Published in final edited form as:
ACS Chem Biol. 2013 October 18; 8(10): . doi:10.1021/cb400335k.

Discovery of a Substrate Selectivity Motif in Amino Acid Decarboxylases Unveils a Taurine Biosynthesis Pathway in Prokaryotes

Giulia Agnello[†], **Leslie L. Chang**[‡], **Candice M. Lamb**[‡], **George Georgiou**[‡,§,†,*], and **Everett M. Stone**[‡,*]

[‡]Departments of Biomedical and Chemical Engineering, University of Texas, Austin, Texas 78712

[§]Section of Molecular Genetics and Microbiology, University of Texas, Austin, Texas 78712

[†]Institute for Cellular and Molecular Biology, University of Texas, Austin, Texas 78712

Abstract

Taurine, the most abundant free amino acid in mammals, with many critical roles such as neuronal development, had so far only been reported to be synthetized in eukaryotes. Taurine is the major product of cysteine metabolism in mammals, and its biosynthetic pathway consists of cysteine dioxygenase and cysteine sulfinic acid decarboxylase (hCSAD). Sequence, structural, and mutational analyses of the structurally and sequentially related hCSAD and human glutamic acid decarboxylase (hGAD) enzymes revealed a three residue substrate recognition motif ($X_1aa_{19}X_2aaX_3$), within the active site that is responsible for coordinating their respective preferred amino acid substrates. Introduction of the cysteine sulfinic acid (CSA) motif into hGAD (hGAD-S192F/N212S/F214Y) resulted in an enzyme with a >700 fold switch in selectivity towards the decarboxylation of CSA over its preferred substrate, L-glutamic acid. Surprisingly, we found this CSA recognition motif in the genome sequences of several marine bacteria, prompting us to evaluate the catalytic properties of bacterial amino acid decarboxylases that were predicted by sequence motif to decarboxylate CSA but had been annotated as GAD enzymes. We show that CSAD from *Synechococcus sp. PCC 7335* specifically decarboxylated CSA and that the bacteria accumulated intracellular taurine. The fact that CSAD homologues exist in certain bacteria and are frequently found in operons containing the recently discovered bacterial cysteine dioxygenases that oxidize L-cysteine to CSA, supports the idea that a *bona fide* bacterial taurine biosynthetic pathway exists in prokaryotes.

Cysteine sulfinic acid decarboxylase (CSAD, EC 4.1.1.29), a member of the aspartate aminotransferase fold type I superfamily (AAT_I), catalyzes the rate limiting step in the biosynthesis of taurine (2-aminoethanesulfonic acid), the most abundant non-peptidic amino acid in the human body and the main end-product of cysteine metabolism in mammals (1). Taurine is involved in a variety of physiological functions (2) including neurotransmission (3), neuromodulation, neuroprotection (4), osmoregulation (5), modulation of protein phosphorylation (6), trophism in the development of central nervous system (CNS), regulation of calcium homeostasis (3), and bile salt formation in the liver. The synthesis of this amino acid starts with the oxidation of L-cysteine (L-Cys) to cysteine sulfinic acid (CSA) by cysteine dioxygenase (CDO, EC 1.13.11.20), followed by CSAD-mediated decarboxylation to hypotaurine, and subsequent oxidation to taurine (7) (Figure 1). CSAD

[*]To whom correspondence should be addressed: George Georgiou, PhD, Chemical Engineering/UT Austin, 1 University Station C0400, Austin, TX 78712-0231 Ph: (512)-471-6975. gg@che.utexas.edu And Everett M. Stone, PhD, Biomedical Engineering/UT Austin, 1 University Station C0400, Austin, TX 78712-0231 Ph: (512)-232-4103, stonesci@.utexas.edu.

also decarboxylates the oxidized form of CSA, cysteic acid (CA), providing another route for taurine synthesis. Formation of taurine was initially thought to be catalyzed by L-glutamic acid (L-Glu) decarboxylase (GAD, EC 4.1.1.15). However subsequent studies (9) revealed that the main function of GAD is the synthesis of γ-aminobutyric acid (GABA). It is now established that in mammals, CSA/CA decarboxylation and L-Glu decarboxylation are in fact catalyzed by different enzymes (8). Specifically, CSAD decarboxylates CSA and CA but has no detectable activity towards L-Glu (9). Commensurate with its physiological importance, the synthesis of taurine is best documented in higher eukaryotes. Plants do not appear to commonly synthesize taurine (10) although notable exceptions exist in certain seaweeds (11) and cacti (12) and its presence in lower eukaryotes was only recently been reported in the fungus *Yarrowia lipolytica* (13). There are no reports of a taurine biosynthetic pathway in bacteria or archea and the current annotations of prokaryotic and archeal genomes do not indicate the existence of CSAD homologues.

Similarities in the amino acid sequence (~46 % amino acid identity) of GAD and CSAD suggest that they derive from a common ancestor. CSAD is thought to be more distant from this common ancestor than GAD, consistent with its narrower substrate specificity (14). GAD enzymes are widely prevalent amongst many eukaryotes and prokaryotes but serve different functions in different organisms. In mammals, GAD exists in two isoforms of different molecular weight: GAD-65 kDa and GAD-67 kDa. The two GAD isozymes play a critical role in the function of the CNS as their product, GABA, is an inhibitory neurotransmitter that regulates important physiological processes including neurogenesis, movement and tissue development (15–18). In bacteria, GAD is involved in maintaining the cellular pH at physiological levels under acid stress (19–22). Neither the taurine biosynthetic enzymes nor taurine itself were believed to exist in prokaryotes or archea, although many species of bacteria can use taurine as a source of sulfur, carbon or nitrogen (23).

Despite the importance of taurine in human health there has been little biochemical characterization of human CSAD (hCSAD), although a recent advance is the deposition of the hCSAD structure (PDB:2JIS). Comparison of the structure of hGAD-67 kDa (PDB: 2OKJ) (24) with the structure of hCSAD (PDB:2JIS), recently deposited by the Structural Genomics Consortium, shows a high degree of structural homology (RMS = 0.68 Å) with a largely conserved active site. However, the F94, S114, and Y116 residues in the active site pocket of hCSAD that accommodates the amino acid substrate side chain are occupied by S192, N212, and F214, respectively in the hGAD-67 kDa isoform. We hypothesized that the identity of the amino acids occupying these three positions in the active site might dictate the substrate selectivity of GAD and CSAD. To test this idea we constructed a set of single, double, and triple amino acid substitutions to replace the nonconserved residues in the hCSAD active site with the respective amino acids in hGAD-67 kDa and *vice versa*. We show that GAD S192F/N212S/F214Y displays >700 fold switch in catalytic selectivity towards the decaboxylation of CSA over L-Glu, whereas the Y116F mutation enables CSAD to decarboxylate L-Glu, an activity not observed with the wild type (wt) CSAD enzyme. Examination of the amino acid occupancy of these three positions within the AAT_I family further suggests that they comprise a substrate recognition motif ($X_1aa_{19}X_2aaX_3$) which may be important not only for the discrimination of L-Glu and CSA, but also for the substrate selectivity in other AAT_I family members. A search for the CSA/CA recognition motif ($F_1aa_{19}S_2aaY_3$) unexpectedly revealed that certain bacteria encode AAT_I enzymes with a CSA recognition motif that had been previously annotated as GAD enzymes. This was surprising as the decarboxylation of CSA/CA was believed to be exclusive to the eukaryotic domain of life (25). Testing of recombinantly expressed proteins of two of these putatively misannotated enzymes from *Polaribacter irgensii 23-P* and *Synechococcus sp. PCC 7335* demonstrated that they selectively catalyze the decarboxylation of CSA but not L-Glu. Moreover we showed that taurine accumulates in the

cytosol of *Synechococcus sp. PCC 7335* under physiological conditions. We found that bacterial CSAD-like genes are localized within putative operons containing CDO homologues (CDO isozymes were recently discovered in bacteria (25)), indicating that bacteria encode the catalytic machinery for the biosynthesis of taurine and hence that the production of this amino acid is not limited to eukaryotes but is instead distributed throughout the tree of life.

RESULTS

Expression and Biochemical Characterization of hCSAD

The gene for hCSAD was fused to nucleotides encoding an N-terminal His_6 affinity tag and expressed from a T7 promoter in *E. coli* (DE3) Rosetta(2) cells. Western blot analysis revealed the expression of a ~55kDa His_6-tagged protein in both the soluble and insoluble fractions. Soluble hCSAD was isolated by immobilized metal ion affinity chromatography (IMAC) purification with a yield of ~4–6 mg/L culture. Incubation of recombinant hCSAD with CSA under physiological conditions (37 °C, pH 7.4) resulted in the formation of two products that could be readily resolved by high performance liquid chromatography (HPLC) and were shown to correspond to hypotaurine and its oxidized form, taurine. However, hCSAD could not catalyze the decarboxylation of L-Glu as no GABA formation was observed even after 12 hrs incubation with 0–4mM of L-Glu. hCSAD catalyzed the decarboxylation of CSA and CA with well-behaved Michaelis-Menten kinetics giving k_{cat}/K_M values of 3550 ± 880 $M^{-1}s^{-1}$ and 122 ± 46 $M^{-1}s^{-1}$, respectively (Table 1). The enzyme was found to be stable in buffer at 37°C, displaying a deactivation $t_{1/2}$ of ~84 hrs. Circular dichroism analysis following the change in molar ellipticity as a function of temperature, revealed that the enzyme has a T_M ~ 70 °C.

Expression and Biochemical Characterization of hGAD-67 kDa

The hGAD-67 kDa enzyme was similarly expressed with an N-terminal His_6 fusion tag from the T7 promoter in *E. coli* (DE3) Rosetta(2). Consistent with earlier reports (26, 27) hGAD-67 kDa was found to be largely insoluble in bacteria. However, expression in the *E. coli* strain NiCo21 (DE3), which is optimized for IMAC purification purposes (28) resulted in higher soluble expression as determined by Western blot analysis and thus GAD-67 kDa could be purified with a yield of ~2 mg/L culture. Earlier studies had reported k_{cat}/K_M values for the decarboxylation of L-Glu by hGAD-67 kDa that varied between 3,080 $M^{-1}s^{-1}$ at pH 6.5 and 43–50 °C (26) and 226 $M^{-1}s^{-1}$ at pH 7.2 and 22 °C (29). We found that recombinant hGAD-67 kDa displays a k_{cat}/K_M = 790 ± 180 $M^{-1}s^{-1}$ with L-Glu as the substrate at 37 °C, pH 7.4. The enzyme also displayed significant decarboxylation activity with CSA (k_{cat}/K_M = 470 ± 160 $M^{-1}s^{-1}$) whereas the decarboxylation of CA occurred with a >20 fold reduced catalytic efficiency (k_{cat}/K_M = 34.0 ± 0.4 $M^{-1}s^{-1}$).

Expression and kinetic analysis of hCSAD and hGAD-67 kDa variants

Overlays of the hCSAD (PDB:2JIS) and the GABA-bound hGAD-67 kDa (PDB:2OKJ) crystal structures (Figure 2) revealed a high degree of structural homology with an α-carbon RMS = 0.68 Å. Notably, the active site residues coordinating and orienting the pyridoxal-phosphate (PLP) cofactor were found to be identical; the only differences in the composition of the active sites in the two enzymes were observed in residues that lie within the pocket accommodating the substrate side chain (F94, S114, and Y116 in hCSAD; S192, N212, and F214 in hGAD-67 kDa). Multiple sequence alignments of > 900 enzyme sequences from the AAT_I family (36% average pairwise amino acid homology to hCSAD and hGAD-67 kDa) suggested that these three active site residues form a linear substrate recognition motif ($X_1aa_{19}X_2aaX_3$, where $X_\#$ indicates the active site amino acid 1, 2 & 3 involved in substrate recognition, and aa indicates any amino acid). To evaluate if the hCSAD ($F_1aa_{19}S_2aY_3$) and

hGAD-67 kDa ($S_1aa_{19}N_2aF_3$) motifs play a role in substrate selectivity we constructed a set of hCSAD and hGAD-67 kDa single-, double- and triple-mutant variants where the putative motif recognition elements were interchanged between hCSAD and hGAD-67 kDa (Table 1 & Figure 2). Steady-state kinetics of the various hCSAD and hGAD-67 kDa variants with CSA, CA or L-Glu were determined at physiological conditions (37 °C, PBS buffer at pH 7.4) (M&M) (Table 1). As discussed above, wt hGAD-67 kDa displays only a 1.7 fold catalytic selectivity for L-Glu over CSA [$(k_{cat}/K_M)_{L\text{-}Glu}/ k_{cat}/K_M)_{CSA}$] and a 23-fold preference for L-Glu over CA. The kinetic analyses of the single point variants show that within the $X_1aa_{19}X_2aaX_3$ motif, the identity of X_3 is the most critical determinant of L-Glu or CSA/CA selectivity. The hGAD-67 kDa -F214Y variant (whereby the amino acid in the X_3 position of GAD had been switched to the respective amino acid in hCSAD) displayed a ~3-fold increase in k_{cat}/K_M (1,240 ± 50 $M^{-1}s^{-1}$) for CSA and ~ 7-fold decrease in k_{cat}/K_M for L-Glu (107 ± 3 $M^{-1}s^{-1}$) compared to the parental enzyme. More remarkably, the equivalent amino acid substitution in the R_3 position of hCSAD in CSAD-Y116F enabled the enzyme to decarboxylate L-Glu (k_{cat}/K_M = 2.2 ± 0.8 $M^{-1}s^{-1}$), an activity absent in the parental protein. CSAD-Y116F also displayed a ~7-fold decrease in the k_{cat}/K_M for CSA and a ~2-fold increase in k_{cat}/K_M for CA. These results highlight the crucial role of a phenolic side chain at motif position X_3 in coordinating the methane-sulfinate group of CSA. Presumably the larger propionate side chain of L-Glu clashes with the phenolic side chain at the X_3 position of CSAD resulting in an unproductive conformation at the active site. The most dramatic change in substrate selectivity was observed when the full hCSAD motif was introduced into hGAD-67 kDa. The resulting GAD-67 kDa (S192F/N212S/F214Y) enzyme showed a >700-fold switch in substrate specificity [$(k_{cat}/K_M)_{CSA}/(k_{cat}/K_M)_{L\text{-}Glu}$] with the catalytic activity for the decarboxylation of L-Glu reduced from 790 ± 180 $M^{-1}s^{-1}$ for the parental enzyme to 0.84 ± 0.02 $M^{-1}s^{-1}$ for the mutant, and k_{cat}/K_M = 625 ± 60 $M^{-1}s^{-1}$ for CSA decarboxylation. These results suggest that the CSAD motif ($F_1aa_{19}S_2aaY_3$) plays a key role in the recognition of CSA and specifically that a phenolic side chain at X_3 is significant in binding CSA.

Glutamic acid inhibition of hCSAD

We further examined whether L-Glu is excluded from the active site of hCSAD by incubating the wt enzyme with CSA in combination with various concentrations of L-Glu (0–9 mM). CSA turnover was inhibited with an apparent IC_{50} value of 1.1 ± 0.1 mM suggesting that although the parental CSAD enzyme doesn't have detectable GAD activity, most likely it binds L-Glu. UV-vis spectroscopic analyses of the hCSAD holoenzyme showed typical PLP-internal aldimine absorption shoulders centered around 335 nm and 410 nm. Scans of hCSAD in a steady-state reaction with CSA revealed a 1.5–2 fold increase in amplitude of the 335 nm shoulder indicative of a reaction intermediate formed by a CSA-PLP-external aldimine. The addition of L-Glu did not appreciably alter the PLP absorption bands of the hCSAD holoenzyme either with or without substrate (SI Figure 1). This suggests that L-Glu inhibits hCSAD non-competitively as it had been demonstrated in an earlier study of partially purified rat brain CSAD (30).

Discovery of CSAD enzymes and taurine synthesis in bacteria

A bioinformatics analysis (Linnaeus Blast Geneious 4.8.5 (31)) of hCSAD homologues revealed a significant number of prokaryotic and archeal AAT_I superfamily enzymes that contained the ($F_1aa_{19}S_2aaY_3$) CSA recognition element motif, within eubacterial phyla, namely *alphaproteobacteria,* in *cyanobacteria*, *bacteroidetes*, in distinct sequences from the marine metagenome, and one from an uncultured marine *archeon*. Interestingly, the CSAD homologues from Polaribacter irgensii 23-P, Synechococcus sp. PCC 7335, and from *Flavobacteriales ALC-1* are located in putative operons with open reading frames encoding proteins having high homology to CDO (Figure 3,) (We note there are other genes that may

lie within these operons such as a predicted methyltransferase in *Polaribacter* and a predicted cobalt transporter in *Moorea product 3L).* CDO is required to oxidize L-Cys to CSA, the substrate of CSAD isozymes, suggesting the existence of a taurine synthetic pathway in bacteria. We explored this hypothesis further by analyzing the substrate specificity of CSAD-like enzymes from *Polaribacter irgensii 23-P* (PiCSAD) and from *Synechococcus sp. PCC 7335* (SsCSAD). Synthetic PiCSAD and SsCSAD codon-optimized genes were constructed, fused to a 5′ sequence encoding a 5′-His_6 affinity tag and expressed from the T7 promoter. The psychrophilic *Polaribacter* enzyme, PiCSAD could only be expressed in *E. coli* BL21(DE3) at 16 °C and even under those conditions the yield of soluble protein was very low (SI Figure 2). Nonetheless, partially purified PiCSAD incubated with CSA, CA or L-Glu in PBS at pH 7.4 and 4 °C, showed decarboxylation of CSA but no detectable activity towards CA or L-Glu. The *cyanobacteria* SsCSAD expressed well at 25 °C in *E. coli* BL21(DE3) and was purified to > 99 % homogeneity as assessed by SDS-PAGE (SI Figure 3) with a yield of 3 mg/L culture. Steady-state kinetic analyses over a (0–4mM) range of CSA and CA concentrations revealed that the enzyme followed Michaelis-Menten kinetics with k_{cat}/K_M values of 1,530 ± 870 $M^{-1}s^{-1}$ and 139 ± 27 $M^{-1}s^{-1}$ for CSA and CA respectively (SI Figure 4). No apparent activity towards L-Glu could be detected.

To determine if taurine could be produced *in vivo, Synechococcus sp. PCC 7335* (ATCC® 29403™) was obtained and grown in the presence or absence of CSA. Examination by HPLC of lysates from *Synechococcus sp. PCC 7335* cultured with CSA showed both consumption of CSA and accumulation of a new peak matching the retention time of taurine that was not detected in cultures grown without CSA (SI Figure 5) (We were unable to assign a clear peak from the lysates to hypotaurine, suggesting that it had all oxidized to taurine or was obscured by neighboring peaks). Analysis by electrospray ionization mass spectrometry (ESI-MS) of either taurine standards or the CSA cultured *Synechococcus* lysates revealed a 327 *m/z* peak corresponding to the calculated mass of the *o*-phthalaldehyde (OPA)/mercaptopropionic-acid/taurine derivative. Collision induced dissociation (CID) of these 327 *m/z* peaks yielded an identical fragmentation pattern for both the taurine standard and the corresponding peak from the *Synechococcus* lysates (Figure 4).

DISCUSSION

Enzymes in the AAT_I superfamily share a highly conserved structural fold, which enables the catalysis of a diverse set of PLP-dependent reactions, including transaminations, racemizations, and decarboxylations. Using a combination of phylogenetic and biochemical analyses we found that several of the AAT_I superfamily amino acid decarboxylases have a distinct sequence motif ($X_1aa_{19}X_2aaX_3$) comprised of three amino acids that in the quarternary structure are part of the substrate binding pocket and dictate recognition of various amino acid substrates. In particular we define a CSAD recognition motif ($F_1aa_{19}S_2aaY_3$) where the Y_3 phenol side chain appears to be highly selective for the sp3 hybridization state of the CSA methane-sulfinate group and to discriminate against the one methylene longer sp2-hybridized L-Glu side chain. We speculate that a phenolic side chain at the X_3 position can form a hydrogen bond with the lone pair of electrons on the sulfur atom of CSA. Consistent with this hypothesis, considering CA as a substrate, where the sulfur atom lone pair is now part of an oxygen-sulfur covalent bond in the methane-sulfonate group, we observed a 5.5-fold drop in k_{cat} and 5.5 fold increase in K_M compared to the decarboxylation of CSA. Similarly, in the hCSAD-Y116F (X_3) variant, the k_{cat}/K_M for CSA is 7-fold lower, the k_{cat}/K_M for CA is slightly improved and L-Glu can be accepted as a substrate. Likewise, introduction of a phenolic side chain at the X_3 position into the GAD motif ($S_1aa_{19}N_2aa(F\rightarrow Y)_3$) potentiated a 12-fold selectivity of CSA over L-Glu, and the complete CSAD motif (hGAD-S192F,N212S,F214Y) resulted in a highly selective

CSA-degrading enzyme with a > 700-fold increase in k_{cat}/K_M towards CSA (625 ± 58 $M^{-1}s^{-1}$) as opposed to L-Glu (0.84± 0.02 $M^{-1}s^{-1}$). Similarities in amino acid sequence (~46 % amino acid identity) and activity have suggested that GAD and CSAD may be derived from a common ancestor (14) and that CSAD may have arisen through a gene duplication event. Our experiments demonstrate a simple mechanism by which the ancestral GAD enzyme likely evolved into a highly specialized CSA-degrading enzyme, possibly to meet high taurine requirements without producing excessive amounts of GABA.

The motif we identified is also germane to recent reports on the specific CSA decarboxylase activity of an enzyme from *Aedes aegypti,* aspartate 1-decarboxylase (*Ae*ADC) and the of the human glutamic acid decarboxylase-like protein 1 (hGADL1) which was shown to display significant CSA decarboxylase activity (32, 33). Interestingly, the *Ae*ADC sequence contains the consensus CSAD recognition motif ($F_1aa_{19}S_2aaY_3$), while hGADL1 has a very similar recognition motif ($Y_1aa_{19}S_2aaY_3$), differing from CSAD only at the R_1 position. These examples further confirm that proteins with CSA decarboxylase activity share the same structural fingerprint and that it is possible to use the CSA recognition motif to identify other proteins in the AAT_I family with this specific activity. Similarly, variations within the recognition motif can be used to predict other substrate specificities. For example, multiple sequence alignment analyses show that bacterial diaminobutyric acid decarboxylase (DBAD) (EC 4.1.1.86) enzymes contain an $H_1aa_{19}S_2aaD_3$ motif (Figure 5). This recognition element for the substrate, L-2,4-diaminobutyric acid, is further observed in structural alignments with the DBAD structure from *Vibrio parahaemolyticus* (PDB:2QMA) and hCSAD (PDB:2JIS) where the aspartyl side chain at position X_3 likely forms a salt-bridge in coordinating the substrate's ethylamine side chain. It appears that with minimal changes in the primary sequence, this class of enzymes can plastically evolve to accommodate different amino acid substrates according to an organism needs.

It has been thought that only eukaryotes synthesize taurine (7). Surprisingly, we found AAT_I superfamily members containing the consensus CSAD motif in several bacterial genomes and indeed CSAD-like proteins from a *Polaribacter* and *Synechococcus* species displayed decarboxylation of CSA, but not of L-Glu. In particular the *Synechococcus* enzyme, exhibited comparable kinetics to hCSAD. These results further support our finding that the $F_1aa_{19}S_2aaY_3$ motif defines enzymes that recognize CSA as a substrate. (We note that the SsCSAD enzyme has a tryptophan residue at the first position of the motif ($W_1aa_{19}S_2aaY_3$) and hGADL1($Y_1X_{19}S_2XY_3$) has a tyrosine thus the motif for CSA recognition is best described by [$(F/W/Y)_1aa_{19}S_2aaY_3$].) We observed that SsCSAD is apparently transcribed as we detected accumulation of taurine in *Synechococcus* cultures incubated with CSA. Our observation that bacterial CSAD genes are found within operons encoding enzymes with homology to CDO suggests that a *bona fide* taurine biosynthetic pathway has evolved in eubacteria and is not exclusive to higher eukaryotes, as had been previously surmised. The evidence of active CSAD proteins in eubacteria raises questions about the potential physiological roles of hypotaurine and taurine in this phylum. Interestingly, high contents of hypotaurine, taurine and taurine derivatives have been reported for a variety of marine invertebrates that establish symbiotic associations with algae and bacteria (34, 35). In fact, for some marine invertebrates, taurine and hypotaurine are the dominant osmolytes and play a key role in counteracting hydrostatic pressure (36). We do not know if taurine is utilized as an osmolyte in bacteria, but it is interesting that the ~35 CSAD motif containing bacterial sequences that we identified are all encoded by marine dwelling organisms. Alternatively, taurine-conjugates could be important detergents and surfactants for bacteria or bacterial hosts, similar to taurine-conjugates in mammalian bile acids. Determining the physiological role that taurine has for the bacteria that can make it may help clear up the question of whether taurine production (easily) evolved in the ancient

eukaryote as our mutagenesis experiments suggest, or if it already existed in a progenitor organism.

MATERIALS AND METHODS

Supporting Information available including: Reagents, *E. coli* strains and media used; *Materials & Methods* of: cloning CSAD-like genes *Polaribacter irgensii 23-P* and *Synechococcus sp. PCC 7335*; SI Table 1 of oligonucleotides used for constructing gene variants, expression and purification of enzymes; methods for mass spectrometry of taurine and SI Figure 1 of UV-Vis absorption scans of hCSAD; SI Figure 2 of SDS-PAGE and western blot of purified recombinant *Polaribacter irgensii 23-P* and SI Figure 3 of an SDS-PAGE of *Synechococcus sp. PCC 7335* CSAD protein; SI Figure 4 of a plot of the initial rates of product formation from recombinant SsCSAD incubated with CSA and CA; SI Figure 5 of HPLC analyses of lysates from Synechococcus sp. PCC 7335 cultured with CSA.

This information is available free of charge via the Internet at http://pubs.acs.org.

Molecular Biology methods

Full-length cDNA encoding for the human cysteine sulfinic acid decarboxylase (hCSAD) and the human glutamic acid decarboxylase (hGAD-67 kDa) were isolated from the Gateway human ORF collection. The hCSAD clone was amplified with 5′ and 3′ oligonucleotides incorporating NheI and EcoRI restriction sites. The hGAD-67 kDa clone was amplified with 5′ and 3′ oligonucleotides incorporating EcoRI and NotI restriction sites. The PCR reaction mixture consisted of the two specific oligonucleotides, standard 5X Phusion HF Buffer and *Phusion* high fidelity DNA polymerase. The reaction was conducted at an initial temperature of 98°C for 30 sec, followed by 30 cycles at 98°C for 10 sec, 60°C for 10 sec in the case of hCSAD and 55°C for 10 sec in the case of hGAD-67 kDa, and 72°C for 40sec, followed by a final polishing step at 72°C for 7 min. The resulting DNA product was gel purified (Zymo), digested with the appropriate restriction enzymes and ligated into pET28a (Novagen), and transformed in *E. coli* MC1061 electrocompetent cells. The gene inserts in the resulting plasmids pCSAD and pGAD were sequenced and transformed into *E. coli* Rosetta (DE3) (pCSAD) or *E. coli* NiCo21(DE3) cells (pGAD) for subsequent expression. In addition, variants of hCSAD and hGAD-67 kDa were constructed using the pCSAD or pGAD as a template and the primer pairs listed in SI Table 1. Each mutant sequence was synthesized by overlap extension PCR consisting of two initial mutagenic PCR reactions (pairing the forward wt primer with the mutant reverse, and the forward mutant primer with the wt reverse), followed by a normal PCR reaction using the two products from the mutagenic PCR reactions and the outermost wt primers. The single point and multiple point mutants were cloned and expressed in a manner analogous to the corresponding wt variants.

Activity assays

Reactions of hCSAD and hGAD-67 kDa variants, and SsCSAD with CSA, CA, L-aspartic acid or L-glutamic acid (L-Glu), were performed using freshly purified enzyme or from aliquots thawed from frozen stocks stored at −80°C. The formation of the decarboxylated products (taurine, hypotaurine, β-alanine, γ-amino butyric acid (GABA)) was assessed by *o*-phthalaldehyde (OPA) derivatization and high performance liquid chromatography (HPLC) analysis essentially as described by Agilent Technologies (37). Reaction of enzymes (concentration ranging from 0.1 to 1 μM) with substrates (concentration from 0 to 5 × K_M) were conducted at 37°C in 100mM PBS, 10μM PLP (pH 7.4), (PiCSAD was reacted with CSA, CA, or L-Glu at 4°C) in a total volume of 50μL, and subsequently quenched with 5μL

of 1M hydrochloric acid to stop the reaction. An aliquot of the quenched reaction mixture was diluted 1:1 with activity buffer, mixed with a molar excess (relative to substrate) of OPA reagent and brought to a final volume of 100μL with 400 mM borate buffer, pH 10.2. The resulting solutions were analyzed by HPLC using an Agilent ZORBAX Eclipse AAA Column (C18 reverse phase, 5 μm, 4.6 mm _ 150 mm). The amount of product was calculated by the following equation, $[P]=[S]\frac{P_{area}}{P_{area}+S_{area}}$, where [*P*] is concentration of product, [*S*] is the concentration of substrate used in the reaction and P_{area} & S_{area} are the product and substrate peak areas. All reactions were conducted at least in triplicate, and the observed rates of product formation were either fit to the Michaelis-Menten equation or when substrate saturation was not observed were fit to a linear equation $v_0/[E] = (k_{cat}/K_M)*[S]$ using software from Kaleidagraph (Synergy).

L-Glu inhibition of hCSAD

The hCSAD enzyme (140 nM final) was incubated with a series of L-Glu concentrations ranging from 0–9 mM in the presence of 0.45 mM CSA at 37°C in 100mM PBS buffer, 10μM PLP (pH 7.4) for 30s and subsequently quenched with hydrochloric acid to stop the reaction. The samples were then analyzed by HPLC as described above to quantitate the amount of hypo/taurine produced and the resulting data was fit to an exponential decay equation to calculate an apparent IC_{50} value.

UV-vis spectroscopy of hCSAD reactions with CSA and L-Glu

Electronic absorption scans were obtained of the hCSAD holoenzyme alone, in the presence of 4 mM CSA, in the presence of 6.7 mM L-Glu and in the presence of a combination of 5.7 mM L-Glu and 2.9 mM CSA at 25 °C. Scans of mixtures containing CSA were obtained during the first 1–2 minutes of the reaction, and the resulting data was then normalized by the absorbance at 280 nm. Scans were also obtained for CSA and L-Glu alone to ensure there were no overlapping absorption bands.

Thermal Stability

A 2 μM sample of hCSAD in a 100 mM phosphate buffer, pH 7.4 was analyzed on a Jasco J-815 CD spectrometer. The change in molar ellipticity at 220 nm (θ_{220}) was monitored as a function of temperature from 20 – 95 °C. The fraction of denatured protein at each temperature was calculated by the ratio of $[\theta_{220}]/[\theta_{220}]_d$ where $[\theta_{220}]_d$ is the molar ellipticity of the completely unfolded protein. The resulting data was fit to a modified logistic equation to determine the thermal transition midpoint.

Culturing Polaribacter irgensii 23-P and Synechococcus sp. PCC 7335

Polaribacter irgensii 23-P was cultured according to the recommendations of ATCC at 4°C in marine broth (BD Difco) supplemented with or without 1mM CSA. *Synechococcus sp. PCC 7335* was cultured in ATCC Medium 957 with or without 1mM CSA at room temperature under fluorescent lighting. After turbid growth was achieved (2–3 weeks for *Polaribacter*, and 1–2 weeks for *Synechococcus*), 4 ml samples of culture were centrifuged to obtain cell pellets. The cell pellets were re-suspended in 500μL dH_2O, heated at 95°C for 10 minutes and lysed by sonication. After centrifugation at 16,000g for 10 min, the soluble fractions of the samples were concentrated by rotary evaporation, derivatized with OPA and analyzed by HPLC as described for the activity assays.

Mass Spectrometry of taurine produced by Synechococcus sp. 7335

Taurine standards and lysate samples from *Synechococcus sp. 7335* cultured with CSA were reacted with OPA/mercaptopropionic acid and chromatographed on an LC system using a gradient comprised of buffer A (H_2O + 0.1% Formic Acid (FA)) and buffer B (acetonitrile +

0.1% FA) with a 1% B (for 5 min) ramp from 1–95% B (10 min), a hold at 95% B for 1 min then back to 1% B for 12min. The samples were then analyzed using a nano-ESI MS (Velos Pro dual-pressure linear ion trap from Thermo) followed by collision induced dissociation (CID) of the 327 *m/z* peak corresponding to the calculated mass of the OPA/mercaptopropionic-acid/taurine derivative. CID data was collected at collision energy (CE) 20% and 25%.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

FINANCIAL DISCLOSURE

This work was supported by grants 1 R01 CA154754 from National Institutes of Health and by the Cancer Prevention and Research Initiative of Texas (CPRIT) grants RP100890 and RP120314. We would also like to thank B. Iverson for many helpful discussions and J. Aponte from the Brodbelt research group for assistance with the ESI-MS and CID analyses.

References

1. de la Rosa J, Stipanuk MH. Evidence for a rate-limiting role of cysteinesulfinate decarboxylase activity in taurine biosynthesis in vivo. Comp Biochem Physiol, Part B: Biochem Mol Biol. 1985; 81:565–571.
2. Lourenco R, Camilo M. Taurine: a conditionally essential amino acid in humans? An overview in health and disease, Nutr Hosp. 2002; 17:262–270.
3. Wu JY, Prentice H. Role of taurine in the central nervous system. J Biomed Sci. 2010; 17:S1. [PubMed: 20804583]
4. Timbrell JA, Seabra V, Waterfield CJ. The in vivo and in vitro protective properties of taurine. Gen Pharmacol. 1995; 26:453–462. [PubMed: 7789717]
5. Lambert IH, Hansen DB. Regulation of taurine transport systems by protein kinase CK2 in mammalian cells. Cell Physiol Biochem. 2011; 28:1099–1110. [PubMed: 22178999]
6. Tang XW, Hsu CC, Schloss JV, Faiman MD, Wu E, Yang CY, Wu JY. Protein phosphorylation and taurine biosynthesis in vivo and in vitro. J Neurosci. 1997; 17:6947–6951. [PubMed: 9278530]
7. Huxtable R. Physiological actions of taurine. Physiol Rev. 1992; 72:101–163. [PubMed: 1731369]
8. Wu JY. Purification and characterization of cysteic acid and cysteine sulfinic acid decarboxylase and L-glutamate decarboxylase from bovine brain. Proc Natl Acad Sci. 1982; 79:4270. [PubMed: 6956856]
9. Oertel W, Schmechel D, Weise V, Ransom D, Tappaz M, Krutzsch H, Kopin I. Comparison of cysteine sulphinic acid decarboxylase isoenzymes and glutamic acid decarboxylase in rat liver and brain. Neuroscience. 1981; 6:2701–2707. 2709–2714. [PubMed: 7322359]
10. Laidlaw S, Grosvenor M, Kopple J. The taurine content of common foodstuffs. J Parenter Enteral Nutr. 1990; 14:183–188.
11. Dawczynski C, Schubert R, Jahreis G. Amino acids, fatty acids, and dietary fibre in edible seaweed products. Food Chem. 2007; 103:891–899.
12. Tesoriere L, Fazzari M, Allegra M, Livrea M. Biothiols, taurine, and lipid-soluble antioxidants in the edible pulp of Sicilian cactus pear (Opuntia ficus-indica) fruits and changes of bioactive juice components upon industrial processing. J Agric Food Chem. 2005; 53:7851–7855. [PubMed: 16190641]
13. Hébert A, Forquin-Gomez M-P, Roux A, Aubert J, Junot C, Heilier J-F, Landaud S, Bonnarme P, Beckerich J-M. New Insights into Sulfur Metabolism in Yeasts as Revealed by Studies of Yarrowia lipolytica. Appl Environ Microbiol. 2013; 79:1200–1211. [PubMed: 23220962]
14. Tappaz M, Bitoun M, Reymond I, Sergeant A. Characterization of the cDNA Coding for Rat Brain Cysteine Sulfinate Decarboxylase. J Neurochem. 1999; 73:903–912. [PubMed: 10461879]

15. Fagiolini M, Fritschy JM, Low K, Mohler H, Rudolph U, Hensch TK. Specific GABAA circuits for visual cortical plasticity. Science's STKE. 2004; 303:1681.

16. Ge S, Goh ELK, Sailor KA, Kitabatake Y, Ming G, Song H. GABA regulates synaptic integration of newly generated neurons in the adult brain. Nature. 2005; 439:589–593. [PubMed: 16341203]

17. Nakatsu Y, Tyndale RF, DeLorey TM, Durham-Pierre D, Gardner JM, McDanel HJ, Nguyen Q, Wagstaff J, Lalando M, Sikela JM. A cluster of three GABAA receptor subunit genes is deleted in a neurological mutant of the mouse p locus. Nature. 1993; 364:448–450. [PubMed: 8392662]

18. Asada H, Kawamura Y, Maruyama K, Kume H, Ding RG, Kanbara N, Kuzume H, Sanbo M, Yagi T, Obata K. Cleft palate and decreased brain γ-aminobutyric acid in mice lacking the 67-kDa isoform of glutamic acid decarboxylase. Proc Natl Acad Sci. 1997; 94:6496. [PubMed: 9177246]

19. Capitani G, De Biase D, Aurizi C, Gut H, Bossa F, Grütter MG. Crystal structure and functional analysis of Escherichia coli glutamate decarboxylase. The EMBO journal. 2003; 22:4027–4037. [PubMed: 12912902]

20. Blankenhorn D, Phillips J, Slonczewski JL. Acid-and base-induced proteins during aerobic and anaerobic growth of Escherichia coli revealed by two-dimensional gel electrophoresis. J Bacteriol. 1999; 181:2209–2216. [PubMed: 10094700]

21. De Biase D, Tramonti A, Bossa F, Visca P. The response to stationary-phase stress conditions in Escherichia coli: role and regulation of the glutamic acid decarboxylase system. Mol Microbiol. 1999; 32:1198–1211. [PubMed: 10383761]

22. Cotter PD, Gahan CGM, Hill C. A glutamate decarboxylase system protects Listeria monocytogenes in gastric fluid. Mol Microbiol. 2001; 40:465–475. [PubMed: 11309128]

23. Cook AM, Denger K, Smits THM. Dissimilation of C 3-sulfonates. Arch Microbiol. 2006; 185:83–90. [PubMed: 16341843]

24. Fenalti G, Law RH, Buckle AM, Langendorf C, Tuck K, Rosado CJ, Faux NG, Mahmood K, Hampe CS, Banga JP. GABA production by glutamic acid decarboxylase is regulated by a dynamic catalytic loop. Nat Struct Mol Biol. 2007; 14:280–286. [PubMed: 17384644]

25. Dominy JE Jr, Simmons CR, Karplus PA, Gehring AM, Stipanuk MH. Identification and characterization of bacterial cysteine dioxygenases: a new route of cysteine degradation for eubacteria. J Bacteriol. 2006; 188:5561–5569. [PubMed: 16855246]

26. Buss K, Drewke C, Lohmann S, Piwonska A, Leistner E. Properties and Interaction of Heterologously Expressed Glutamate Decarboxylase Isoenzymes GAD65kDa and GAD67kDa from Human Brain with Ginkgotoxin and Its 5′-Phosphate. J Med Chem. 2001; 44:3166–3174. [PubMed: 11543686]

27. Bu DF, Erlander MG, Hitz BC, Tillakaratne N, Kaufman DL, Wagner-McPherson CB, Evans GA, Tobin AJ. Two human glutamate decarboxylases, 65-kDa GAD and 67-kDa GAD, are each encoded by a single gene. Proc Natl Acad Sci. 1992; 89:2115. [PubMed: 1549570]

28. Robichon C, Luo J, Causey TB, Benner JS, Samuelson JC. Engineering Escherichia coli BL21 (DE3) Derivative Strains To Minimize E. coli Protein Contamination after Purification by Immobilized Metal Affinity Chromatography. Appl Environ Microbiol. 2011; 77:4634–4646. [PubMed: 21602383]

29. Sha D, Wei J, Wu H, Jin Y, Wu JY. Molecular cloning, expression, purification, and characterization of shorter forms of human glutamic decarboxylase 67 in an *E. coli* expression system. Mol Brain Res. 2005; 136:255–261. [PubMed: 15893607]

30. Jacobsen JG, Thomas LL, Smith LH Jr. Properties and distribution of mammalian L-cysteine sulfinate carboxy-lyases. Biochim Biophys Acta, Spec Sect Enzymol Subj. 1964; 85:103–116.

31. Drummond, A.; Ashton, B.; Cheung, M.; Heled, J.; Kearse, M. Geneious v4. 8. 2009. Biomatters Ltd; Auckland, New Zealand: 2009.

32. Liu P, Ge X, Ding H, Jiang H, Christensen BM, Li J. Role of Glutamate Decarboxylase-like Protein 1 (GADL1) in Taurine Biosynthesis. J Biol Chem. 2012

33. Liu P, Ding H, Christensen BM, Li J. Cysteine sulfinic acid decarboxylase activity of *Aedes aegypti* aspartate 1-decarboxylase: The structural basis of its substrate selectivity. Insect Biochem Mol Biol. 2012

34. Pruski A, Fiala-Médioni A, Fisher C, Colomines J. Composition of free amino acids and related compounds in invertebrates with symbiotic bacteria at hydrocarbon seeps in the Gulf of Mexico. Mar Biol. 2000; 136:411–420.

35. Pruski A, Medioni AF, Prodon R, Colomines J. Thiotaurine is a biomarker of sulfide-based symbiosis in deep-sea bivalves. Limnol Oceanogr. 2000:1860–1867.

36. Yancey PH, Blake WR, Conley J. Unusual organic osmolytes in deep-sea animals: adaptations to hydrostatic pressure and other perturbants. Comp Biochem Physiol, Part A: Mol Integr Physiol. 2002; 133:667–676.

37. http://www.chem.agilent.com/Library/datasheets/Public/5980-3088.PDF.

38. DeLano, WL. The PyMOL molecular graphics system. 2002.

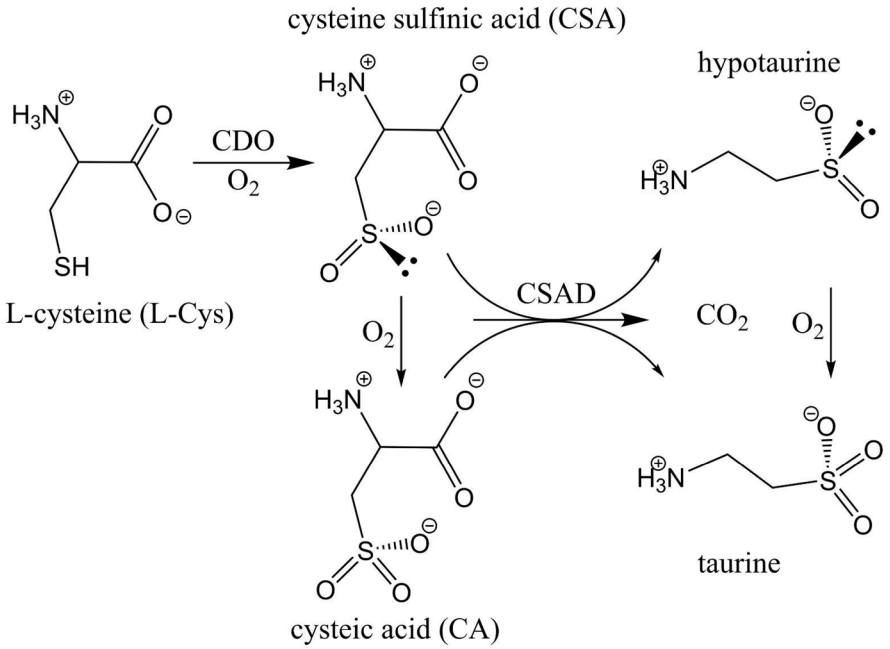

Figure 1.
Reactions catalyzed by cysteine dioxygenase (CDO) and cysteine sulfinic acid decarboxylase (CSAD).

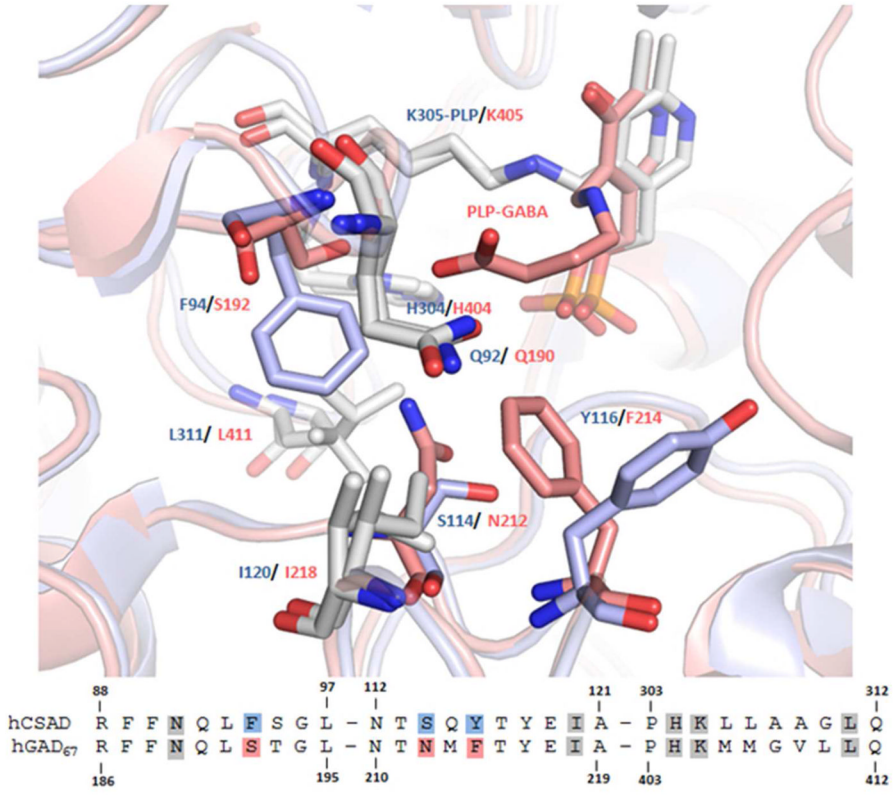

Figure 2.

Human CSAD (hCSAD, in blue, PDB entry 2JIS) active site overlaid against the human hGAD-67 kDa (hGAD-67 kDa, in red, PDB entry 2OKJ (24), with bound γ-aminobutyric acid (GABA)) active site, using PyMol (38). In sticks are amino acids within ~5Å from the PLP-GABA diamine in the hGAD-67 kDa structure. Residues colored in grey are conserved between the two structures, those colored in blue are specific to hCSAD and those colored in red are specific to hGAD-67 kDa. The sequence alignment depicts the linear substrate recognition motif ($X_1aa_{19}X_2aaX_3$).

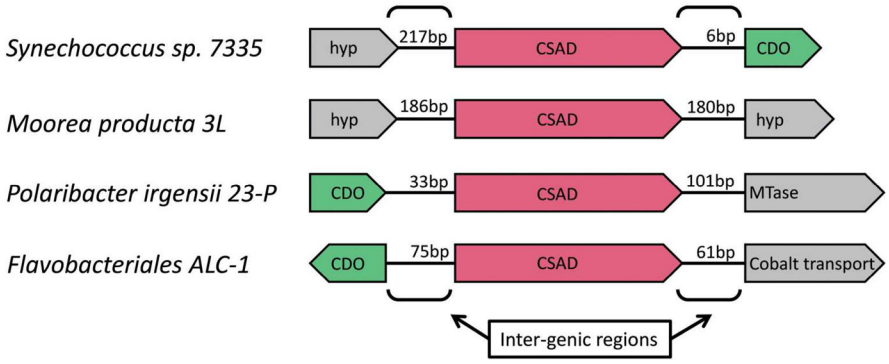

Figure 3.

Comparative schematic of putative bacterial operons with genes containing AAT_1 homologues with CSAD motif [$(F/W/Y)_1aa_{19}S_2aaY_3$]. CSAD = cysteine sulfinic acid decarboxylase, CDO = cysteine dioxygenase (predicted), hyp = hypothetical protein, MTase = methyltransferase (predicted), cobalt transport = gene predicted to code for a cobalt transporter. Brackets indicate number of base pairs (bp) between open reading frames.

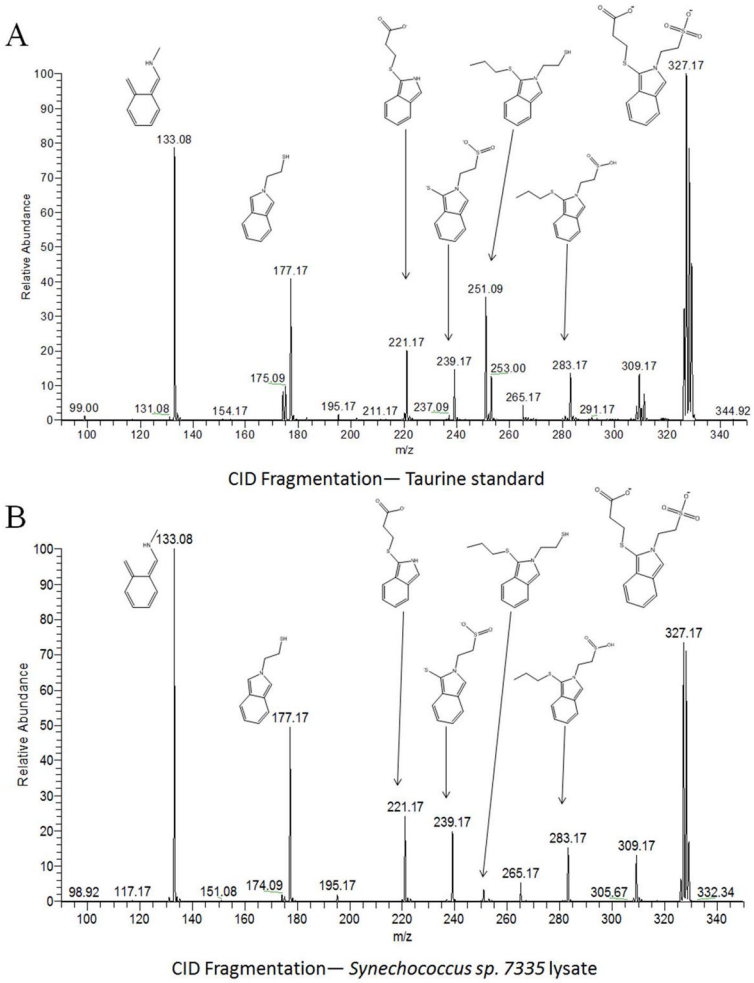

Figure 4.

Collision induced decay fragmentation of the 327 *m/z* peaks corresponding to the calculated mass of an OPA/mercaptopropionic-acid/taurine derivative in a taurine standard (A) and in a *Synechococcus sp. 7335* lysate (B). Structures represent predicted fragmentation species matching the observed *m/z* peaks.

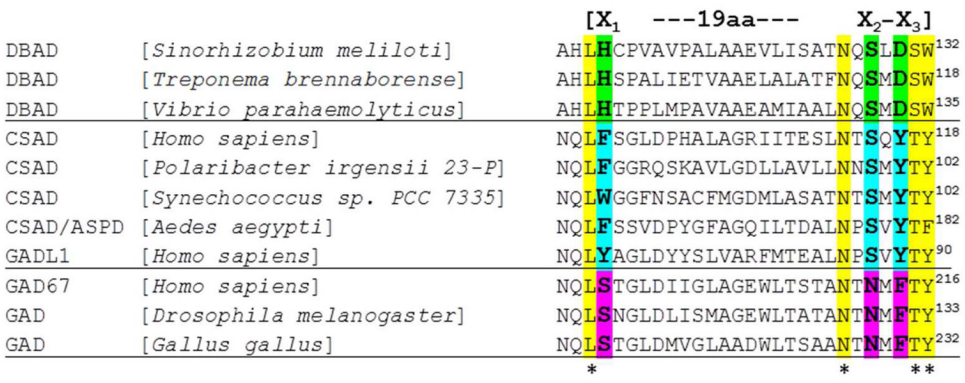

Figure 5.

Representative sequence alignment of AAT_I superfamily members with varying substrate recognition motifs ($X_1aa_{19}X_2aaX_3$). DBAD = diaminobutyric acid decarboxylase; CSAD = cysteine sulfinic acid decarboxylase; GAD = glutamic acid decarboxylase.

Table 1

2nd Order Rate Constants of hCSAD and hGAD-67 kDa variants

A. CSA Decarboxylation

Variant	k_{cat}/K_M $M^{-1}s^{-1}$	Variant	k_{cat}/K_M $M^{-1}s^{-1}$
CSAD-WT	3550 ± 880	GAD-WT	470 ± 160
CSAD-F94S	3.4 ± 2.1	GAD-S192F	7.3 ± 0.5 *
CSAD-S114N	1370 ± 520	GAD-N212S	52 ± 2 *
CSAD-Y116F	490 ± 210	GAD-F214Y	1240 ± 50 *
CSAD-S114N-Y116F	45 ± 3 *	GAD-N212S F214Y	36 ± 6 *
CSAD-F94S-S114N-Y116F	0.38 ± 0.01 *	GAD-S192F-N212S-F214Y	625 ± 60 *

B. **L-Glu Decarboxylation**

Variant	k_{cat}/K_M $M^{-1}s^{-1}$	Variant	k_{cat}/K_M $M^{-1}s^{-1}$
CSAD-WT	n.d.	GAD-WT	790 ± 180
CSAD-F94S	n.d.	GAD-S192F	2.2 ± 0.1 *
CSAD-S114N	n.d.	GAD-N212S	57 ± 2 *
CSAD-Y116F	2.2 ± 0.8	GAD-F214Y	107± 3 *
CSAD-S114N-Y116F	0.030 ± 0.002 *	GAD-N212S F214Y	0.30 ± 0.04 *
CSAD-F94S-S114N-Y116F	n.d.	GAD-S192F-N212S-F214Y	0.84 ± 0.02 *

C. CA Decarboxylation

Variant	k_{cat}/K_M $M^{-1}s^{-1}$	Variant	k_{cat}/K_M $M^{-1}s^{-1}$
CSAD-WT	122 ± 46	GAD-WT	34.0 ± 0.4 *
CSAD-F94S	0.020 ± 0.001 *	GAD-S192F	13 ± 1 *
CSAD-S114N	42 ± 2	GAD-N212S	2.4 ± 0.4 *
CSAD-Y116F	220 ± 65	GAD-F214Y	40 ± 3 *
CSAD-S114N-Y116F	77 ± 14 *	GAD-N212S F214Y	0.09 ± 0.01 *
CSAD-F94S-S114N-Y116F	0.08 ± 0.004 *	GAD-S192F-N212S-F214Y	93 ± 5 *

* Substrate saturation kinetics were not observed: Data were instead fit to $v_0/[E] = (k_{Cat}/K_M)*[S]$. n.d. = not detected