana*; EV, empty vector control; *NbGAPDH* was used as an internal control. Data represent mean $\pm$ SD (n=3). The significance of the difference ($P < 0.05$) was labeled with different letters according to Duncan's multiple range test.

CsAlaDC

70 80 90 100 110 120 130 140

CsAlaDC D VDDSTSDMEAFMAGVLVVRKRLTLEIKTKYHLGYPFNLDLDYGP..LAELQHFAINNLGDPFIESNYSNVHQRQFVGVGLD
AtSerDC EVNDEFTGDKEAYMAASVLARVKTLIVERTKNHLYPYNLDFDYGA..LGQLQHFSINNLGDPFIESNYSNVHRPFPPVGVGLD
MetDC GLDDD...RRRLRALDAVDEY...LTRKRKHVLGYQATQDMQGTAALDARFMFPNNINNLGDPFGSGGYKPNKKVVBRVAVGLD
HisDC1 NKLDDEF...WAYCVKNQY...FNIGYVESADFDYTI..LERFMRFISINNCGDWAAYSNYLLNFDFDFKEVME
TrpDC ELRSSVVPGMTHWASPNNFAF...FPSTNSAAAI..AGDLIASAMNTVGFFTWQAS...PAATEMVAVGLD
HisDC2 DIERIIMPVGVIHQSPHHMAY...YPALTSSWPSL..LGDMLADAICNLGFTWASS...PATCELMNMVD
TyrDC NIFGSMCSNVLPITRKIVIDIF...RLET...NLGDPGLFK...GTKLLEKAVA
GluDC QILDEGNRSRLNLATFCQTYMBE...PEAEVELM..KDTLEKNAIDKSEYP...RTAETENRCVN
consensus>70 d...y...yp...l...in.g...E...ld

CsAlaDC 150 160 170 180

CsAlaDC W F A R L W E I E Q K E Y W G Y I T N . G G T C E N L H G L L I G R E V F P G G I F
AtSerDC W F A R L W E I E R D D Y W G Y I T N . C G T E G N L H G L L I G R E M F P G G I L
MetDC Y Y A R L W E I R H P D A D P E S Y W G Y I T N . M G S T E G N M Y A L W N A R D Y L S G K A L I Q P T A P F D A V R Y K A D P D R R N P N A H . H P V
HisDC1 Y F A D L F K I P F E D S W G Y I T N . G G T E S N M F G V I L G R E L F P G G I L
TrpDC W L A O M L N L P T S F M N R T G E G R G T G G G V I L G T T E A M L V T L V A A R D A A L R R S G S D G V A . G L H . R L A
HisDC2 W L A K M L G L P E H F L H H P S S Q G G V I L G T V S E S T L I A L L A A R K N K I L E M K T S E P D A D E S S I N A R L V
TyrDC L L G S L L N N K D A Y G H I V S . G G A L E M L M A L R C I K N I W R E K R . R K G L S . K N E . H P K
GluDC I A N L W H A P E A E S F T G T S I G S S E C M L A G L A M K F A W R K R A K A G L D L T A H . Q P N
consensus>70 ..a.\$..g.i..E.\$..re..dg.

CsAlaDC 190 200 210 220 230

CsAlaDC J Y T R E S H Y S I F K A A R M Y R M E C V K V G T I L N G E I I C A D F K A K D L S N K
AtSerDC J Y A S R E S H Y S V F K A A R M Y R M E C E K V D T L M S G E I I C D L R K K L L A N K
MetDC A F Y S E D T H Y S F A K A V A V L G V E T F H A V G L E K Y A D C P L D V P V T G L R T W P T E V P S R P G P S G L S W D G P E I V D D A L A V I V E F F A
HisDC1 J Y Y S K D T H Y S V A K I V K L L R I K S Q L V D S L P N G E I I Y D D I S K I K Q D D
TrpDC V Y A A D Q T H S T F F K A C R L A G F P P A N I R S I P T G A E T D Y G L P A R L E A M Q A D A
HisDC2 A Y A A D Q A H S S V E K A G L S L V K M K F P V D D N F S L R G E A L G A I E E D K
TyrDC I I V P I T A H S F E K G R E M D L G Y I Y A P I K E D Y T I E K V F K A V A D E Y D
GluDC I I V S A G Y Q V C W E K F C V Y W D I L M H V F P M D D D H M S L N V D V L D Y V D Y
consensus>70 .y.s.e.h.s.K..e.v..d..

CsAlaDC 240 250 260 270 280

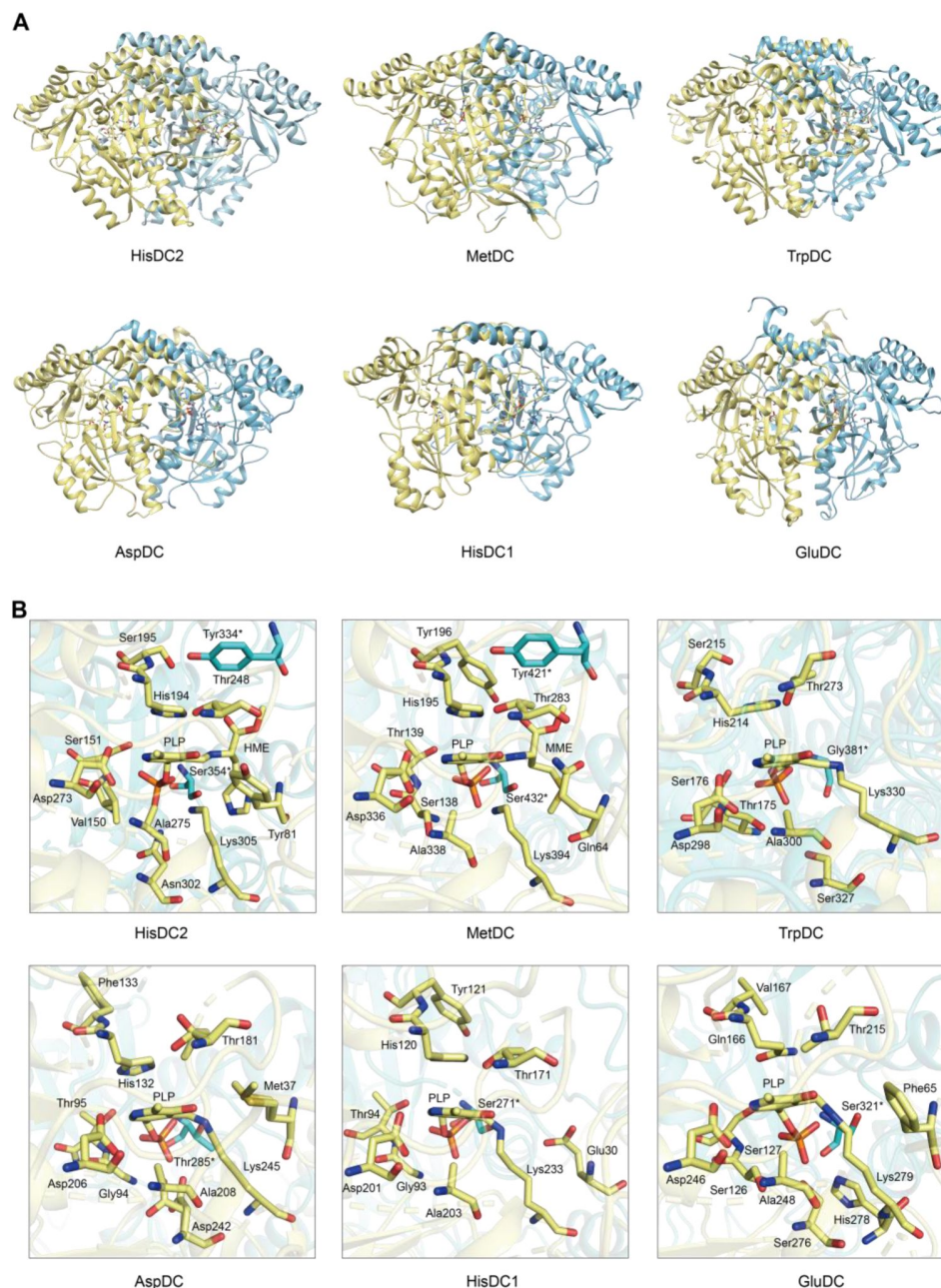
CsAlaDC D K . . . P A I L N L N I G T V K G A V D D I D L V I Q T T E E C G F S H D R . F Y I H C D G A L F G . . F M M F F
AtSerDC D K . . . P A I L N V N I G T V K G A V D D L D L V I K T T E E C G F S H D R . F Y I H C D G A L F G . . L M M F F
MetDC A K G H . P V E V N L N L G S T F K G A V D D V R A V C E R L I P E R H G L V Q R E V V Y G S C P T G R P L V D V R R G F W I H C D G A L G A . G Y A F F
HisDC1 E K . . . H P I I F A N I G T V R G A D D I S K I Q A M I G E L G I K R E D . Y Y I H A D A A L G S . M I I F F
TrpDC D A G L V P T V C A T V G T S S N A V D P V G A V A D V A A R F A W V H V D A A Y G A S C I T P E
HisDC2 Q R G L V P V C A T L G T G V C A C X L S E L G P I C A R E G L W I H I D A A Y G A S T A L P E
TyrDC V D G I I G I A C E L T G I D N I E E L S K I A K E N N I Y I H V D A A F G C . L V I P F
GluDC T I G I V G I M C T Y T Q V Y D I A R L D A V V E R Y N R I T K F P V Y I H V D A A S G G . F Y I P E
consensus>70 ...1.Gt.r.ga.Dd...v.Si.s.hk...cg...n.Yl...

CsAlaDC 290 300 310 320 330

CsAlaDC I N R G P K I T F K K P I G S V S V S G H K F M G C P T P C G V Q I I L E H . . I N A D S R N V E Y L A S .
AtSerDC V K R A P K V T F N K P I G S V S V S G H K F V G C P M P C G V Q I I T M E H . . I K V L S R N V S Y L A S .
MetDC I L R A A E D P E G V G W T P E A L P E F D F G L R I P T A G H G E V D M S S T A M S G H K A G A P W C G V I Y M T K V Y . . Q I S P S Q N P Y I G A .
HisDC1 V D E P Q G F N F A D G I D S V S G H K M I G S P I P C G I V A K R N . . V D A I S E I D Y I S A .
TrpDC F R H H L D V E R V D S I S M S P F K W L M T C D T C L Y V D T H R L T G S L E T N P R I K N H .
HisDC2 F R G F L K G E Y A D S F T F N P K M M V M H F C T G F W V D K Y K L Q T F S V N P I Y R H .
TyrDC L D D K Y K K G V N Y K F D F S L G V D S I T I D P K M G H C P I S G G I L F K I D Y . K R Y L D V D A P Y L T E .
GluDC I E P E L K W D F R L N N V I S N A S G H K I G L V P Y G V W V I W D D Q Q Y L P K E L V F K V S Y I G G E
consensus>70 ...f.v.Si.s.hk...cg...n.Yl...

CsAlaDC 340 350 360 370 380 390 400

CsAlaDC . R D A T I T G S R N G H S F I L W Y A L N R K F K G F Q K E V Q K C L R N A H Y I K D R I R A G I S A M L N E L . . S S
AtSerDC . R D A T I T G S R N G H A P L F L W Y T L N R K Y K G F Q K E V Q K C L R N A H Y I K D R I R A G I S A M L N E L . . S S
MetDC . P D T I F A G S R N G F S L L L W D H L S R Y Y R D Q V E R I R E A Q E L A A Y I E R L T A M E R L G V E L W P A R T P G A V
HisDC1 . H D K I T I G S R N G H T P L M M V A V K S H A D F K R R I N R S L D A A H A V Q R L Q T A G I N A K N N . . S I
TrpDC A S D S G E V I D L K D M Q V G V G R R F R G L K L M V M Y R I V G A K L Q E H I R S D V M A K V P E D L R G D R E V V I V P R N . . F A
HisDC2 . A N S G V A D F M H W Q I P L S R R F S R V K L V F I R S F G V K N I Q A H V R H G T E M A K Y F E S I V R N D P S F I P A K R H . . L G
TyrDC T R Q A T I L G T R V G F G S C T Y A V L R Y S F G R E G Q R K I V N E C M E N T L Y I Y K K L K E N N K P V I E P I . . L N
GluDC L P T M A I F S S A S Q L I Q Q Y Y N F I R F G P D Y G R I T E K T H D V A R Y I A K S L T K I G G F S L I N D G H E . L P
consensus>70 ...d.i...g...lw...g...v...a.y...l...e...



Supplementary figure 4.

Structures of HisDC2, MetDC, TrpDC, AspDC, HisDC1 and GluDC.

(A) The Overall Structures of HisDC2, MetDC, TryDC, AspDC, HisDC1 and GluDC. Chain A is shown in khaki, chain B is shown in cyan. (B) Amino acid residues in the substrate binding pocket of HisDC2, MetDC, TryDC, AspDC, HisDC1 and GluDC. The amino acid residues in chain A are shown in khaki, and the amino acid residues in chain B are shown in cyan.

	CsAlaDC	CsAlaDC-EA complex	AtSerDC
Data Collection			
Space Group	<i>P3121</i>	<i>P3121</i>	<i>C2</i>
Cell Dimensions			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	163.5,163.5, 50.2	164.0,164.0,50.2	115.4,149.5,139.1
α , β , γ (°)	$\alpha=\beta=90$, $\gamma=120$	$\alpha=\beta=90$, $\gamma=120$	$\alpha=\gamma=90$, $\beta=109.2$
Wavelength (Å)	0.979	0.979	0.979
Resolution (Å)	50-2.50	50-2.60	50-2.85
Completeness (%)	99.3 (99.9)	99.6 (100.0)	99.6 (98.1)
Redundancy	5.4 (5.1)	13.0 (12.5)	5.9 (5.0)
$\langle I \rangle / \langle \sigma I \rangle$	9.3 (1.9)	18.9 (4.6)	10.6 (2)
<i>R</i> _{merge}	0.128 (0.603)	0.156 (0.562)	0.173 (0.551)
Refinement			
<i>R</i> _{work} / <i>R</i> _{free}	0.173/0.222	0.172/0.223	0.214/0.262
No. reflections	25175	22813	49543
B-facator (Å ²)			
Protein	43.3	48.3	57.6
PLP	29.5	31.5	52.8
Zn ²⁺	58.7	43.8	71.6
Ca ²⁺	-	55.5	-
Ethylamine	-	27.5	-
Glycerol	-	-	51.7
Water	38.5	40.9	-
R.m.s. deviations			
Bond lengths (Å)	0.008	0.008	0.007
Bond angles (°)	1.531	1.601	1.410
No. of atoms			
Protein	3292	3265	3250
PLP	15	15	60
Zn ²⁺	2	1	4
Ca ²⁺	-	1	-
Ethylamine	-	5	-
Glycerol	-	-	24
Water	115	60	-
Ramachandran statistics (%)			
Favoured	94.01	93.58	91.52
Allowed	5.74	5.68	7.61
Outliers	0.25	0.74	0.87

$R_{\text{merge}} = \sum hkl \sum i |I_i(hkl) - \langle I(hkl) \rangle| / \sum hkl \sum i I_i(hkl)$, where $\langle I(hkl) \rangle$ is the main value of $I(hkl)$.

$R_{\text{work}} = \sum ||F_{\text{obs}}| - |F_{\text{calc}}|| / \sum |F_{\text{obs}}|$, where F_{obs} and F_{calc} are observed and calculated structure

Supplementary table 1.

Data collection and refinement statistics

factors.
The free R factor was calculated using 5% of reflections omitted from the refinement.
Highest resolution shell is shown in parenthesis

Supplementary table 1. (continued)

Name	Forward primers (5'-3')	Reverse primers (5'-3')
<i>CsAlaDC</i>	GGAATTCCATATGACAACAAGCCTCAC CATCACGG	GATCTCGAGTCATTTATGAAGATCACAA TCACAATTCTCACTTC
<i>AtSerDC</i>	GGAATTCCATATGACCACCAGCCTGGC CG	GATCTCGAGTTATTTATGGGCCGGACA AATGCAATTATTG
<i>CsSerDC</i>	ATGGTGGGGAAGTGTGGGGT	TCACTTGTGTAGTGCACAAAGACAGT

Supplementary table 2.

Primers used for gene cloning.

Name	Forward primers (5'-3')	Reverse primers (5'-3')
<i>CsAlaDC</i>	CACTGTGATGGGGCTCTGTT	TGTTATCTGGACACCGCACG
<i>AtSerDC</i>	TCACTCGCTGTAACGGAACC	AACTGACCAAGCGCACCATA
<i>NbGAPDH</i>	AGCTCAAGGGAATTCTCGATG	AACCTTAACCATGTCATCTCCC

Supplementary table 3.

Primers used for real-time PCR.

References

1. Drew L (2019) **The growth of tea** *Nature* **566**:S2–S4
2. Wheeler D. S., Wheeler W. J (2004) **The medicinal chemistry of tea** *Drug Dev. Res* **61**:45–65
3. Xia E. H., et al. (2017) **The tea tree genome provides insights into tea flavor and independent evolution of caffeine biosynthesis** *Mol. Plant* **10**:866–877
4. Yang C. S., Hong J. G (2013) **Prevention of Chronic Diseases by Tea: Possible Mechanisms and Human Relevanc** *Annu. Rev. Nut* **33**:161–181
5. Xia E., et al. (2020) **The reference genome of tea plant and resequencing of 81 diverse accessions provide insights into its genome evolution and adaptation** *Mol. Plant* **13**:1013–1026
6. Juneja L. R., Chu D. C., Okubo T., Nagato Y., Yokogoshi H (1999) **L-theanine - a unique amino acid of green tea and its relaxation effect in humans** *Trends Food Sci. Technol* **10**:425–425
7. Lin J. Z., et al. (2022) **Effect of the presence of stem on quality of oolong tea** *Foods* **11**
8. Haskell C. F., Kennedy D. O., Milne A. L., Wesnes K. A., Scholey A. B (2008) **The effects of L-theanine, caffeine and their combination on cognition and mood** *Biol. Psychol* **77**:113–122
9. Kakuda T., et al. (2000) **Protective effect of gamma-glutamylethylamide (theanine) on ischemic delayed neuronal death in gerbils** *Neuro. Sci. Lett* **289**:189–192
10. Liu Q., Duan H., Luan J., Yagasaki K., Zhang G (2009) **Effects of theanine on growth of human lung cancer and leukemia cells as well as migration and invasion of human lung cancer cells** *Cytotechnology* **59**:211–217
11. Lu K., et al. (2004) **The acute effects of L-theanine in comparison with alprazolam on anticipatory anxiety in humans** *Hum. Psychopharmacol* **19**:457–465
12. Takagi Y., et al. (2010) **Combined administration of (L)-cystine and (L)-theanine enhances immune functions and protects against influenza virus infection in aged mice** *J. Vet. Med. Sci* **72**:157–165
13. Zheng G., Sayama K., Okubo T., Juneja L. R., Oguni I (2004) **Anti-obesity effects of three major components of green tea, catechins, caffeine and theanine, in mice** *In Vivo* **18**:55–62
14. Lin S., et al. (2023) **Theanine metabolism and transport in tea plants (*Camellia sinensis* L.): advances and perspectives** *Crit. Rev. Biotechnol* **43**:327–341
15. Deng W., Ogita S., Ashihara H (2008) **Biosynthesis of theanine (γ -ethylamino-L-glutamic acid) in seedlings of *Camellia sinensis*** *Phytochem. Lett* **1**:115–119
16. Deng W.-W., Ogita S., Ashihara H (2009) **Ethylamine content and theanine biosynthesis in different organs of *camellia sinensis* seedlings** *Z. Naturforsch. C* **64**:906–906

17. Cheng S., et al. (2019) **Differential accumulation of specialized metabolite L-theanine in green and albino-induced yellow tea (*Camellia sinensis*) leaves** *Food Chem* **276**:93–100
18. Takeo T (1974) **L-alanine as a precursor of ethylamine in *Camellia sinensis*** *Phytochem* **13**:1401–1406
19. Zhu B., et al. (2021) **CsAlaDC and CsTSI work coordinately to determine theanine biosynthesis in tea plants (*Camellia sinensis* L.) and confer high levels of theanine accumulation in a non-tea plant** *Plant Biotechnol. J* **19**:2395–2397
20. Bai P., et al. (2021) **Biochemical characterization of specific Alanine Decarboxylase (ADC) and its ancestral enzyme Serine Decarboxylase (SDC) in tea plants (*Camellia sinensis*)** *BMC Biotechnol* **21**
21. Han S. W., Shin J. S (2022) **Aromatic L-amino acid decarboxylases: mechanistic features and microbial applications** *Appl. Microbiol. Biotechnol* **106**:4445–4458
22. Berger M., Gray J. A., Roth B. L (2009) **The expanded biology of serotonin** *Annu. Rev. Med* **60**:355–366
23. Khan M. Z., Nawaz W (2016) **The emerging roles of human trace amines and human trace amine-associated receptors (hTAARs) in central nervous system** *Biomed. Pharmacother* **83**:439–449
24. Ng J., Papandreou A., Heales S. J., Kurian M. A (2015) **Monoamine neurotransmitter disorders--clinical advances and future perspectives** *Nat. Rev. Neurol* **11**:567–584
25. Rontein D., et al. (2001) **Plants synthesize ethanolamine by direct decarboxylation of serine using a pyridoxal phosphate enzyme** *J. Biol. Chem* **276**:35523–35529
26. Gibellini F., Smith T. K (2010) **The Kennedy pathway-De novo synthesis of phosphatidylethanolamine and phosphatidylcholine** *IUBMB Life* **62**:414–428
27. Mudd S. H., Datko A. H (1989) **Synthesis of ethanolamine and its regulation in *Lemna paucicostata*** *Plant Physiol* **91**:587–597
28. Zinser E., et al. (1991) **Phospholipid synthesis and lipid composition of subcellular membranes in the unicellular eukaryote *Saccharomyces cerevisiae*** *J. Bacteriol* **173**:2026–2034
29. Yunus I. S., Liu Y.-c., Nakamura Y. (2016) **The importance of SERINE DECARBOXYLASE1 (SDC1) and ethanolamine biosynthesis during embryogenesis of *Arabidopsis thaliana*** *Plant J* **88**:559–569
30. Paiardini A., Giardina G., Rossignoli G., Voltattorni C. B., Bertoldi M (2017) **New insights emerging from recent investigations on human group II pyridoxal 5'-phosphate decarboxylases** *Curr. Med. Chem* **24**:226–244
31. Torrens-Spence M. P., et al. (2020) **Structural basis for divergent and convergent evolution of catalytic machineries in plant aromatic amino acid decarboxylase proteins** *Proc. Natl. Acad. Sci. U. S. A* **117**:10806–10817

32. Hagihara R., Ohno S., Hayashi M., Tabata K., Endo H (2021) **Production of L-theanine by *Escherichia coli* in the absence of supplemental ethylamine** *Appl. Environ. Microbiol* **87**:e00031–21
33. Yamamoto S., Wakayama M., Tachiki T (2007) **Characterization of theanine-forming enzyme from *Methylovorus mays* no. 9 in respect to utilization of theanine production** *Biosci. Biotechnol. Biochem* **71**:545–552
34. Holm L (2022) **Dali server: structural unification of protein families** *Nucleic. Acids Res* **50**:W210–W215
35. Okawa A., et al. (2021) **Structural basis for substrate specificity of L-methionine decarboxylase** *Protein Sci* **30**:663–677
36. Fenalti G., et al. (2007) **GABA production by glutamic acid decarboxylase is regulated by a dynamic catalytic loop** *Nat. Struct. Mol. Biol* **14**:280–286
37. Gayathri S. C., Manoj N (2019) **Structural insights into the mechanism of internal aldimine formation and catalytic loop dynamics in an archaeal Group II decarboxylase** *J. Struct. Biol* **208**:137–151
38. Torrens-Spence M. P., von Guggenberg R., Lazear M., Ding H., Li J. (2014) **Diverse functional evolution of serine decarboxylases: identification of two novel acetaldehyde synthases that uses hydrophobic amino acids as substrates** *BMC Plant Biol* **14**
39. Qu L., Gu S., Zhang J., Zhao C., Deng X (2021) **Determination of 18 amino acids in three different kinds of milk powder by ultra performance liquid chromatography coupled with pre-column derivatization** *Chin. J. Chromatogr* **39**:472–477
40. Madeira F., et al. (2022) **Search and sequence analysis tools services from EMBL-EBI in 2022** *Nucleic Acids Res* **50**:W276–W279
41. Dong C., et al. (2020) **Theanine transporters identified in tea plants (*Camellia sinensis* L** *Plant J* **101**:57–70
42. Schmittgen T. D., et al. (2000) **Quantitative reverse transcription-polymerase chain reaction to study mRNA decay: comparison of endpoint and real-time methods** *Anal. Biochem* **285**:194–204
43. Robert X., Gouet P (2014) **Deciphering key features in protein structures with the new ENDscript server** *Nucleic Acids Res* **42**:W320–W324
44. Camacho C., et al. (2009) **BLAST+: architecture and applications** *BMC Bioinf* **10**
45. Sievers F., Higgins D. G (2018) **Clustal Omega for making accurate alignments of many protein sequences** *Protein Sci* **27**:135–145
46. Larkin M. A., et al. (2007) **Clustal W and clustal X version 2.0** *Bioinformatics* **23**:2947–2948
47. Yu G., Lam T. T.-Y., Zhu H., Guan Y (2018) **Two methods for mapping and visualizing associated data on phylogeny using ggtree** *Mol. Biol. Evol* **35**:3041–3043

Article and author information

Hao Wang

Department of Life Sciences and Medicine, University of Science and Technology of China, Hefei, China

ORCID iD: [0009-0003-8119-9882](https://orcid.org/0009-0003-8119-9882)

Biyang Zhu

State Key Laboratory of Tea Plant Biology and Utilization, Anhui Agricultural University, Hefei, Anhui 230036, China

Siming Qiao

State Key Laboratory of Tea Plant Biology and Utilization, Anhui Agricultural University, Hefei, Anhui 230036, China

Chunxia Dong

State Key Laboratory of Tea Plant Biology and Utilization, Anhui Agricultural University, Hefei, Anhui 230036, China

Xiaochun Wan

State Key Laboratory of Tea Plant Biology and Utilization, Anhui Agricultural University, Hefei, Anhui 230036, China

ORCID iD: [0000-0001-5294-786X](https://orcid.org/0000-0001-5294-786X)

Weimin Gong

Department of Life Sciences and Medicine, University of Science and Technology of China, Hefei, China

For correspondence: wgong@ustc.edu.cn

ORCID iD: [0000-0002-6723-1298](https://orcid.org/0000-0002-6723-1298)

Zhaoliang Zhang

State Key Laboratory of Tea Plant Biology and Utilization, Anhui Agricultural University, Hefei, Anhui 230036, China

For correspondence: zhlzhang@ahau.edu.cn

ORCID iD: [0000-0001-6615-1598](https://orcid.org/0000-0001-6615-1598)

Copyright

© 2023, Wang et al.

This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

Editors

Reviewing Editor

Amie Boal

Pennsylvania State University, United States of America

Reviewer #1 (Public Review):

In this study, the structural characteristics of plant AlaDC and SerDC were analyzed to understand the mechanism of functional differentiation, deepen the understanding of substrate specificity and catalytic activity evolution, and explore effective ways to improve the initial efficiency of theanine synthesis.

On the basis of previous solid work, the authors successfully obtained the X-ray crystal structures of the precursors of theanine synthesis-CsAlaDC and AtSerDC, which are key proteins related to ethylamine synthesis, and found a unique zinc finger structure on these two crystal structures that are not found in other Group II PLP- dependent amino acid decarboxylases. Through a series of experiments, it is pointed out that this characteristic zinc finger motif may be the key to the folding of CsAlaDC and AtSerDC proteins, and this discovery is novel and prospective in the study of theanine synthesis.

In addition, the authors identified Phe106 of CsAlaDC and Tyr111 of AtSerDC as key sites of substrate specificity by comparing substrate binding regions and identified amino acids that inhibit catalytic activity through mutation screening based on protein structure. It was found that the catalytic activity of CsAlaDC L110F/P114A was 2.3 times higher than that of CsAlaDC. At the same time, CsAlaDC and AtSerDC substrate recognition key motifs were used to carry out evolutionary analysis of the protein sequences that are highly homologous to CsAlaDC in embryos, and 13 potential alanine decarboxylases were found, which laid a solid foundation for subsequent studies related to theanine synthesis.

In general, this study has a solid foundation, the whole research idea is clear, the experimental design is reasonable, and the experimental results provide strong evidence for the author's point of view. Through a large number of experiments, the key links in the theanine synthesis pathway are deeply studied, and an effective way to improve the initial efficiency of theanine synthesis is found, and the molecular mechanism of this way is expounded. The whole study has good novelty and prospectivity, and sheds light on a new direction for the efficient industrial synthesis of theanine.

Reviewer #2 (Public Review):**Summary:**

The manuscript focuses on the comparison of two PLP-dependent enzyme classes that perform amino acyl decarboxylations. The goal of the work is to understand the substrate specificity and factors that influence the catalytic rate in an enzyme linked to theanine production in tea plants.

Strengths:

The work includes x-ray crystal structures of modest resolution of the enzymes of interest. These structures provide the basis for the design of mutagenesis experiments to test hypotheses about substrate specificity and the factors that control catalytic rate. These ideas are tested via mutagenesis and activity assays, in some cases both in vitro and in plants.

Weaknesses:

The manuscript could be more clear in explaining the contents of the x-ray structures and how the complexes studied relate to the reactant and product complexes. The structure and mechanism section would also be strengthened by including a diagram of the reaction mechanism and including context about reactivity. As it stands, much of the structural results section consists of lists of amino acids interacting with certain ligands without any explanation of why these interactions are important or the role they play in catalysis. The

experiments testing the function of a novel Zn(II)-binding domain also have serious flaws. I don't think anything can be said at this point about the function of the Zn(II) due to a lack of key controls and problems with experimental design.

Reviewer #3 (Public Review):

In the manuscript titled "Structure and Evolution of Alanine/Serine Decarboxylases and the Engineering of Theanine Production," Wang et al. solved and compared the crystal structures of Alanine Decarboxylase (AlaDC) from *Camellia sinensis* and Serine Decarboxylase (SerDC) from *Arabidopsis thaliana*. Based on this structural information, the authors conducted both in vitro and in vivo functional studies to compare enzyme activities using site-directed mutagenesis and subsequent evolutionary analyses. This research has the potential to enhance our understanding of amino acid decarboxylase evolution and the biosynthetic pathway of the plant-specialized metabolite theanine, as well as to further its potential applications in the tea industry.