Structural and Functional Characterization of β-Cyanoalanine Synthase from *Tetranychus urticae*

Leily Daneshian[1], Isabella Renggli[1], Ryan Hanaway[1], Lesa R. Offermann[1], Caleb R. Schlachter[1], Ricardo A. Hernandez[1], Shannon Henry[1], Rahul Prakash[1], Nicky Wybouw[2], Wannes Dermauw[3,4], Linda S. Shimizu[1], Thomas Van Leeuwen[4], Thomas M. Makris[1,5], Vojislava Grbic[6], Miodrag Grbic[6,7], Maksymilian Chruszcz[1]*

[1]Department of Chemistry and Biochemistry, University of South Carolina, Columbia, SC 29208, USA
[2]Terrestrial Ecology Unit, Department of Biology, Ghent University, Ghent, B-9000, Belgium
[3]Flanders research institute for Agriculture, Fisheries and Food (ILVO), Plant Sciences Unit, Merelbeke, B-9820, Belgium
[4]Department of Plants and Crops, Ghent University, Ghent, B-9000, Belgium
[5]Department of Molecular and Structural Biochemistry, NC State University, Raleigh, NC 27607, USA
[6]Department of Biology, Western University, London, Ontario, N6A 5B7, Canada
[7]University of La Rioja, Logrono, Spain

*To whom correspondence should be addressed: Maksymilian Chruszcz, Department of Chemistry and Biochemistry, University of South Carolina, Columbia, SC 29208 USA; chruszcz@mailbox.sc.edu; Tel: (803) 777-7399, FAX: (803) 777-9521.

Keywords: β-cyanoalanine synthase, cyanide detoxification, cyanoalanine, herbivory, spider mite

Abstract

Tetranychus urticae is a polyphagous spider mite that can feed on more than 1100 plant species including cyanogenic plants. The herbivore genome contains a horizontally acquired gene *tetur10g01570* (*TuCAS)* that was previously shown to participate in cyanide detoxification. To understand the structure and determine the function of TuCAS in *T. urticae*, crystal structures of the protein with lysine conjugated pyridoxal phosphate (PLP) were determined. These structures reveal extensive TuCAS homology with the β-substituted alanine synthase family, and they show that this enzyme utilizes a similar chemical mechanism involving a stable α-aminoacrylate intermediate in β-cyanoalanine and cysteine synthesis. We demonstrate that TuCAS is more efficient in the synthesis of β-cyanoalanine, which is a product of the detoxification reaction between cysteine and cyanide, than in the biosynthesis of cysteine. Also, the enzyme carries additional enzymatic activities that were not previously described. We show that TuCAS can detoxify cyanide using O-acetyl-L-serine as a substrate, leading to the direct formation of β-cyanoalanine. Moreover, it catalyzes the reaction between the TuCAS-bound α-aminoacrylate intermediate and aromatic compounds with a thiol group. In addition, we have tested several compounds as TuCAS inhibitors. Overall, this study identifies additional functions for TuCAS and provides new molecular insight into the xenobiotic metabolism of *T. urticae*.

1. Introduction

Cyanide (CN^-) is a carbon-nitrogen anion that is found in a wide range of species such as bacteria, fungi, arthropods and plants (Yi et al., 2012). The CN^- is toxic due to its high reactivity. It reacts with keto compounds and Schiff bases leading to the production of cyanohydrins and nitrile derivatives, respectively. It also chelates di- and trivalent metal ions in the prosthetic groups of metalloproteins (Arenas-Alfonseca et al., 2018; Donato et al., 2007). In eukaryotic cells, it mainly targets mitochondria as it binds to the heme iron of cytochrome c oxidase and blocks the respiratory chain by inhibiting electron transport (Arenas-Alfonseca et al., 2018; Cooper and Brown, 2008; Isom and Way, 1984). In addition, in plants, CN^- binds to copper in plastocyanin, inhibiting plastocyanin-dependent electron transport to photosystem I (Berg and Krogmann, 1975) and dark CO_2 assimilation (Bishop et al., 1955; Trebst et al., 1960). CN^- toxicity in chloroplasts occurs in the dark and is partially reversible by illumination in the presence of an electron acceptor (Bishop and Spikes, 1955; Cohen and McCarty, 1976). As a result of this broad toxicity, the generation of CN^-/HCN necessitates keeping its concentration below toxic levels. Some organisms like plants and insects, use CN^- as a defense mechanism. In this case the cyanide is no accumulated, but is stored in a form of cyanogenic precursors that are activated, for example, after tissue damage (Zagrobelny et al., 2008). CN^- detoxification occurs through different metabolic pathways, including the degradation of CN^- and the formation of simple nitrogenous compounds such as formamide and ammonium (Watanabe et al., 1998), or incorporation of CN^- into primary metabolites. The incorporation of CN^- into primary metabolites is supported by either sulfur transferases like 3-mercaptopyruvate transferase and rhodanese, which are involved in the transfer of thiol (-SH) groups to cyanide to produce thiocyanate (Nagahara et al., 1999; Wing et al., 1992), or via a major pathway where cyanide is incorporated into nitrogen metabolism through the synthesis of aspartate and asparagine. The latter reaction is carried out by a β-cyanoalanine synthase (CAS) enzyme that belongs to the family of β-substituted alanine synthase (βSAS) enzymes (Ikegami et al., 1988; Maruyama et al., 2001; Piotrowski, 2008; Yi et al., 2012). CAS is

a pyridoxal phosphate-dependent enzyme (Arenas-Alfonseca et al., 2018) that catalyzes the substitution of L-cysteine sulfhydryl moiety with cyanide, forming the non-protein amino acid, β-cyanoalanine with the release of hydrogen sulfide (Figure 1A) (Okonji et al., 2018). Released H_2S can block the mitochondrial respiratory pathway and can be further detoxified by O-acetyl-L-serine sulfhydrylase (OASS) (Alvarez et al., 2012). OASS also belongs to the βSAS enzyme family, which incorporate H_2S into O-acetyl-L-serine (OAS) to produce L-cysteine (Figure 1B). Enzymes that catalyze the formation of cysteine (OASS) and β-cyanoalanine (CAS) are highly conserved (Burkhard et al. 1998; Bonner et al. 2005; Watanabe et al. 2008), so that individual enzymes typically carry both cysteine and β-cyanoalanine synthase activities, albeit with different efficiencies. For example, in *Arabidopsis thaliana* there are eight loci encoding enzymes with OASS and/or CAS activities (Li et al. 2021; Watanabe et al. 2008). Of these, four genes are highly expressed and encode three OASS that are localized in the cytosol, plastids, and mitochondria and one CAS enzyme that localizes in the mitochondria. Previous work showed that *CAS* genes have been acquired in animal lineages such as herbivorous arthropods by horizontal gene transfer (Wybouw et al 2014; Li et al. 2021). Similar to *Arabidopsis*, OASS/CAS genes in Lepidopterans also underwent duplications. Paralogous genes subsequently diversified expression patterns and enzymatic properties (Li et al. 2021). OASS genes are widespread in Lepidoptera, while enzymes with CAS activity are found in few lepidopteran insects, which enables them to feed on cyanogenic plants or produce their own cyanogenic glucosides as defense metabolites (Beran et al. 2019; Stauber et al. 2012; Van Ohlen et al. 2016; Zagrobelny et al., 2008).

The two-spotted spider mite (TSSM), *Tetranychus urticae* Koch (Arthropoda: Acari: Acariformes: Tetranychoidea) is one of the most destructive agricultural pests worldwide that feeds on more than 1,100 plant species belonging to more than 140 different plant families of vegetables, fruits, ornamentals and field crops (Jeppson et al., 1975; Rioja et al., 2017). Spider mites have stylets that protrude through the leaf epidermis, apparently without damaging these cells, to reach leaf mesophyll cellular layer where they feed from a single cell (Bensoussan et al., 2016). The host plant range of TSSM includes a number of plants that synthesize cyanogenic glucosides. These compounds are not toxic on their own and require hydrolysis to generate a reactive hydroxynitrile that releases toxic HCN (Okonji et al., 2018). The hydrolysis occurs upon tissue disruption that can occur during herbivore feeding. Despite the relatively minimal leaf damage caused by mite feeding, HCN is released (Balhorn et al., 2006). Comparative genomics showed that TSSM and other *Tetranychus* mites have a single OASS/CAS enzyme that they acquired through a horizontal gene transfer from free-living bacteria (Wybouw et al., 2014a). Here, we determine the spectroscopic, catalytic and structural features of TuCAS. We describe the crystal structure of TuCAS and further analyze its role in spider mites. Our report characterizes the enzymatic properties of TuCAS and confirms the role of the enzyme as a β-cyanoalanine synthase in the CN^- and H_2S detoxification system of plant-feeding spider mites. In addition, we identified additional TuCAS substrates, that widens possibilities for additional functions of TuCAS.

2. Materials and methods

2.1 Cloning of TuCAS into pMCSG53

Codon optimized *tetur10g01570* in the cloning pUC57 vector (Wybouw et al., 2014b) was subcloned into a pMCSG53 vector (Eschenfeldt et al., 2013) obtained from DNASU Plasmid Repository (Tempe, AZ). Polymerase chain reaction (PCR) was performed following the KOD polymerase (Millipore, Billerica, MA) protocol with the following additions: 1 M betaine

monohydrate, 2 mM magnesium sulfate, and 0.2 ng of pMCSG53 in a 50 µL reaction. Thermocycling was conducted as follows: initial polymerase activation at 95°C for 2 minutes followed by 35 cycles of 95°C for 20 seconds, 61°C for 10 seconds, 70°C for 2 minutes, and final elongation at 70°C for 10 minutes. TuCAS was amplified using Phusion polymerase (Thermo Fisher Scientific, Waltham, MA) following the standard protocol. Conditions for thermal cycling were as follows: initial denaturation 98°C for 30 seconds followed by 35 cycles of 98°C for 10 seconds, 68°C for 20 seconds, 72°C for 20 seconds, and final elongation at 72°C for 10 minutes. Each PCR reaction was run on a 1% agarose gel and the band corresponding to pMCSG53 or *tetur10g01570* was excised using a GeneJET Gel Extraction Kit (Thermo Fisher Scientific, Waltham, MA) and eluted with sterile water.

Sticky ends for ligation Independent cloning (LIC) were made by incubating gel excised PCR products with T4 DNA polymerase (NEB, Ipswich, MA). A 40 µL reaction containing 38 fmoles of pMCSG53, 5.0 mM DTT, 2.5 mM dCTP, 1X T4 polymerase buffer and 3 units of T4 DNA polymerase, and a 40 µL reaction containing 230 fmoles of *tetur10g01570*, 5 mM DTT, 2.5 mM dGTP, 1X T4 polymerase buffer and 3 units of T4 DNA polymerase were incubated at room temperature for 30 minutes. The T4 DNA polymerase was inactivated at 75°C for 20 minutes and the reactions were mixed 1:1, 1:2, and 1:4 at room temperature for 5 minutes. Subsequently, 1 µL of 25 mM EDTA was added and the reaction was incubated for an additional 5 minutes at room temperature. Following this final incubation, the LIC reaction was then transformed into DH5α cells following the heat shock protocol (Froger and Hall, 2007).

2.2 Site-directed mutagenesis of TuCAS K52A

For a single mutation of K52A in the active site of the protein or surface entropy reduction (SER) mutations for improving the crystallization (Goldschmidt et al., 2007) of TuCAS, primers were designed using NEBaseChanger (http://nebasechanger.neb.com/) (Table S1). To perform the PCR, a 50 µL reaction contacting 10 mM dNTPs, 10 µM forward and reverse primers, 1 ng of pCRS001 (pMCSG53 plasmid containing *tetur10g01570*), 10 µL 5X Q5 reaction buffer (Thermo Fisher Scientific), and 0.5 µL Q5 High-Fidelity DNA Polymerase (Thermo Fisher Scientific) were mixed in a PCR tube (Thermo Fisher Scientific). Thermocycling conditions were initial denaturation at 98°C for 30 seconds followed by 35 cycles of 10 seconds at 98°C, 30 seconds at 64°C, and 5 minutes at 72°C, and a final extension was for 2 minutes at 72°C. The PCR was performed using a PTC-100 PCR Programmable Thermal Controller in Hard Shell 96-Well PCR plate (Bio-Rad) sealed with Bio-Rad Microseal PCR Plate Sealing Film.

2.3 Expression and purification of recombinant proteins

Production of TuCAS and TuCAS mutants. The pCRS001 plasmid was transformed into BL-21(DE3) *E. coli* cells. Cells were grown to an OD of 0.8 at 37°C with shaking in the presence of 10 mg of pyridoxal phosphate (PLP) per 1 L culture. Next, cells were induced using 0.4 mM IPTG at 25°C and were grown for 12 h at 16°C with shaking. Cells were centrifuged and the pellet was frozen at -80°C until needed.

For purification, each gram of the pellet was resuspended in 5 mL of lysis buffer (50 mM HEPES, 500 mM NaCl, 0.05 mg PLP, pH 7.5) and lysed by sonication with a Qsonica Sonicator for 10 cycles (10 sec on, 40 sec off) at 60% power. Cell debris was separated by centrifugation and the crude extract was applied to Ni-NTA column (Thermo Fisher Scientific, Rockford, IL) which was equilibrated in 5 column volumes (CV) of binding buffer (50 mM HEPES, 500 mM NaCl, 40 mM imidazole, pH 7.5). The resin was washed with 5 CV binding buffer and the protein

was eluted using elution buffer (50 mM HEPES, 500 mM NaCl, 500 mM imidazole, pH 7.5). Next, the N-terminal hexahistidine-tag was cleaved using the tobacco etch virus (TEV). Protein and TEV were combined in a 1:25 (w/w) protease/protein ratio, added to SnakeSkin dialysis tubing with a 10 kDa molecular weight cutoff (MWCO) (Thermo Fisher Scientific), and allowed to equilibrate in dialysis buffer (50 mM HEPES, 50 mM NaCl, pH 7.5) for 12 h at 4°C. To separate the cleaved his-tag from protein, the mixture was applied to a pre-equilibrated Ni-NTA column and washed with the binding buffer. The flow-through and wash were collected and concentrated with Amicon Ultra concentrators (EMD Millipore, Billerica, MA, USA) with a 3 kDa MWCO. To further purify the sample, the concentrated protein was run on a Superdex 200 column attached to an ÄKTA Pure FPLC system (Cytiva, Marlborough, MA) previously equilibrated with 25 mM HEPES, 50 mm NaCl, pH 7.5. Protein concentration was determined using the molecular weight and extinction coefficient estimated by ProtParam (Gasteiger et al., 2003) with NanoDrop 2000c Spectrophotometer (Thermo Fisher Scientific). K52A and SERp mutants were expressed and purified using the same procedure as the wild type of protein.

Production of Tetur36g00900. Tetur36g00900 was identified as a putative sulfur carrier potein based on a text search of Uniprot (Consortium, 2019) with query: ubiquitin *Tetranychus urticae*, and later identification of a conserved C-terminal GG motif. Only Tetur36g00900 (Uniprot code: T1L3S0) satisfied the search conditions. BLAST searches (Altschul et al. 1990) using the *Mycobacterium tuberculosis* CysO sequence (Jurgenson et al., 2008) as query returned Tetur36g00900 as the only hit when E-value of 1000 was used. To test whether TuCAS interacts with Tetur36g00900*,* a putative sulfur carrier protein, a construct was ordered from Bio Basic (Amherst, NY) in a pET21b (+) vector. The vector containing tDNA coding for Tetur36g00900 gene was transformed into BL-21(DE3) *E. coli* strain. BL-21(DE3) cells were grown to an OD of 0.8 at 37°C while shaking. Expression and purification of the protein were performed similarly to TuCAS only that instead of 25 mM HEPES buffer 20 mM Na_2HPO_4 buffer was used.

2.4 Differential scanning fluorimetry

Differential Scanning Fluorimetry was performed as described previously (Schlachter et al., 2019; Daneshian et al., 2021). Briefly, 1 µL of SYPRO® Orange Dye (Thermo Fisher, Waltham, MA) was added to a 1.0 mL solution of 1 mg/mL CAS protein for dilution of 1:1000. Next, 10 µL of the mixture of dye and protein were incubated with 10 µL of each pH and salt screen condition for a 20 µL total reaction volume in a Bio-Rad Hardshell 96-well RT-PCR plate. The plate was sealed with Bio-Rad Microseal PCR Plate Sealing Film (Hercules, CA). Fluorescence data was collected by a Bio-Rad CFX96 RT-PCR instrument (Hercules, CA) as the temperature was increased in the range of 30-90°C with a 1°C/min in 2°C increments. Emission was measured at 590 nm (excitation at 488 nm). Screen conditions had working concentrations of 50 mM buffer with a pH range of 4.0-9.5 in 0.5 pH unit increments and a sodium chloride range from 0-1.0 M (No salt, 0.05 M, 0.10 M, 0.15 M, 0.20 M, 0.25 M, 0.50 M and 1.0 M). Buffers used to maintain the pH ranges were acetate (pH 4.0-5.0), Bis-Tris (pH 5.5-6.5), Tris (pH 7.0-8.0) and CHES (pH 8.5-9.5).

2.5 Enzymatic activity assays

β-Cyanoalanine synthesis with L-cysteine and KCN as substrates. Kinetic assays for TuCAS activity were performed under pseudo-first-order conditions by measuring the product of the reaction between L-cysteine (Sigma Aldrich) and KCN (Alfa Aesar) (Yi et al., 2012). The reaction was performed in 500 µL reaction mixture containing buffer (100 mM Tris, 50 mM MES,

and 50 mM acetic acid, pH 6.5-9.5 (Ellis and Morrison, 1982)), L-cysteine (0 to 4 mM), KCN (0 to 5 mM), and 1.2 µg protein. Reactions were terminated after 15 min of incubation at 25°C by the addition of 50 µL of 30 mM $FeCl_3$ (ACROS Organics) in 1.2 M HCl and 50 µL of 20 mM N,N-dimethyl-*p*-phenylenediamine dihydrochloride (ACROS Organics) in 7.2 M HCl (Yi et al., 2012). Next, the mixture was incubated in the dark for 10 min. The reaction between L-cysteine and KCN releases hydrogen sulfide. One molecule of sulfide and two molecules of N, N-dimethyl-*p*-phenylenediamine dihydrochloride produces methylene blue. The amount of methylene blue produced was determined at 670 nm using a Beckman DU800 UV/visible spectrophotometer (Yi et al., 2012). Data was processed using OriginPro software (Northampton, MA) and steady-state kinetic parameters were determined by the Michaelis–Menten or Hill equations.

Cysteine synthesis with O-acetyl-L-serine and Na_2S as substrates. To investigate if TuCAS also acts as an OASS to detoxify H_2S, O-acetyl-L-serine (Sigma Aldrich) and Na_2S (Sigma Aldrich) were used in (0 to 9 mM) and (0 to 15 mM) concentrations, respectively. The reaction was performed in 1000 µL reaction mix containing buffer (100 mM Tris, 50 mM MES, and 50 mM acetic acid, pH 6.5-9.5 (Ellis and Morrison, 1982)), and 12.7 µg protein and was terminated after 10 mins of incubation at 25°C in a closed 1000 µL tube by addition of 50 µL of 20% trichloroacetic acid. Next 250 µL of the reaction solution was mixed with 250 µL ninhydrin reagent. Ninhydrin reagent was prepared by dissolving 1.25 g of ninhydrin in a mixture of 30 mL of acetic acid and 20 mL of 6 M phosphoric acid. The mixture of the reaction and the ninhydrin reagent was heated for 5 mins at 98°C on a heat block (Thermolyne, MD). Then, 50 µL cold 100% ethanol was added to the mixture and it was incubated on ice for 15 mins to measure the absorbance at 560 nm using a Beckman DU800 UV/visible spectrophotometer (Bonner et al., 2005).

Compound 1 and 2 synthesis. 3-(3-(3,4-dichlorophenyl)ureido)benzoic acid (Compound 1) and 3-(3-(4-chlorophenyl)ureido)benzoic acid (Compound 2) were synthesized from commercial isocyanates and aryl amines, similar to prior reports (Brunner et al., 2016a). Compound 1: ^{1}H NMR (300 MHz, DMSO) δ 12.95 (s, 1H), 9.04 (s, 2H), 8.13 (s, 1H), 7.89 (d, J = 2.4 Hz, 1H), 7.63 (d, J = 8.6 Hz, 1H), 7.57 (d, J = 7.8 Hz, 1H), 7.53 (d, J = 8.8 Hz, 1H), 7.41 (t, J = 7.8 Hz, 1H), 7.34 (dd, J = 8.8, 2.4 Hz, 1H). ESI-MS *m/z* expected 325.01, found 325.0142 $(M+H)^+$ Compound 2: ^{1}H NMR (300 MHz, DMSO) δ 12.93 (s, 1H), 8.94 (s, 1H), 8.85 (s, 1H), 8.12 (s, 1H), 7.63 (d, J = 8.0 Hz, 1H), 7.55 (d, J = 7.8 Hz, 1H), 7.50 (d, J = 8.9 Hz, 2H), 7.40 (t, J = 7.9 Hz, 1H), 7.33 (d, J = 8.9 Hz, 2H). ESI-MS expected *m/z* 291.05, found *m/z* 291.0529 $(M+H)^+$. See supplementary data for further information (Figures S1-S4).

Inhibition of TuCAS activity. O-phospho-L-serine (Sigma Aldrich), NH_4SCN (ThermoFisher Scientific), D-serine, L-serine, D-cycloserine, L-cycloserine, (Sigma Aldrich), S-methyl-L-cysteine (Sigma Aldrich), cyanate (Sigma Aldrich), reduced glutathione (Sigma Aldrich), triclocarban (Sigma Aldrich), 3-(3-(3,4-dichlorophenyl)ureido)benzoic acid (Compound 1) (Figure S5B) and 3-(3-(4-chlorophenyl)ureido)benzoic acid (Compound 2) (Brunner et al., 2016a) (Figure S5C) were used in different concentrations to check their inhibition effect on the β-cyanoalanine synthase and cysteine synthase reactions. Lineweaver Burk equation (Lineweaver and Burk, 1934) was used to calculate inhibition constants (K_i).

Insights into the role of Tetur36g00900 in the biosynthesis of cysteine and its interactions with TuCAS. To investigate whether TuCAS uses small sulfur carrier proteins as a sulfur donor (Jurgenson et al., 2008), the binding interaction of TuCAS and Tetur36g00900 was studied using native gel electrophoresis. Tetur36g00900 is a ubiquitin-like protein with a predicted thiol carboxylate group at the C-terminus. The two proteins were incubated together at 4°C overnight in a mixture supplemented with 500 µM PLP, and 5 mM L-cysteine/10 mM O-acetyl-L-serine/10

mM O-phospho-L-serine. In addition, we have used thioacetic acid and thiobenzoic acid (Sigma Aldrich) as potential sulfur donors and mimickers of the C-terminal thiocarboxylate group of the sulfur carrier protein. 10 mM O-acetyl-L-serine with 10 mM of each mentioned compound were incubated in a 1000 µL reaction mixture containing 20 mM Tris buffer pH 9.0 and 20 µg TuCAS protein for 16 h at 25°C. Later the formation of the cysteine was investigated using the ninhydrin assay as described previously. 250 µL of the reaction solution was mixed with 250 µL ninhydrin reagent. The mixture of the reaction and the ninhydrin reagent was heated for 5 mins at 98°C on a heat block. Then, 50 µL cold 100% ethanol was added to the mixture and it was incubated on ice for 15 mins to measure the absorbance at 560 nm using a Beckman DU800 UV/visible spectrophotometer.

Formation of cysteine derivatives. The activity of TuCAS was also investigated toward thiol containing compounds such as thiosalicylic acid, 3-mercaptobenzoic acid, and 4-mercaptobenzoic acid (all from Sigma Aldrich). Similarly, to the experiments involving thioacetic acid and thiobenzoic acid, 10 mM O-acetyl-L-serine was incubated with 10 mM of each mentioned compound in a 1000 µL reaction mix containing 20 mM Tris buffer pH 9.0 and 20 µg protein for 16 hours at 25°C. Later, the protein was removed from the sample using Amicon Ultrafilters (3 kDa) and the formation of cysteine derivatives was confirmed with mass spectroscopy (Sigure S6).

β-*Cyanoalanine synthesis with O-acetyl-L-serine and KCN as substrates.* Here we investigated if TuCAS also detoxifies KCN using O-acetyl-L-serine as a substrate. The reaction was performed in a 1000 µL reaction mix containing 20 mM Tris buffer pH 9.0 and 20 µg protein and was incubated for 16 h at 25°C in the dark. Later, the protein was removed from the sample using Amicon Ultrafilters (3 kDa MWCO) and the formation of β-cyanoalanine was confirmed with mass spectroscopy. The mass spectrometry was completed on two samples in three replicates: 1) A control sample which is the mixture of 10 mM O-acetyl-L-serine and 5 mM KCN in 20 mM Tris buffer pH 9.2) The reaction sample which is the same as the control sample with the addition of 20 µg TuCAS. The peak at 113 Da corresponding to β-cyanoalanine was observed only in the replicated reaction samples and not in the control samples (Figure S6C).

2.6 UV/visible spectroscopy

Absorption spectra (350 to 600 nm) of TuCAS (1.3 mg/mL) in the absence or presence of L-cysteine, KCN, NH_4SCN, O-acetyl-L-serine (Thermo Fisher Scientific), S-methyl-L-cysteine (Thermo Fisher Scientific), and reduced glutathione (Thermo Fisher Scientific) were obtained in 25 mM HEPES, 50 mM NaCl, pH 7.5 by binding the substrates to the protein and was observed by monitoring the PLP spectrum using a Beckman DU800 UV/visible spectrophotometer. Spectroscopic analysis indicates that the wild-type protein reported here has 50% PLP content and no PLP is bound by the mutant. First, the concentration of TuCAS was calculated with the estimated extinction coefficient of 29910 M^{-1} cm^{-1} at 280 nm. Next, the concentration of TuCAS bound to PLP was measured at 412 nm using the estimated extinction coefficient 5900 M^{-1} cm^{-1} (Sang et al., 2007). The ratio of the two measurements showed 50% PLP content.

2.7 Crystallization of TuCAS

The recombinant TuCAS produced by us with an N-terminal cleavable his-tag was used for crystallization. In addition, surface entropy reduction (SER) mutations (E151A/E152A, and K207A/E208A/K209A, K252A/K253A, and K276A/E277A) for crystallization were designed using the SERp Server (Goldschmidt et al., 2007). After trying various conditions with the different designed proteins, we were able to crystallize native TuCAS (pMCSG53) using a sample

containing protein at a high concentration (~63 mg/mL) with the purification tag cleaved. Crystallization experiments were performed by the sitting-drop vapor diffusion method at 4°C using MRC 2-drop 96-well crystallization plates. The crystallization plates and the PEG-ION crystal screen were purchased from Hampton Research (Aliso Viejo, CA). A drop of protein (63 mg/mL) mixed with mother liquor 1:1 was set on the plate. Yellow crystals were grown in 0.1 M sodium malonate pH 6.0, 12% w/v PEG 3350 (Protein Data Bank (PDB) code: 6PMU) and 0.2 M ammonium chloride, 12% w/v PEG 3350, pH 7.0 (PDB code: 6XO2, and 7MFJ).

2.8 Data collection, structure determination, and refinement

Crystals were taken directly from the crystallization solution and cooled in liquid nitrogen. Diffraction data were collected using Southeast Regional Collaborative Access Team (SER-CAT) 22ID beamline at the Advanced Photon Source (APS), Argonne National Lab (Argonne, IL). Data were processed with the HKL-2000 software package (Otwinowski and Minor, 1997). The structure was determined by molecular replacement using HKL-3000 (Minor et al., 2006) and MOLREP (Vagin and Teplyakov, 1997). Soybean β-cyanoalanine synthase (PDB code: 3VBE) was used as a search model for the determination of the 6PMU structure, while the 6PMU model and molecular replacement were used for the determination of 6XO2 and 7MFJ (Table S2). Refinement for all structures was performed using REFMAC (Murshudov et al., 2011) and HKL-3000. In the case of the 6PMU and 7MFJ structures, non-crystallographic symmetry was used during the whole process of refinement. TLS refinement and TLS Motion Determination server were used during the last stages of refinement (Painter and Merritt, 2006). The models were validated with COOT (Emsley and Cowtan, 2004) and MOLPROBITY (Davis et al., 2007). Various programs from the CCP4 package (Winn et al., 2011) were used for handling diffraction data and coordinates. Table S2 shows the data collection, refinement, and validation statistics for the structures. All crystal structures for TuCAS were deposited in the Protein Data Bank (PDB) with the accession codes 6PMU, 6XO2, and 7MFJ.

2.9 Computational analysis of the TuCAS crystal structures

The programs ProFunc and PDBePISA were used to analyze the crystal structures (Krissinel and Henrick, 2007; Laskowski, 2017). COOT, PyMOL (DeLano W. 2002), and UCSF-Chimera (Pettersen et al., 2004) were used to visualize and analyze the structures, as well as to generate figures. DALI (Holm and Rosenstrom, 2010) and PDBeFold (Krissinel and Henrick, 2004) were used to search for similar structures. Program PDBePISA was used to analyze the oligomeric structure of TuCAS (Berman et al., 2012). The pI values were calculated using the ExPASy ProtParam tool.

3. Results

3.1 Protein production

The TuCAS coding sequence (codon optimized for *E. coli*) was cloned into pMCSG53 vector with a cleavable histidine (His)-tag, which resulted in the production of protein containing an N-terminal His-tag, TEV cleavage site ((SSGVDLGTENLYFQ/S), and the amino acid sequence (residues 1-320) from *T. urticae*. The histidine tag was cleaved for all further studies. Overexpression of both wild type and the TuCAS K52A (inactive) mutant yielded approximately 20 mg per 1.0 L of culture and 5-10 mg per 1.0 L culture, respectively. Wild-type TuCAS and the PLP binding site mutant K52A (Yi et al., 2012) were purified and eluted as monomers after SEC (Figure S7). However, native gel electrophoresis, which was at pH 8.3, indicates the formation of

dimers for both the wild type TuCAS and the TuCAS K52A mutant (Figure S7A). The dimeric form of the enzyme is also observed in the crystal form.

Tetur36g00900 production yielded approximately 15 mg per 1.0 L culture. The protein was soluble and was present in solution in a monomeric form. We were able to cleave the N-terminal purification tag, and protein with the His-tag removed was used for all experiments involving Tetur36g00900.

3.2 UV spectroscopy

Optical spectroscopy reveals that ~ 50 % of wild-type TuCAS has the PLP cofactor bound. The addition of L-cysteine or O-acetyl-L-serine to TuCAS lower the absorbance (A_{max}) of PLP at 412 nm and increase the absorbance at 470 nm (Figure 2), which is consistent with the formation of an α-aminoacrylate intermediate from either molecule (Bonner et al., 2005). K52A spectra shows a flat line which indicated the absence of the PLP cofactor and the α-aminoacrylate intermediate. It is reported that *Arabidopsis* OASS (Bonner et al., 2005), as well as soybean OASS and CAS (Yi et al., 2012), show the signal for the α-aminoacrylate intermediate (A_{max} = 470 nm) after addition of O-acetyl-L-serine or L-cysteine. These experiments demonstrate that TuCAS forms a common α-aminoacrylate intermediate after addition of L-cysteine or O-acetyl-L-serine (Figure 1D).

3.3 Stability of TuCAS

Differential Scanning Fluorimetry (DSF) was performed to compare the thermal stability of wild-type enzyme with the stability of the active-site mutant. TuCAS is relatively stable in a broad pH range from 5 to 9 with melting temperatures of 52°C and 54°C in the presence of high salt (Figure 3A). The addition of 500 μM PLP appears to increase the thermal stability of the enzyme as much as 8°C (Figure 3B), which is consistent with the observation that the recombinant wild-type TuCAS we produced is not fully loaded with the cofactor. DSF results indicate that the active site mutation K52A reduces the thermal stability of TuCAS as much as 8°C and the addition of PLP does not change this stability (Figure 3C, D), which is expected as the mutant enzyme is not able to bind the cofactor.

3.4 TuCAS enzymatic activity and inhibition

3.4.1 Cyanide detoxification and cysteine synthesis

The optimum pH for TuCAS activity for both cyanide detoxification and cysteine biosynthesis is 8.5 (Figure 4A, D), which is consistent with an increased enzymatic activity at basic pH that has been previously observed for TuCAS (Wybouw et al., 2014a) and other CAS enzymes (van Ohlen et al., 2016; Yi and Jez, 2012; Yi et al., 2012). The kinetic data were fitted using the Hill equation with Hill coefficients of 2.6 and 2.7 for KCN and Na_2S respectively (Table 1), indicating a positive cooperativity between enzyme subunits was occurring when KCN, Na_2S, L-cysteine, or O-acetyl-L-serine were used as substrates (Figure 4). Values of k_{cat} or catalytic efficiencies for substrates in the β-cyanoalanine synthesis reaction are 9 and 6 times higher than the equivalent values for substrate in the cysteine synthesis reaction. The specificity of the enzyme for substrates involved in β-cyanoalanine synthesis is approximately 25 and 12 times higher than for substrate in the cysteine synthesis reaction (Table 1). The K_M for L-cysteine is 1.6 ± 0.2 mM and for O-acetyl-L-serine is 4.3 ± 0.3 mM. Comparison of TuCAS activity and some of the previously characterized βSAS proteins are summarized in Table S3.

Interestingly, we have also shown that TuCAS can detoxify cyanide without producing potentially toxic sulfide. In this case, the enzyme is using O-acetyl-L-serine and cyanide as substrates, which leads to production of β-cyanoalanine and acetate (Figures 1C and S6C). Such a reaction was not previously described for any CAS or OASS.

3.4.2 Rection between O-acetyl-L-serine and compounds containing thiol groups

It has been shown that OASS can conjugate aromatic thiol compounds with O-acetyl-L-serine and produce cysteine derivatives (Burow et al., 2007). Therefore, we have tested whether TuCAS can perform a similar reaction, which is does not involve cyanide as a substrate. To test this hypothesis, we utilized thiosalicylic acid, 3-mercaptobenzoic acid, or 4-mercaptobenzoic acid as TuCAS substrates (Figure S8). The reaction products were observed only when thiosalicylic acid or 4-mercaptobenzoic acid were used (Figure S9). The formation of the cysteine derivatives was confirmed using mass spectrometry (Figure S6A, B).

3.4.3 Search for an alternative sulfur donor in reaction of cysteine synthesis

M. tuberculosis uses an enzyme (CysM) that is homologous to TuCAS during cysteine biosynthesis (Jurgenson et al., 2008; Schnell et al., 2007). *In vivo,* CysM works together with CysO, which is a sulfur carrier protein. CysO is a small, ubiquitin-like protein that has a characteristic C-terminal sequence motif with two glycine residues. The carboxyl group of the CysO C-terminal glycine is modified to thiocarboxyl group that serves as the sulfur donor during the cysteine synthesis reaction catalyzed by CysM. We have identified Tetur36g00900 as a putative orthologue of CysO, hypothesized to work with TuCAS during cysteine synthesis in TSSM. Tetur36g00900 shares a 25% amino acid sequence identity (37% similarity) with *M. tuberculosis* CysO, it has a very similar length and also has the conserved C-terminal sequence motif including two glycine residues. However, we were not able to demonstrate a direct interaction between Tetur36g00900 and TuCAS using native electrophoresis (Figure S10). In addition, we have not seen any influence of Tetur36g00900 on TuCAS enzymatic activity. Therefore, we have used thioacetic acid and thiobenzoic acid (Figure S8) as potential mimics of the C-terminal thiocarboxyl group of the sulfur carrier protein and potential sulfur donors. However, again in this case we did not observe evidence for cysteine formation.

3.4.4. TuCAS inhibitors

In order to identify potential TuCAS inhibitors, the kinetic parameters associated with the modulation of the β-cyanoalanine synthase activity were determined (Table 2, Figures S11 and S12). In the reaction with L-cysteine and KCN as the substrates, S-methyl-L-cysteine (K_i = 3.7 ± 0.2 mM) and O-phospho-L-serine (K_i = 25 ± 3 mM) were competitive inhibitors of L-cysteine. L-cycloserine (K_i = 130 ± 8 μM) was mixed inhibitor toward L-cysteine. L-serine showed mixed noncompetition toward L-cysteine (K_i = 4.3 ± 0.3 mM). Thiocyanate (K_i = 3.5 ± 0.3 mM) and cyanate (K_i = 9.0 ± 1.0 mM) were competitive inhibitors of CN^-. In addition, D-serine, D-cycloserine, and reduced L-glutathione were tested as inhibitors. However, they did not have inhibitory effects on the formation of β-cyanoalanine.

The kinetic parameters associated with the modulation of L-cysteine synthesis reaction by the inhibitors were also determined. In this reaction, where O-acetyl-L-serine and Na_2S served as substrates, L-serine (K_i = 4.3 ± 0.1 mM), L-cycloserine (K_i = 74.0 ± 8.5 μM), and O-phospho-L-serine (K_i = 7.0 ± 0.8 mM) showed competitive inhibition against O-acetyl-L-serine. Thiocyanate (K_i = 6.4 ± 1.9 mM) and cyanate (K_i = 19.4 ± 2.4 mM) showed noncompetitive inhibition against

Na_2S. D-serine, D-cycloserine, and reduced L-glutathione were tried as inhibitors, but they did not show any effects on the formation of L-cysteine.

Compound 1, Compound 2, and triclocarban, shown in Figure S5, were chosen as potential TuCAS inhibitor based on studies by Brunner *et al*. (2016). These studies showed that *M. tuberculosis* OASS was inhibited by molecules similar to Compounds 1 and 2. However, only compound 1 was found to inhibit TuCAS. Compound 1 had a mixed inhibition effect with K_i = 150 ± 10 μM for β-cyanoalanine synthesis and K_i = 140 ± 20 μM for cysteine synthesis.

3.5 Crystal structures of TuCAS

3.5.1 Overall structure

The first crystal form of TuCAS ($P4_32_12$ space group; PDB code: 6PMU) contains a protein dimer in the asymmetric unit (Figure 5). The crystal structure obtained from this form was determined at 2.1 Å resolution. Chain A of the dimer includes residues 5-221 and 257-318, while chain B includes residues 0-220 and 257-319, where the very first residue of chain B comes from the purification tag. In the case of this structure, residues Glu221 to Asp257 could not be modeled (Figure 5A) due to the absence of interpretable electron density which indicates a significant flexibility of this TuCAS fragment.

The second crystal structure of TuCAS ($P4_12_12$ space group; PDB code: 6XO2) contains only one protein chain in the asymmetric unit, and this structure was determined at 1.6 Å resolution. While there is one protein chain in the asymmetric unit, the analysis of the crystal structure reveals that TuCAS is present in the form of a dimer, and the dimer symmetry coincides with the crystal symmetry. A single chain of the dimer is composed of residues (-2)-319, where residues -2, -1, and 0 originate from the purification tag. In both crystal forms, the C-terminal Glu320 could not be modeled. Similar to the structure with the dimer in the asymmetric unit, a significant fraction of the fragment formed by residues 221-257 is not visible in the electron density and was not modeled. However, in this case, residues 225-230 were visible in the electron density. Generally, the protein chains in both crystal forms have a similar conformation and superpose with rmsd (root mean square deviation) values of approximately 0.6Å.

The third crystal form of TuCAS ($P4_32_12$ space group; PDB code: 7MFJ) contains a protein dimer in the asymmetric unit. This crystal was grown using the same condition as crystals with PDB code 6PMU, except the protein used for crystallization was incubated with 10 mM O-phospho-L-serine. The crystal structure obtained from this form was determined at 2.35 Å resolution. Chain A of the dimer includes residues 0-221 and 255-320, while chain B includes all residues 0-318, where the very first residue of chains comes from the purification tag.

The single TuCAS chain can be divided into two domains (Figure 5A). The smaller domain (residues 51–163) consists of four-stranded parallel β sheets surrounded by four α helices. The larger domain (residues 1-50 and 164–319) consists of six-stranded β sheets (five parallel β-strands and one antiparallel strand) surrounded by four α helices. PLP is present in the active site of the protein in a cleft between the two protein domains. In each crystal form, there is clear electron density corresponding to PLP that is covalently linked to Lys52, and PLP was modeled in all protein chains with full occupancy. Additionally, in the higher resolution structure (PDB code: 6XO2) and the structure that contains the missing fragment (PDB code: 7MFJ), we have identified an acetate ion that forms hydrogen bonds with the side chains of Ser80, Asn106, as well as the main chain of Ser103.

The TuCAS dimeric assembly is very similar to one observed in homologous proteins that had their structures determined (Figure 5C) (Francois et al., 2006; Jurgenson et al., 2008; Yi et al.,

2012). The interface area between TuCAS chains forming dimers is 1984Å^2 for 6PMU, 1965Å^2 for 6XO2, and 1999Å^2 for 7MFJ which is significantly over the cutoff value (856 Å^2) proposed for discrimination between homodimeric and monomeric proteins (Ponstingl et al., 2000). The analysis of the dimer interface shows that the protein chains forming the dimer interact through hydrophobic interactions, as well as a network of hydrogen bonds. The most important residues for the hydrophobic interactions are the following: Ile12, Leu19, Pro24, Ile26, Leu45, Pro47 Ala92, Val93, Met113, Leu117, Leu279, and Leu316. Based on PDBePISA analysis of the structure, the shortest hydrogen bonds corresponding to the strongest hydrogen bonds that stabilize the dimer are formed by the main chain and/or side-chain atoms of Ile9, Thr13, Pro14, Leu25, Ala27, Arg30, and Gln181. Most of the residues stabilizing the dimer originate from the larger domain and they cluster mainly in the N-terminal portion of the protein. Some of these residues, including Pro24, Leu25, Arg30, Pro47, Gln181, and Leu279, are conserved in almost all the enzymes in this family (Figure 6). In addition, several water molecules are trapped between protein chains that form the dimer, and these water molecules provide additional hydrogen bonds that bridge the interface. Interestingly, even though crystal structures, as well as native electrophoresis, clearly indicate that TuCAS forms dimers, the results of the size exclusion chromatography indicate that the protein elutes at the volume corresponding to the monomeric form (Figure S7). To test the effect of pH and salt concentration on the quaternary TuCAS structure several conditions were tested. Buffers with pH 6.0, 7.0, and 8.0, as well as NaCl concentrations in the 0-500 mM range were tested. However, in all of the tested conditions, SEC results show that TuCAS runs as a monomer (data is not shown), which suggests that the interactions between dimer forming chains are relatively weak.

3.5.2 Comparison of TuCAS with homologous proteins

The overall conformations of the TuCAS chains in all obtained crystal forms are very similar (Figure 5). Comparison of the sequence and structure of TuCAS against other entries deposited to the PDB reveals a relatively large group of proteins with high structural similarity. However, none of the homologous proteins that have their structures determined have a sequence similarity higher than 40%. Despite the relatively low sequence identity, there are numerous examples of structures that superpose onto TuCAS models with rmsd values lower than 1.7 Å. For example, PDBeFOLD (Krissinel and Henrick, 2004) showed that TuCAS shares a very similar overall fold with bacterial enzymes of the cysteine synthase family originating from *Leishmania major* (PDB code: 4AIR), *Fusobacterium nucleatum* (PDBe code: 5XEO), *Mycobacterium tuberculosis* (PDBe code: 5IW8) or *Salmonella typhimurium* (PDB code: 2JC3), as well as plant proteins from soybean (PDB code: 3VBE) or *Arabidopsis thaliana* (PDB code: 1Z7W). In addition, both structure and/or sequence-based searches indicate a significant similarity between TuCAS and proteins that have their function annotated as cystathionine β-synthase (CBS), threonine synthase (e.g. *Bacillus subtilis* protein; PDB code: 6CGQ), or proteins involved in the production of siderophores, like *Staphylococcus aureus* SbnA (PDB code: 5D86), which participates in the production of staphyloferrin B (Petronikolou et al., 2019). Moreover, TuCAS has a 31% sequence identity with *Streptomyces lavendulae* O-ureido-L-serine synthase (PDB code: 3X43) which is necessary for D-cycloserine biosynthesis (Uda et al., 2015).

TuCAS is the first β-cyanoalanine synthase that is encoded by a metazoan genome whose structure was determined. Cystathionine β-synthases from βSAS family in honeybee (PDB code: 5OHX) and human (PDB code: 1M54) are the most similar to TuCAS in terms of their overall fold among the animal enzymes whose structures were determined. However, the cystathionine β-

synthases are significantly larger, as they include an additional domain that is responsible for heme binding (Taoka et al., 2002). Nonetheless, the overall sequence identity between the cystathionine β-synthases and TuCAS is low and in the cases of the honeybee and human enzymes (less than 41%). TSSM also has cystathionine β-synthase (Tetur06g06721) that shares 50% sequence identity with the homolog from honeybee and 55% sequence identity with the human protein. At the same time, TuCAS and Tetur06g06721 have only 36% identical residues. However, TuCAS is very similar to the homolog(s) from the spider mite *Tetranychus evansi* (93% sequence identity), the false spider mite *Brevipalpus yothersi* (bryot455g00050 and bryot357g00050 showing 85 and 89% identity to TuCAS, respectively; https://bioinformatics.psb.ugent.be/orcae/overview/Bryot) (Liu et al., 2021), as well as other arthropod and insect CAS enzymes that originate from various agricultural pests, like *Pieris rapae* (67% sequence identity) (Figure 6).

3.5.3 Active site

The electron densities in both TuCAS wild type crystal structures reported here are consistent with PLP in the active site of the enzyme, which forms a covalent bond with Lys52 (Figure 5). The lysine residue, which is critical for protein function (Yi et al., 2012), is located at the bottom of a cleft that is formed by the small and large protein domains (Figure 5A). In addition to the covalent linkage, Asn82 and Thr284 form hydrogen bonds with the PLP hydroxyl group and nitrogen from the pyridine ring, respectively. Asn82 is conserved in enzymes originating from very diverse organisms, and its amide group not only participates in hydrogen bond mediated interaction with PLP, but it also forms hydrogen bonds with the side chains of Asp313 and Lys37. The PLP phosphate group interactions with neighboring residues involve a network of hydrogen bonds. These bonds involve main chain amide groups of Gly194, Thr195, Gly196, and Cys198, as well as the hydroxyl group of Thr195 (Figure 5B). Moreover, two water molecules in the active site form hydrogen bonds with the phosphate moiety while simultaneously forming hydrogen bonds with the main chain atoms of Gly192 and Phe199, or side chain of His170.

The mode of PLP binding in the TuCAS active site is similar to homologous proteins. The PLP phosphate group forms hydrogen bonds with a highly conserved glycine- and threonine-rich loop (GTGGT) which is (194-GTGGC-198) in TuCAS. The 3-hydroxyl group of PLP forms a hydrogen bond to the side chain of an Asn82 from a highly conserved sequence stretch (EPTSGNTG) which is (77-EVTSGNQG-84) in TuCAS (Yi et al., 2012).

It was shown that in the case of OASS from *Arabidopsis*, residues Thr74 and Ser75 were involved in sulfur incorporation, while Asn77 and Gln147 play a role in O-acetyl-L-serine binding (Francois et al., 2006). These residues are conserved in TuCAS and correspond to Thr79, Ser80, Asn82, and Gln160, respectively. Thr79, Ser80, and Asn82 form a highly conserved loop (79-TSGN-82; Figure 6). In the higher resolution structure of TuCAS, we determined that there is an acetate ion in the vicinity of Ser80, which may mimic cyanide or sulfide ions. In addition, the acetate ion interacts with Asn106, which belongs to the sequence fragment (102-MSKGNS-107). This fragment is conserved in CASs originating from arthropods and insects (Figure 6). Gln160 is also a part of a highly conserved sequence fragment (Figure 6).

3.5.4. Flexible fragment of TuCAS

Co-crystallization of TuCAS with O-phospho-L-serine resulted in the stabilization of one of the disordered fragments in the TuCAS dimer. Residues 221-257 that form this part of the enzyme are critical for its activity as they close the opening to the active site next to the PLP. Comparison of TuCAS structure with the ordered loop (PDB code: 7FMJ) and similar structures

showed that residues 221-257 of TuCAS aligned well with residues 205-239 of *Planctomyces limnophila* OASS (PDB code: 5XA2) and residues 208-242 of *A. thaliana* OASS (PDB code: 2ISQ). The conformation of this region is very similar between these structures, and several residues forming this fragment are conserved. The overall sequence identities between TuCAS-*P. limnophila* OASS and TuCAS-*A. thaliana* OASS are 39% and 41% respectively. Residue 221-257 of TuCAS have also similar conformation to one observed for residues 231-265 of soybean CAS (PDB code: 3VBE) with conserved TuCAS Gly228, 240, 242, and Gly244 corresponding to residues Gly238, 248, 250, and 253 in soybean CAS. TuCAS and soybean CAS share 39% sequence identity. Interestingly, a comparison of another structure of soybean CAS (PDB code: 3VC3), which represents a complex with cysteine that is covalently linked to PLP, shows that the highly conserved loop (79-TSGN-82) moves toward the intermediate. The movement of Ser80 (Ser124 in soybean protein) is especially pronounced and corresponds to ~5Å shift of the Cα carbon of this residue.

On the other hand, comparison of the loops between TuCAS and *M. tuberculosis* OASS (PDB code: 3FGP) shows that residues 211-238 are aligned with residues 221-257 of TuCAS. In chain A, this fragment resembles the one observed in TuCAS. However, this segment of the *M. tuberculosis* OASS in chain B adopts a different conformation. While the function of this flexible fragment is not fully understood, in the crystal structure of a *M. tuberculosis* OASS in complex with a sulfur carrier protein (CysO) (PDB code: 3DWG), this loop is involved in the binding of the CysO. Furthermore, the flexible fragment in chain A and B of *M. tuberculosis* OASS complexed with CysO show a short alpha helix similar to the one seen in 3FGP. Chain A in 3DWG interacts with CysO through residues 206-246 which corresponds to residues 216-265 in TuCAS. Residue Val218, Pro225, Gly228, Leu238, Gly244, Val246, Pro247, Lue249, Asp250, Ser259, Ser260 in TuCAS corresponds to residues Val209, Pro216, Gly219, Leu225, Gly231, Val233, Pro234, Lue236, Asp238, Ser247, Ser248 of *M. tuberculosis* OASS.

4. Discussion

Cyanogenic precursors are important components of chemical plant defenses against herbivores. At the same time many mites and lepidopterans are able to live on cyanogenic plants. To further elucidate the molecular basis of the TSSM cyanide detoxification system we have focuses on TuCAS that was implicated in protection of *T. urticae* against cyanide poisoning (Wybouw et al., 2014). We produced a recombinant version of TuCAS is both enzymatically active and also forms well-diffracting crystals. The enzymatic activity of the recombinant TuCAS presented here is similar to one observed in previously reported recombinant TuCAS (Wybouw et al., 2014). In addition, it was shown that enzyme is most active in solutions with pH ~8.5, which is also in agreement with previously published data for CAS and OASS enzymes (Wybouw et al., 2014; Yi et al., 2012). Mutation of the PLP binding Lys52 resulted in an inactive enzyme, indicating that this residue is critical for TuCAS activity. This mutation also significantly decreases thermal stability of the enzyme.

The presented here recombinant TuCAS has a strong preference for the synthesis of β-cyanoalanine over cysteine. The nature of this preference is not well understood. Previous studies indicate that in plant β-substituted alanine synthases TSGN(M/T) motif is conserved. The Met residue from this motif is conserved in enzymes with higher efficiency in producing β-cyanoalanine using cysteine, while the threonine residue is present in enzymes with higher efficiency producing cysteine using OAS (Yi et al., 2012). However, in mite and insect CASs, there is a Gln (residue 83 in TuCAS) instead of Met residue (Figure 6). Also, the

V(M/S)GIG(S/T)G motif, which is conserved in plant βSASs is only partially conserved in TuCAS (residues 190-LVGCGTG-196). Our structural studies combined with mapping of the conserved residues on the TuCAS structure reveal that many highly conserved residues cluster not only in the PLP binding region but also further from the active site (Figure 7). The distribution of the conserved residues suggests that enzyme-substrate preferences may be controlled by molecule regions that are located further from the PLP binding site.

In vitro experiments utilized sulfide as a sulfur source in the reaction of cysteine synthesis. However, it is not known whether this is the only form of sulfur that is used by TuCAS *in vivo*. It is possible, despite the fact that we have not observed complex formation between TuCAS and Tetur36g00900, that *in vivo* the sulfur atom is donated through a thiocarboxyl group of a sulfur carrier protein. This was demonstrated in the case of the *M. tuberculosis* OASS (CysM) and CysO (Jurgenson et al., 2008). Our attempts in the use of thiocarboxylates instead of the sulfur carrier protein were also not successful, as thioacetic and thiobenzoic acids were not able to act as potential sulfur donors and mimickers of the C-terminal thiocarboxylate group of the sulfur carrier protein. Nonetheless, it is possible, that the TSSM genome encodes another sulfur carrier protein that acts together with TuCAS, but we were not able to identify it.

While thiocarboxylates did not act as surrogate substrates for TuCAS cysteine synthesis, we have been able to demonstrate that the protein can produce a number of cysteine derivatives. Such observations were reported previously for the enzymes from βSAS family, e.g. OASS from *Salmonella typhimurium* is able to use 5-thio-2-nitrobenzoate together with OAS to produce an L-cysteine derivative (Rabeh and Cook, 2004). We have tested three different mercaptobenzoic acids, and reaction products were observed when thiosalicylic acid (2-mercaptobenzoic acid) or 3-mercaptobenzoic acid were used. However, no reaction was observed when 4-mercaptobenzoic acid was used as substrates. In addition, we have observed more products formed when thiosalicylic acid was the substrate. Reaction with 5-thio-2-nitrobenzoate was reported in the case of *E. coli* OASS (CysM) and this compound was modeled in the active site of the bacterial enzyme (Zocher et al., 2007); however, there are no reports on reactivity of other thionitrobenzoates. Therefore, currently we do not know why TuCAS preferentially reacts with thiosalicylic acid or 3-mercaptobenzoic acid. It is possible that the flexible fragment of TuCAS, which closes on the active site plays a role in selection of the enzyme substrate.

The reaction involving a thiol group that leads to cysteine derivatives formation was not previously described for TuCAS. This advances our insights into the functional role(s) of laterally acquired CAS enzymes in herbivorous arthropods. For example, plant glucosinolates are an important class of defense compounds that are activated by endogenous myrosinases and produce unstable intermediates which are later converted to isothiocyanates or nitrils (Barba et al., 2016). These intermediates contain a sulfur atom that can potentially be linked to α-aminoacrylate that is generated by TuCAS (Figure 1D), leading to the formation of a cysteine derivative.

Until now the studies of CAS and OASS enzymes have mainly focused on reactions of cysteine synthesis and detoxification of cyanide with cysteine as a substrate (Figure 1A, B). The β-cyanoalanine synthesis results in the generation of sulfide that also has to be detoxified (Alvarez et al., 2012). However, we have demonstrated that the detoxification of CN^-, for example generated from cyanogenic precursors produced by plants, can be performed by TuCAS with the use of O-acetyl-L-serine. This reaction does not involve the formation of the sulfide, and therefore may be an interesting alternative in CN^- detoxification. In fact, the reaction, in which cyanide and OAS are used as TuCAS substrates can be considered as a sum of reactions in which cysteine or β-cyanoalanine are generated (Figure 1C).

TuCAS is a critical component of the xenobiotic metabolism of TSSM, and most likely in related plant-feeding mite species. For instance, a CAS homolog is also present in the genome of *Brevipalpus yothersi*, a false spider mite belonging to the Tenuipalpidae that diverged from the Tetranychidae more than 150 MYA (Liu, et al., 2021). However, there are likely other independent pathways that can be used to detoxify CN^-. For example, it was proposed that cyanide may be oxidized to cyanate by an unidentified enzyme or during a light catalyzed reaction, and later the cyanate is decomposed by cyanase (Wybouw et al., 2012; Schlachter et al., 2017). It is also plausible that TSSM has an additional way of detoxifying CN^- through a rhodanese. Rhodanese detoxifies cyanide by converting it to thiocyanate. While there is not yet such enzyme identified in TSSM, there is for example Tetur25g00350 (Uniprot id: T1KWX3) that has a rhodanese-like domain. TuCAS action on the cyanide and cysteine results not only in the generation of β-cyanoalanine, but it also generates an equivalent amount of sulfide. It is not exactly known what impact this compound has on mites. However, it was shown that in the case of plants, a higher concentration of HS^- has a toxic effect (Alvarez et al., 2012; Birke et al., 2012). Therefore, one can assume that the double activity of TuCAS allows not only for the detoxification of cyanide but at the same time allows for the detoxification of another potentially toxic compound, sulfide, that it can convert into cysteine. Importantly, it can also catalyze the reaction of β-cyanoalanine formation and detoxify cyanide by using OAS as a substrate, which allows to avoid the production of the toxic HS^-.

The structural and enzymatic activity studies provide a foundation for the development of novel chemical control agents that can target TuCAS activity and may help control TSSM populations feeding on cyanogenic plants. In pilot studies, we have tested several compounds (Table 2) that mostly showed relatively weak inhibitory properties with L-cycloserine and Compound 1 being the most effective inhibitors we characterized. Interestingly, our results indicate that TuCAS can differentiate between L and D isomers of cycloserine or serine, as the D forms of these compounds had no impact on the enzyme activity. We have also shown that TuCAS cannot use O-phospho-L-serine as an alternative to O-acetyl-L-serine substrate in cysteine synthesis. However, some bacterial OASSs are able to utilize O-phospho-L-serine in this reaction. In fact, in the case of TuCAS, O-phospho-L-serine has some weak inhibitory properties. TuCAS activity was also not significantly impacted by the presence of cyanates or thiocyanates. These compounds were investigated as they may appear as products of other alternative pathways leading to cyanide detoxification. Interestingly, cyanates and thiocyanates were competitive inhibitors against CN^- and they were noncompetitive inhibitors against sulfide ions. This strongly suggests that TuCAS may discriminate between cyanide and sulfide ions, and the preference toward reaction of β-cyanoalanine formation is associated with the selectivity towards small nucleophiles. For example, it was suggested that in the case of MtOASS (CysM) the C-terminal end of the protein and conformational changes of surface loops act together as a selectivity filter, which protects the α-aminoacrylate intermediate against attack of various nucleophiles and formation of unwanted products (Ågren et al, 2009). The sequence comparison (Figure 6) indicates that bacterial and plant OASSs have different C-terminal ends when compared with animal CASs, which may indicate that this fragment of the proteins can potentially play a role in selection of the substrates, and therefore bias enzyme preferences toward cysteine or β-cyanoalanine formation.

Compounds 1 and 2 were shown to bind *M. tuberculosis* OASS with dissociation constants of 0.32±0.01 μM and 3.4±0.1 μM respectively (Brunner et al., 2016). The compounds from this group also displayed antibacterial activity and were able to inhibit *M. tuberculosis* growth when applied in low micromolar concentrations. Compound 1 was also able to inhibit the activity of

TuCAS; however, we have not observed such inhibition in the case of Compound 2. Comparison of the crystal structure of the Compound 1 in complex with *T. tuberculosis* OASS (PDB code: 7I7A) with TuCAS structure allows for the identification of residues from the TSSM enzyme that most likely are involved in interaction with the inhibitor (Figure 8). Many of these residues are identical or very similar in the bacterial and TSSM enzymes. However, the major difference is related to the flexible region that closes the active site. It is likely that this region is responsible for selectivity that we observe in TuCAS interactions with Compound 1 and 2. However, despite many attempts, we were not able to obtain crystal structures of the TSSM enzyme in the complex with any inhibitors. Therefore, the interactions between TuCAS and these compounds remain unknown. Compound 1 will have to be optimized in order to improve its inhibitory properties, as well as tested if it (or its derivatives) can impact performance of TSSM feeding on cyanogenic plants, before it can be considered as a potential lead for the development of acaricides.

In summary, we have shown that TuCAS is a versatile enzyme that is able to modify many substrates. It is likely that the ability for TuCAS to process a wide range of substrates allows TSSM to use this enzyme in different processes, and detoxification of cyanide is only one of them.

Acknowledgments

We would like to thank Dr. Caryn Outten for access to the microplate reader, Dr. Robert Jedrzejczak for help with the cloning protocol, and Safaa Kader for reading the manuscript and valuable comments.

Structural results shown in this report are derived from data collected at Southeast Regional Collaborative Access Team (SER-CAT; 22 ID) beamline at the Advanced Photon Source, Argonne National Laboratory. Supporting institutions may be found at www.ser-cat.org/members.html. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract Nos. DE-AC02-06CH11357 and W-31-109-Eng-38. This work was partially supported by ASPIRE II and ASPIRE III grants from the Office of the Vice President of Research at the University of South Carolina. M.G. and V.G. acknowledge funding by the Ontario Research Fund−Research Excellence Round 8 (RE08-067). N.W. was supported by a BOF post-doctoral fellowship (Ghent University, 01P03420). T.V.L. was supported by a Fund for Scientific Research Flanders (FWO) grant G009312N while T.V.L. and W.D. were supported by FWO grant G053815N. T.V.L., V.G., and M.G. were supported by the European Commission (EC contract 618105) via FACCE ERA-NET Plus and FACCE-JP (Genomite, project ID 137). T.V.L. has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement n° 772026). This project was partially funded by USDA's National Institute of Food and Agriculture, award #2020-67014-31179 through the NSF/NIFA Plant Biotic Interactions Program.

Table 1. Summary of enzymatic activity parameters for TuCAS catalyzed reactions.

	β-Cyanoalanine synthases		**Cysteine synthases**	
Substrate	L-cysteine	KCN	O-acetyl-L-serine	Na_2S
Hill coefficient	2.40 ± 0.50	2.6 ± 0.4	2.0 ± 0.2	2.7 ± 0.7
k_{cat} (min^{-1})	31.7	31.0	3.60	5.10
K_M (mM)	1.60 ± 0.20	2.30 ± 0.20	4.3 ± 0.3	4.5 ± 0.4
Vmax (μM min^{-1})	1.03 ± 0.09	1.01 ± 0.06	1.22 ± 0.03	1.76 ± 0.12
k_{cat}/K_M ($mM^{-1}min^{-1}$)	19.8	13.5	0.80	1.10

Table 2. Inhibitor constants (K_i) for inhibitors of β-cyanoalanine and L-cysteine synthesis reactions.

Inhibitor	**K_i (β-cyanoalanine synthesis)**	**K_i (L-cysteine synthesis)**
L-Serine	0.9 ± 0.2 mM	4.3 ± 0.3 mM
L-Cycloserine	130 ± 8 μM	74.0 ± 8.5 μM
O-Phospho-L-serine	53 ± 5 mM	7.0 ± 0.8 mM
S-Methyl-L-cysteine	3.7 ± 0.2 mM	N/A*
Sodium thiocyanate	3.5 ± 0.3 mM	13.2 ± 0.4 mM
Sodium cyanate	9.0 ± 1.0 mM	15.9 ± 3.5 mM
Compound 1**	150 ± 10 μM	140 ± 20 μM

* Our enzyme activity assay is not compatible with this compound.
** 3-(3-(3,4-dichlorophenyl)ureido)benzoic acid

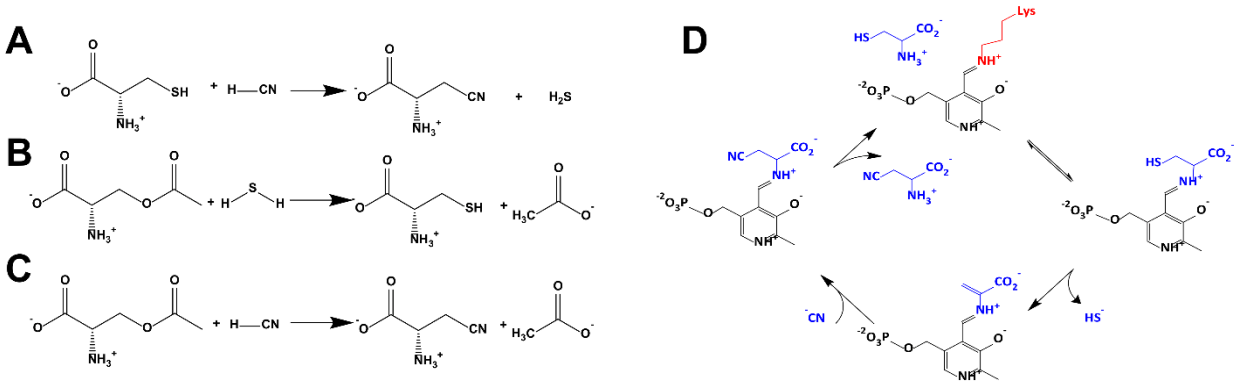


Figure 1. Reactions catalyzed by TuCAS. A) Reaction of cyanide detoxification with cysteine as a substrate. B) Cysteine synthesis catalyzed by the enzyme. C) A new alternative route for cyanide detoxification. D) Reaction mechanism of CAS. In the case of OASS activity, cysteine and cyanide are replaced with O-acetyl-L-serine and sulfide, respectively (Yi et al., 2012).

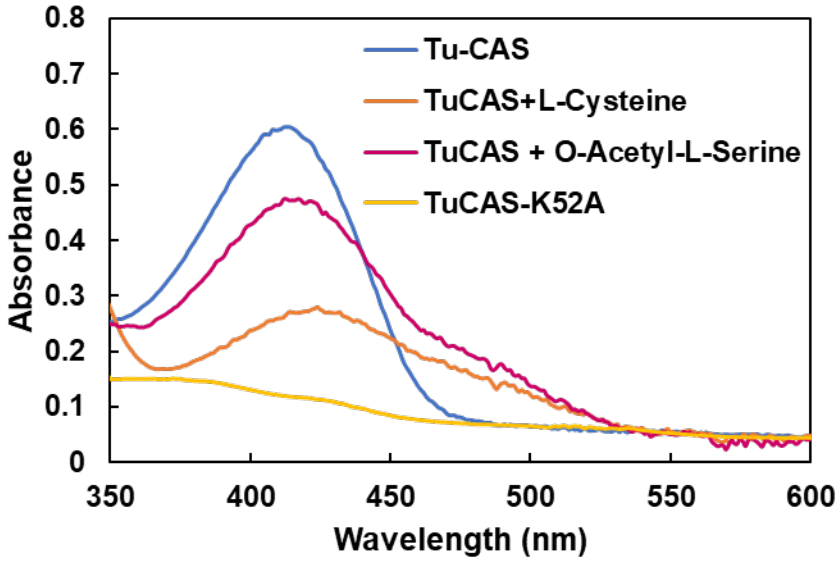


Figure 2. UV spectroscopy analysis of TuCAS. Absorption spectra of 35 μM TuCAS (blue), TuCAS with the addition of 1 mM L-cysteine (orange), and TuCAS with the addition of 1mM O-acetyl-L-serine (magenta). Absorption spectra of 35 μM K52A TuCAS are shown in yellow.

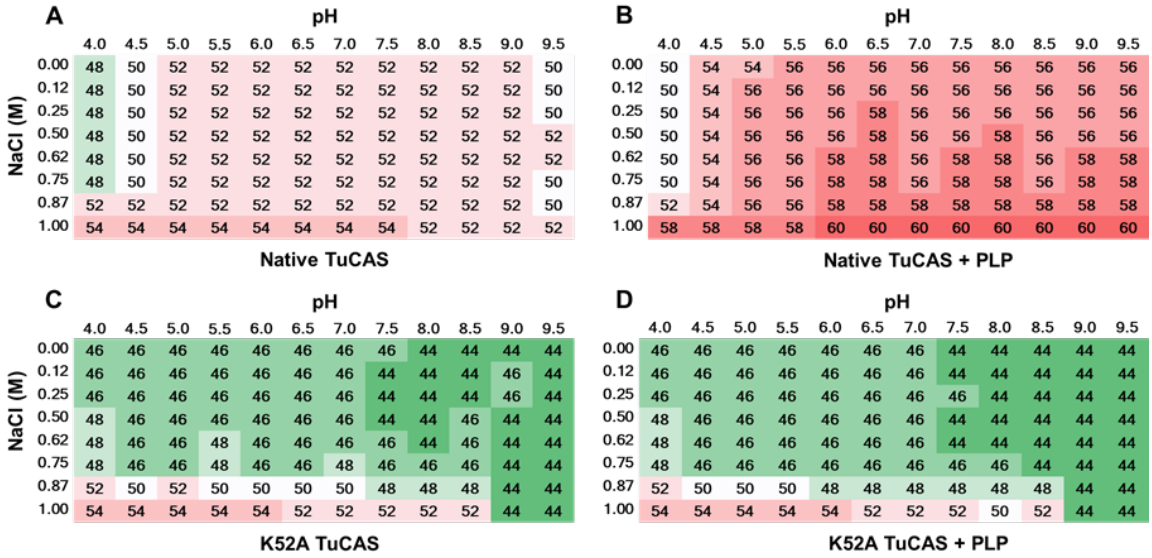


Figure 3. DSF results showing differences in TuCAS thermal stability. A) Native TuCAS. B) Native TuCAS incubated with 500 µM PLP. C) TuCAS K52A. D) TuCAS K52A incubated with 500 µM PLP.

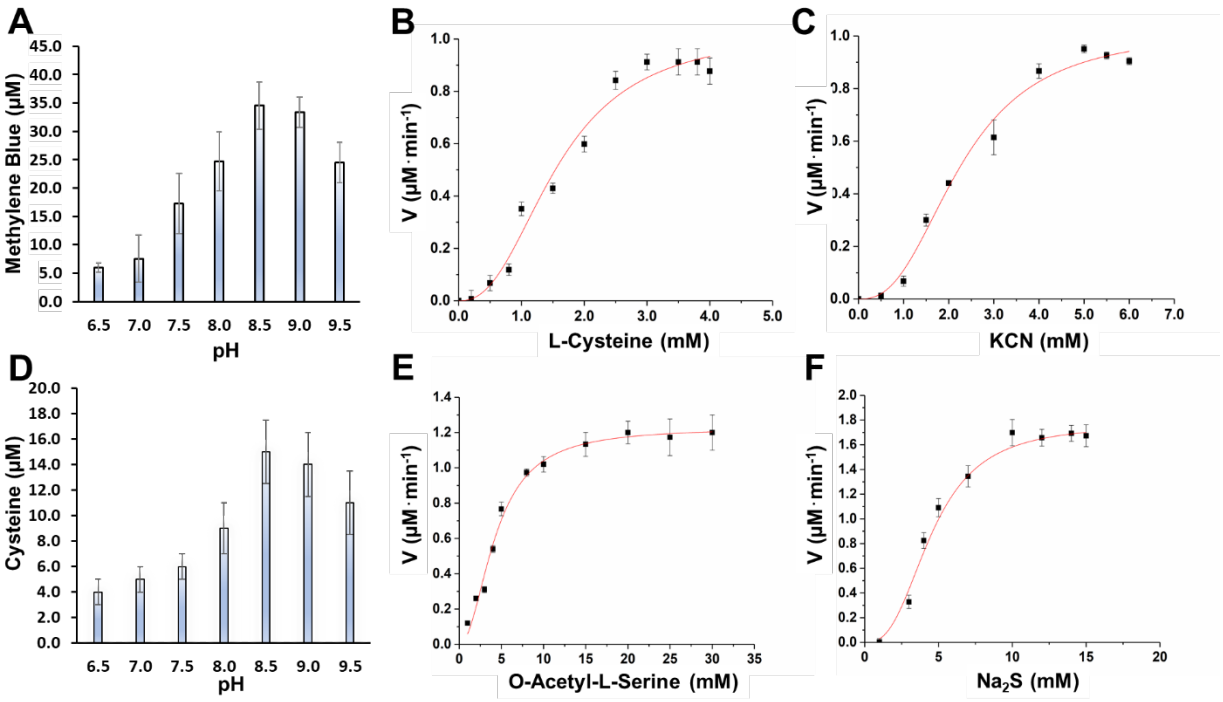


Figure 4. Summary on enzyme activity and kinetics. A) pH profile of TuCAS in the formation of β-cyanoalanine. B) Hill equation graph representing the initial velocity ($\mu M \cdot min^{-1}$) vs L-cysteine concentration (mM) in pH 8.5. C) Hill equation graph representing the initial velocity ($\mu M \cdot min^{-1}$) vs KCN concentration (mM) in pH 8.5. D) pH activity profile of TuCAS in the formation of L-cysteine. E) Hill equation graph representing the initial velocity ($\mu M \cdot min^{-1}$) vs O-acetyl-L-serine concentration (mM) in pH 8.5. F) Hill equation graph representing the initial velocity ($\mu M \cdot min^{-1}$) vs Na_2S concentration (mM) in pH 8.5.

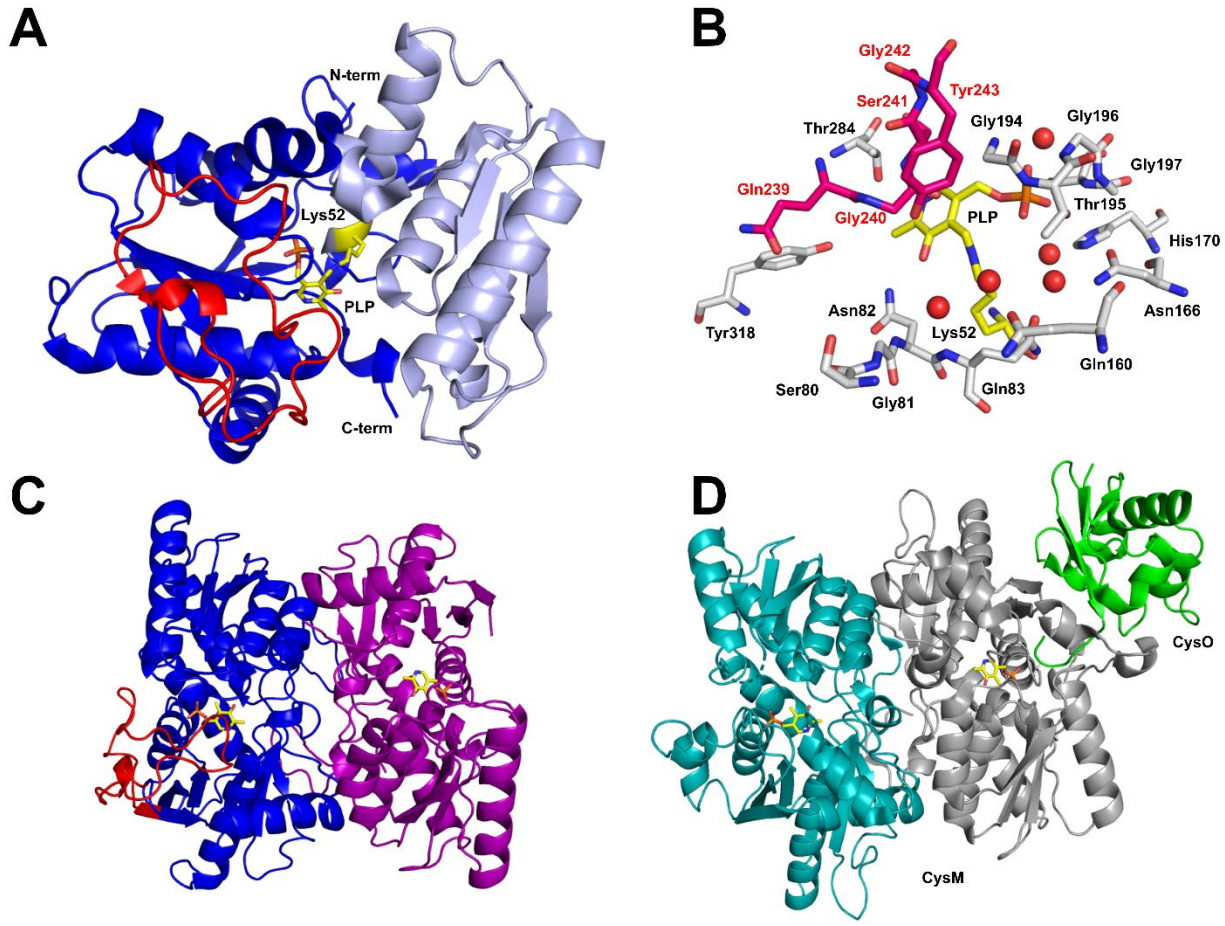


Figure 5. Structural features of TuCAS. A) Crystal structure of a single TuCAS chain. The smaller domain is marked in light blue, and the larger domain is shown in blue and red. The flexible fragment is marked in red. PLP and Lys52 are shown in stick representation and colored according to the type of the atom with carbon atoms in yellow. B) Residues forming TuCAS active center. Water molecules are shown as red spheres. The residues originating from the flexible fragment are labeled in red. C) Dimer of TuCAS. Only one protein chain has the flexible fragment ordered. This fragment is marked in blue. D) Dimer of *M. tuberculosis* CysM in complex with the sulfur carrier protein (CysO; green; PDB code: 3DWG).

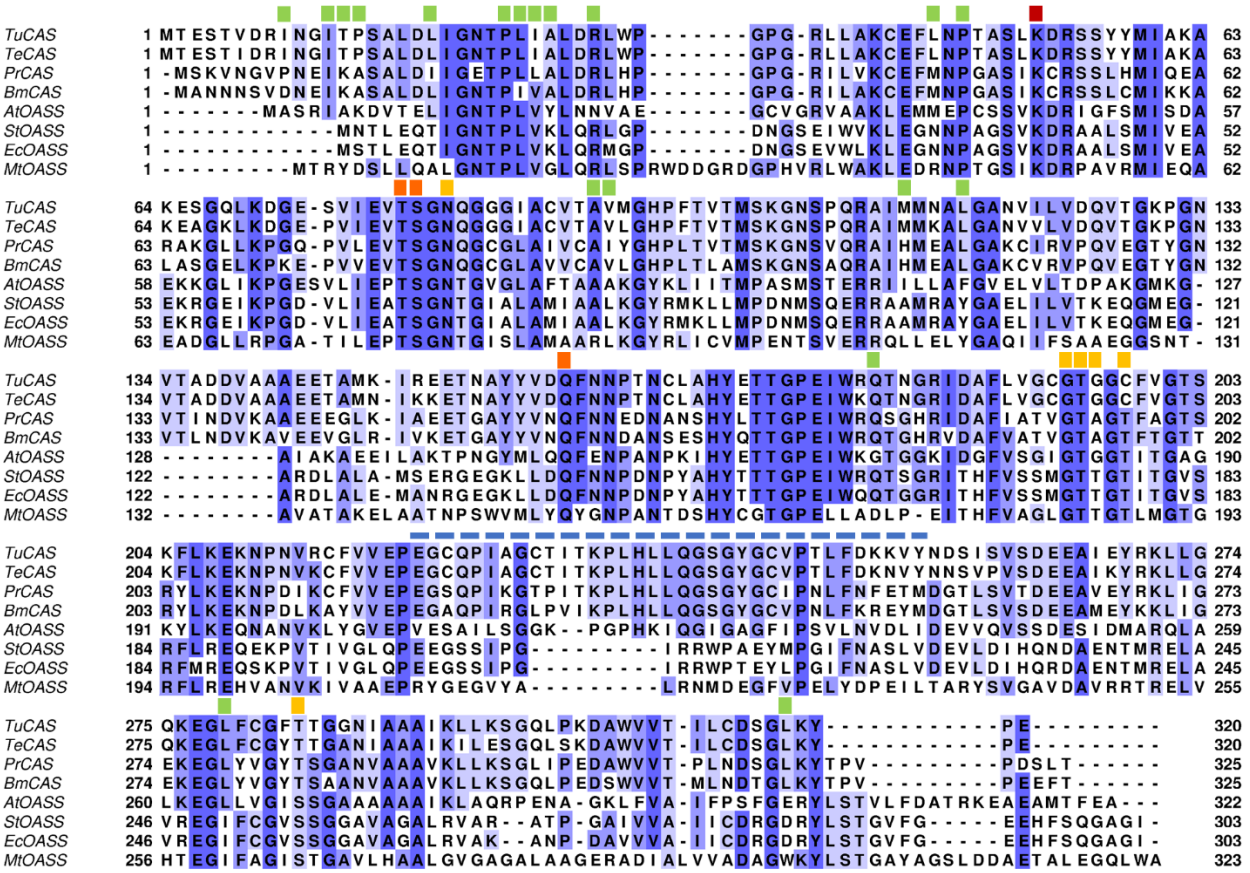


Figure 6. Sequence alignment of several proteins from the βSAS family. TuCAS – CAS from *T. urticae*, TeCAS – CAS from *T. evansi*, TuCAS – CAS from *P. rapae*, BmCAS – CAS from *Bombyx mori*, AtOASS – OASS from *A. thaliana*, StOASS – OASS from *Salmonella typhimurium*, EcOASS – OASS from *E. coli*, and MtOASS – OASS from *M. tuberculosis*. Green squares mark residues participating in dimer formation. Lys52 is marked with a dark red square, and residues forming hydrogen bonds with PLP ring fragment are marked with light red squares. Orange squares indicate residues responsible for anchoring of the PLP phosphate moiety. The flexible TuCAS region is marked with dashed line.

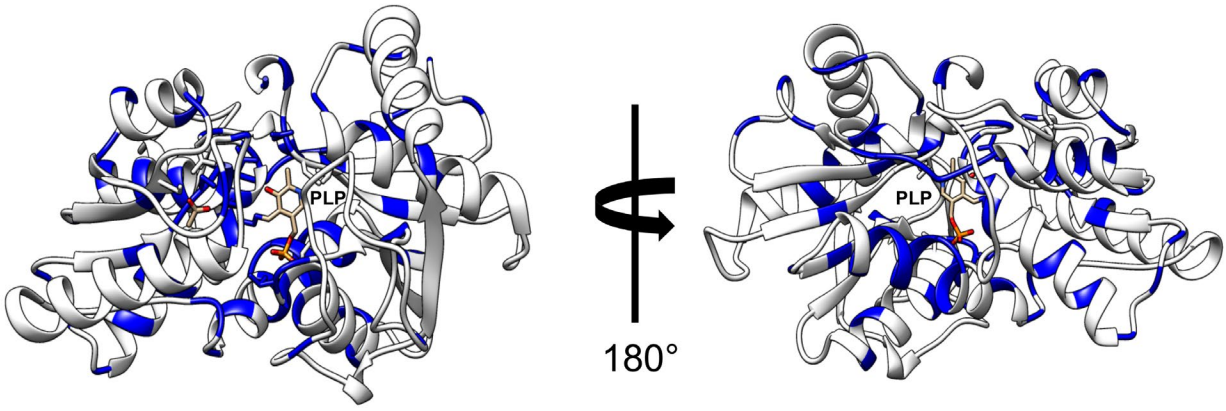
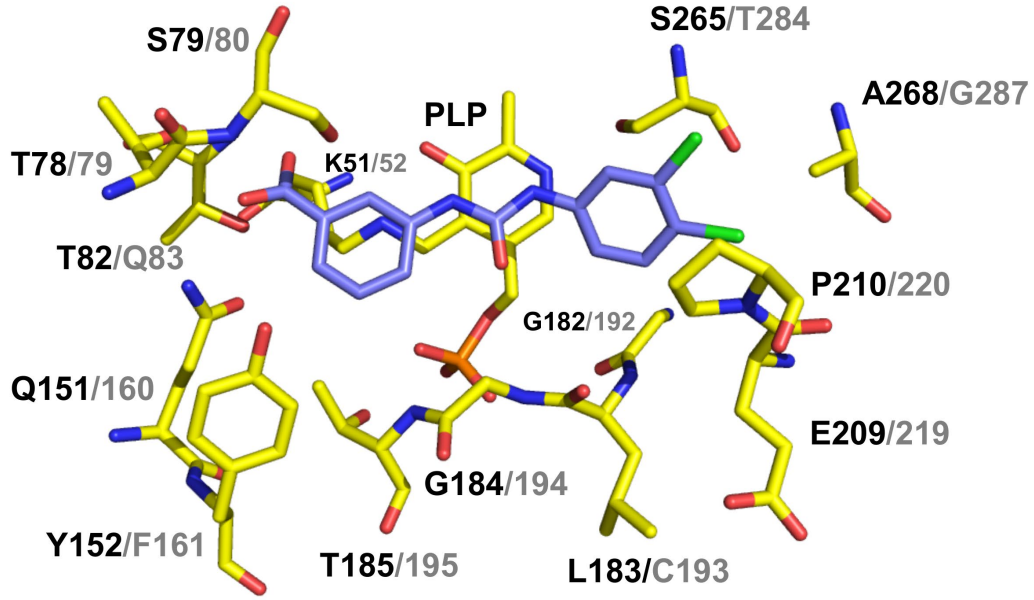


Figure 7. Cartoon representation of TuCAS with conserved residues shown in blue. The conserved residues were selected based on the sequence alignment presented in Figure 6.



825
826 **Figure 8.** Compound 1 in complex with *M. tuberculosis* CysM (PDB code: 5I7A). Numbering of
827 residues in black corresponds to CysM, while the numbering in grey corresponds to TuCAS.
828
829
830

References

Altschul, S., F., Gish, W., Miller, W., Myers E., W., Lipman, D., J., 1990. Basic local alignmentsearch tool. J. Molec. Biol., 215, 403-410.

Álvarez, C., García, I., Romero, L.C., Gotor, C., 2012. Mitochondrial sulfide detoxification requires a functional isoform O-acetylserine (thiol) lyase C in Arabidopsis thaliana. Molec. Plant 5(6):1217-1226.

Arenas-Alfonseca, L., Gotor, C., Romero, L.C., García, I., 2018. ß-Cyanoalanine synthase action in root hair elongation is exerted at early steps of the root hair elongation pathway and is independent of direct cyanide inactivation of NADPH oxidase. Plant Cell Physiol. 59, 1072-1083.

Ågren, D., Schnell, R., Schneider, G., 2009. The C-terminal of CysM from Mycobacterium tuberculosis protects the aminoacrylate intermediate and is involved in sulfur donor selectivity. FEBS letters. 583(2):330-336.

Ballhorn D.J., Heil M., Lieberei R., 2006. Phenotypic plasticity of cyanogenesis in lima bean Phaseolus lunatus—activity and activation of β-glucosidase. J. Chem. Ecol. 32(2):261-275.

Barba, F.J., Nikmaram, N., Roohinejad, S., Khelfa, A., Zhu, Z., Koubaa, M., 2016. Bioavailability of glucosinolates and their breakdown products: Impact of processing. Front. Nutr. 3, 24.

Beran, F., Köllner, T.G. Gershenzon, J., Tholl, D., 2019. Chemical convergence between plants and insects: biosynthetic origins and functions of common secondary metabolites. New Phytol., 223, 52-67.

Berg, S.P., Krogmann, D., 1975. Mechanism of KCN inhibition of photosystem I. J. Biol. Chem. 250, 8957-8962.

Birke, H., Haas, F.H., De Kok, L.J., Balk, J., Wirtz, M., Hell, R., 2012. Cysteine biosynthesis, in concert with a novel mechanism, contributes to sulfide detoxification in mitochondria of Arabidopsis thaliana. Biochem. J. 445(2):275-283.

Berman, H., Henrick, K., Kleywegt, G., Nakamura, H., Markley, J., 2012. PDBe.

Bethke, P.C., Libourel, I.G., Reinöhl, V., Jones, R.L., 2006. Sodium nitroprusside, cyanide, nitrite, and nitrate break Arabidopsis seed dormancy in a nitric oxide-dependent manner. Planta 223, 805-812.

Bishop, N.I., Lumry, R., Spikes, J.D., 1955. The mechanism of the photochemical activity of isolated chloroplasts. I. Effect of temperature. Archives of biochemistry and biophysics 58, 1-18.

Bishop, N.I., Spikes, J.D., 1955. Inhibition by cyanide of the photochemical activity of isolated chloroplasts. Nature 176, 307-308.

Bleecker, A.B., Kende, H., 2000. Ethylene: a gaseous signal molecule in plants. Annu. Rev. Cell Dev. Biol. 16, 1-18.

Bonner, E.R., Cahoon, R.E., Knapke, S.M., Jez, J.M., 2005. Molecular basis of cysteine biosynthesis in plants structural and functional analysis of o-acetylserine sulfhydrylase from Arabidopsis thaliana. J. Biol. Chem. 280, 38803-38813.

Böttcher, C., Westphal, L., Schmotz, C., Prade, E., Scheel, D., Glawischnig, E., 2009. The multifunctional enzyme CYP71B15 (PHYTOALEXIN DEFICIENT3) converts cysteine-indole-3-acetonitrile to camalexin in the indole-3-acetonitrile metabolic network of Arabidopsis thaliana. The Plant Cell 21, 1830-1845.

Brunner, K., Maric, S., Reshma, R.S., Almqvist, H., Seashore-Ludlow, B., Gustavsson, A.L., Poyraz, O., Yogeeswari, P., Lundback, T., Vallin, M., Sriram, D., Schnell, R., Schneider, G., 2016b. Inhibitors of the Cysteine Synthase CysM with Antibacterial Potency against Dormant Mycobacterium tuberculosis. J. Med. Chem. 59, 6848-6859.

876 Burow, M., Bergner, A., Gershenzon, J., Wittstock, U., 2007. Glucosinolate hydrolysis in
877 Lepidium sativum—identification of the thiocyanate-forming protein. Plant. Mol. Biol. 63, 49-61.
878 Consortium, U., 2019. UniProt: a worldwide hub of protein knowledge. Nucleic acids
879 Res.47(D1):D506-D515.
880 Chivasa, S., Carr, J.P., 1998. Cyanide restores N gene–mediated resistance to tobacco mosaic virus
881 in transgenic tobacco expressing salicylic acid hydroxylase. The Plant Cell 10, 1489-1498.
882 Cohen, W.S., McCarty, R.E., 1976. Reversibility of the cyanide inhibition of electron transport in
883 spinach chloroplast thylakoids. Biochem. Biophys. Res. Comms 73, 679-685.
884 Cooper, C.E., Brown, G.C., 2008. The inhibition of mitochondrial cytochrome oxidase by the
885 gases carbon monoxide, nitric oxide, hydrogen cyanide and hydrogen sulfide: chemical
886 mechanism and physiological significance. J. Bioenerg. . Biomembr. 40, 533.
887 Daneshian, L., Schlachter, C., Timmers, L., F., S., M., Radford, T., Kapingidza, B., Dias, T, Liese,
888 J., Sperotto, R., A., Grbic, V., Grbic, M., 2021. Delta class glutathione S-transferase (TuGSTd01)
889 from the two-spotted spider mite Tetranychus urticae is inhibited by abamectin. Pesticide
890 Biochem. Physiol.176: 104873
891 Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W.,
892 Arendall III, W.B., Snoeyink, J., Richardson, J.S., 2007. MolProbity: all-atom contacts and
893 structure validation for proteins and nucleic acids. Nucleic Acids Res. 35, W375-W383.
894 Donato, A.J., Eskurza, I., Silver, A.E., Levy, A.S., Pierce, G.L., Gates, P.E., Seals, D.R., 2007.
895 Direct evidence of endothelial oxidative stress with aging in humans: relation to impaired
896 endothelium-dependent dilation and upregulation of nuclear factor-κB. Circulation Res. 100,
897 1659-1666.
898 Ellis, K.J., Morrison, J.F., 1982. [23] Buffers of constant ionic strength for studying pH-dependent
899 processes, Methods in enzymology. Elsevier, pp. 405-426.
900 Emsley, P., Cowtan, K., 2004. Coot: model-building tools for molecular graphics. Acta
901 Crystallogr. Sec. D: Biol. Crystallogr. 60, 2126-2132.
902 Eschenfeldt, W.H., Makowska-Grzyska, M., Stols, L., Donnelly, M.I., Jedrzejczak, R.,
903 Joachimiak, A., 2013. New LIC vectors for production of proteins from genes containing rare
904 codons. J. Struct.. Funct. Gen. 14, 135-144.
905 Francois, J.A., Kumaran, S., Jez, J.M., 2006. Structural basis for interaction of O-acetylserine
906 sulfhydrylase and serine acetyltransferase in the Arabidopsis cysteine synthase complex. The Plant
907 Cell 18, 3647-3655.
908 Froger, A., Hall, J.E., 2007. Transformation of plasmid DNA into E. coli using the heat shock
909 method. J.Visual. Exper., e253.
910 García-Sánchez, S., Bernales, I., Cristobal, S., 2015. Early response to nanoparticles in the
911 Arabidopsis transcriptome compromises plant defence and root-hair development through
912 salicylic acid signalling. BMC genomics 16, 341.
913 García, I., Arenas-Alfonseca, L., Moreno, I., Gotor, C., Romero, L.C., 2019. HCN regulates
914 cellular processes through posttranslational modification of proteins by S-cyanylation. Plant
915 Physiol. 179, 107-123.
916 García, I., Castellano, J.M., Vioque, B., Solano, R., Gotor, C., Romero, L.C., 2010. Mitochondrial
917 β-cyanoalanine synthase is essential for root hair formation in Arabidopsis thaliana. The Plant Cell
918 22, 3268-3279.
919 García, I., Rosas, T., Bejarano, E.R., Gotor, C., Romero, L.C., 2013. Transient transcriptional
920 regulation of the CYS-C1 gene and cyanide accumulation upon pathogen infection in the plant
921 immune response. Plant Physiol. 162, 2015-2027.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., Bairoch, A., 2003. ExPASy: The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. 31, 3784-3788.
Goldschmidt, L., Cooper, D.R., Derewenda, Z.S., Eisenberg, D., 2007. Toward rational protein crystallization: A Web server for the design of crystallizable protein variants. Protein Sci. 16, 1569-1576.
Grbic, M., Van Leeuwen, T., Clark, R.M., Rombauts, S., Rouze, P., Grbic, V., Osborne, E.J., Dermauw, W., Ngoc, P.C., Ortego, F., Hernandez-Crespo, P., Diaz, I., Martinez, M., Navajas, M., Sucena, E., Magalhaes, S., Nagy, L., Pace, R.M., Djuranovic, S., Smagghe, G., Iga, M., Christiaens, O., Veenstra, J.A., Ewer, J., Villalobos, R.M., Hutter, J.L., Hudson, S.D., Velez, M., Yi, S.V., Zeng, J., Pires-daSilva, A., Roch, F., Cazaux, M., Navarro, M., Zhurov, V., Acevedo, G., Bjelica, A., Fawcett, J.A., Bonnet, E., Martens, C., Baele, G., Wissler, L., Sanchez-Rodriguez, A., Tirry, L., Blais, C., Demeestere, K., Henz, S.R., Gregory, T.R., Mathieu, J., Verdon, L., Farinelli, L., Schmutz, J., Lindquist, E., Feyereisen, R., Van de Peer, Y., 2011. The genome of Tetranychus urticae reveals herbivorous pest adaptations. Nature 479, 487-492.
Hatzfeld, Y., Lee, S., Lee, M., Leustek, T., Saito, K., 2000. Functional characterization of a gene encoding a fourth ATP sulfurylase isoform from Arabidopsis thaliana. Gene 248, 51-58.
Holm, L., Rosenstrom, P., 2010. Dali server: conservation mapping in 3D. Nucleic Acids Res. 38, W545-W549.
Ikegami, F., Takayama, K., Tajima, C., Murakoshi, I., 1988. Purification and properties of β-cyano-L-alanine synthase from Spinacia oleracea. Phytochem. 27, 2011-2016.
Isom, G.E., Way, J.L., 1984. Effects of oxygen on the antagonism of cyanide intoxication: cytochrome oxidase, in vitro. Toxicol. Appl. Pharmacol. 74, 57-62.
Jeppson, L.R., Keifer, H.H., Baker, E.W., 1975. Mites injurious to economic plants. Univ of California Press.
Jurgenson, C.T., Burns, K.E., Begley, T.P., Ealick, S.E., 2008. Crystal structure of a sulfur carrier protein complex found in the cysteine biosynthetic pathway of Mycobacterium tuberculosis. Biochem. 47, 10354-10364.
Krissinel, E., Henrick, K., 2004. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr. 60, 2256-2268.
Krissinel, E., Henrick, K., 2007. Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. 372, 774-797.
Lai, K.W., Yau, C.P., Tse, Y.C., Jiang, L., Yip, W.K., 2009. Heterologous expression analyses of rice OsCAS in Arabidopsis and in yeast provide evidence for its roles in cyanide detoxification rather than in cysteine synthesis in vivo. J. Exper. Botany. 60, 993-1008.
Laskowski, R.A., 2017. The ProFunc Function Prediction Server. Methods Mol. Biol. 1611, 75-95.
Li, Y., Zhou, Y., Jing, W., Xu, S., Jin, Y., Xu, Y., 2021. Horizontally acquired cysteine synthase genes undergo functional divergence in lepidopteran herbivores. *Heredity*:1-14.
Lineweaver, H., Burk, D., 1934. The determination of enzyme dissociation constants. J. . Amer. Chem. Soc. 56, 658-666.
Liu, Q., Deng Y., Song, A., Xiang, Y., Chen, D., Wei, L., 2021. Comparative analysis of mite genomes reveals positive selection for diet adaptation. Commun. Biol. 4(1):1-10.
Maruyama, A., Saito, K., Ishizawa, K., 2001. β-Cyanoalanine synthase and cysteine synthase from potato: molecular cloning, biochemical characterization, and spatial and hormonal regulation. Plant Mol. Biol. 46, 749-760.

967 Minor, W., Cymborowski, M., Otwinowski, Z., Chruszcz, M., 2006. HKL-3000: the integration
968 of data reduction and structure solution–from diffraction images to an initial model in minutes.
969 Acta Crystallographica Section D: Biol. Crystallogr. 62, 859-866.
970 Murshudov, G.N., Skubak, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn,
971 M.D., Long, F., Vagin, A.A., 2011. REFMAC5 for the refinement of macromolecular crystal
972 structures. Acta Crystallogr. Sec. D-Biol. Crystallogr. 67, 355-367.
973 Nagahara, N., Ito, T., Minami, M., 1999. Invited Reviews-Mercaptopyruvate sulfurtransferase as
974 a defense against cyanide toxication: Molecular properties and mode of detoxification. Histol. and
975 Histopathol. 14, 1277-1286.
976 Okonji, R., Adediji, O., Itakorode, B., Torimiro, N., Onwudiegwu, C., 2018. Purification and
977 characterization of betacyanoalanine synthase from Pseudomonas straminea. Intern. J. Biol. .
978 Chem. Sci. 12, 1086-1101.
979 Otwinowski, Z., Minor, W., 1997. Processing of X-ray diffraction data collected in oscillation
980 mode. Macromol. Crystallogr. A 276: 307–326.
981 Painter, J., Merritt, E.A., 2006. TLSMD web server for the generation of multi-group TLS models.
982 J. Appl. Crystallogr. 39, 109-111.
983 Petronikolou, N., Ortega, M.A., Borisova, S.A., Nair, S.K., Metcalf, W.W., 2019. Molecular basis
984 of Bacillus subtilis ATCC 6633 self-resistance to the phosphono-oligopeptide antibiotic
985 rhizocticin. ACS cChem. Biol. 14, 742-750.
986 Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin,
987 T.E., 2004. UCSF Chimera--a visualization system for exploratory research and analysis. J.
988 Comput. Chem. 25, 1605-1612.
989 Piotrowski, M., 2008. Primary or secondary? Versatile nitrilases in plant metabolism. Phytochem.
990 69, 2655-2667.
991 Ponstingl, H., Henrick, K., Thornton, J.M., 2000. Discriminating between homodimeric and
992 monomeric proteins in the crystalline state. Proteins 41, 47-57.
993 Rabeh, W.M., Cook, P.F., 2004. Structure and mechanism of O-acetylserine sulfhydrylase. J. .Biol.
994 Chem. 279, 26803-26806.
995 Rioja, C., Zhurov, V., Bruinsma, K., Grbic, M., Grbic, V., 2017. Plant-Herbivore Interactions: A
996 Case of an Extreme Generalist, the Two-Spotted Spider Mite Tetranychus urticae. Mol. Plant
997 Microbe. Interact. 30, 935-945.
998 Sang, Y., Barbosa, J.M., Wu, H., Locy, R.D., Singh, N.K., 2007. Identification of a pyridoxine
999 (pyridoxamine) 5′ -phosphate oxidase from Arabidopsis thaliana. FEBS letters 581, 344-348.
1000 Schnell, R., Oehlmann, W., Singh, M., Schneider, G., 2007. Structural insights into catalysis and
1001 inhibition of O-acetylserine sulfhydrylase from Mycobacterium tuberculosis: crystal structures of
1002 the enzyme α-aminoacrylate intermediate and an enzyme-inhibitor complex. J. Biol.l Chem. 282,
1003 23473-23481.
1004 Schlachter, C.R., Klapper, V., Wybouw, N., Radford, T., Van Leeuwen, T., Grbic, M., Chruszcz, M., 2017.
1005 Structural Characterization of a Eukaryotic Cyanase from Tetranychus urticae. J Agric Food Chem 65,
1006 5453-5462.
1007 Schlachter, C., R., Daneshian, L., Amaya, J., Klapper, V., Wybouw, N., Borowski, T., Van
1008 Leeuwen, T., Grbic, V., Grbic, M., Makris, T., M., 2019. Structural and functional characterization
1009 of an intradiol ring-cleavage dioxygenase from the polyphagous spider mite herbivore Tetranychus
1010 urticae Koch. Insect Biochem. Molec. Biol. 107, 19-30.
1011 Seo, S., Mitsuhara, I., Feng, J., Iwai, T., Hasegawa, M., Ohashi, Y., 2011. Cyanide, a coproduct
1012 of plant hormone ethylene biosynthesis, contributes to the resistance of rice to blast fungus. Plant
1013 Physiol. 155, 502-514.

Taoka, S., Lepore, B.W., Kabil, Ö., Ojha, S., Ringe, D., Banerjee, R., 2002. Human cystathionine β-synthase is a heme sensor protein. Evidence that the redox sensor is heme and not the vicinal cysteines in the CXXC motif seen in the crystal structure of the truncated enzyme. Biochem. 41, 10454-10461.
Taylorson, R., Hendricks, S., 1973. Promotion of seed germination by cyanide. Plant Physiol. 52, 23-27.
Trebst, A., Losada Villasante, M., Arnon, D.I., 1960. Photosynthesis by isolated chloroplasts XII. Inhibitors of Co2 assimilation in a reconstituted chloroplast system. J. Biol. Chem., 235 (3), 840-844.
Uda, N., Matoba, Y., Oda, K., Kumagai, T., Sugiyama, M., 2015. The structural and mutational analyses of O-ureido-l-serine synthase necessary for d-cycloserine biosynthesis. The FEBS J. 282, 3929-3944.
Vagin, A., Teplyakov, A., 1997. MOLREP: an automated program for molecular replacement. J. Appl. Crystallogr. 30, 1022-1025.
van Ohlen, M., Herfurth, A.-M., Kerbstadt, H., Wittstock, U., 2016. Cyanide detoxification in an insect herbivore: Molecular identification of β-cyanoalanine synthases from Pieris rapae. Insect Biochem. Mol. Biol. 70, 99-110.
Vozdek, R., Hnízda, A., Krijt, J., Šerá, L., Kožich, V., 2013. Biochemical properties of nematode O-acetylserine (thiol) lyase paralogs imply their distinct roles in hydrogen sulfide homeostasis. Bioch. Biophys. Acta. Proteins and Proteomics; 1834(12):2691-2701.
Warrilow, A.G., Hawkesford, M.J., 2000. Cysteine synthase (O-acetylserine (thiol) lyase) substrate specificities classify the mitochondrial isoform as a cyanoalanine synthase. J. Exper. Botany 51, 985-993.
Watanabe, A., Yano, K., Ikebukuro, K., Karube, I., 1998. Cyanide hydrolysis in a cyanide-degrading bacterium, Pseudomonas stutzeri AK61, by cyanidase. Microbiol. 144, 1677-1682.
Watanabe, M., Kusano, M., Oikawa, A., Fukushima, A., Noji, M., Saito, K., 2008. Physiological roles of the β-substituted alanine synthase gene family in Arabidopsis. Plant phys. 146(1):310-320.
Wing, D., Patel, H., Baskin, S., 1992. The effect of picrylsulphonic acid on in vitro conversion of cyanide to thiocyanate by 3-mercaptopyruvate sulphurtransferase and rhodanese. Toxicol. in vitro 6, 597-603.
Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A., 2011. Overview of the CCP4 suite and current developments. Acta Crystallogr. Sec. D: Biol. Crystallogr. 67, 235-242.
Wong, C.E., Carson, R.A., Carr, J.P., 2002. Chemically induced virus resistance in Arabidopsis thaliana is independent of pathogenesis-related protein expression and the NPR1 gene. Mol. Plant-Microbe Interact. 15, 75-81.
Wurtele, E.S., Nikolau, B.J., Conn, E.E., 1985. Subcellular and developmental distribution of β-cyanoalanine synthase in barley leaves. Plant Physiol. 78, 285-290.
Wybouw, N., Balabanidou, V., Ballhorn, D.J., Dermauw, W., Grbic, M., Vontas, J., Van Leeuwen, T., 2012. A horizontally transferred cyanase gene in the spider mite Tetranychus urticae is involved in cyanate metabolism and is differentially expressed upon host plant change. Insect Biochem. Mol. Biol. 42, 881-889.
Wybouw, N., Dermauw, W., Tirry, L., Stevens, C., Grbic, M., Feyereisen, R., Van Leeuwen, T., 2014a. A gene horizontally transferred from bacteria protects arthropods from host plant cyanide poisoning. Elife 3, e02365.

Yamaguchi, Y., Nakamura, T., Kusano, T., Sano, H., 2000. Three Arabidopsis genes encoding proteins with differential activities for cysteine synthase and β-cyanoalanine synthase. Plant Cell Phys. 41(4):465-476.
Yi, H., Jez, J.M., 2012. Assessing functional diversity in the soybean β-substituted alanine synthase enzyme family. Phytochem. 83, 15-24.
Yi, H., Juergens, M., Jez, J.M., 2012. Structure of soybean β-cyanoalanine synthase and the molecular basis for cyanide detoxification in plants. The Plant Cell 24, 2696-2706.
Zagrobelny, M., Bak, S., Møller, B.L. 2008 Cyanogenesis in plants and arthropods. Phytochem., 69, 1457-1468.
Zocher, G., Wiesand, U., Schulz, G.E., 2007. High resolution structure and catalysis of O-acetylserine sulfhydrylase isozyme B from Escherichia coli. The FEBS J. 274(20):5382-5389.

Structural and Functional Characterization of β-Cyanoalanine Synthase from *Tetranychus urticae*

Leily Daneshian[1], Isabella Renggli[1], Ryan Hanaway[1], Lesa R. Offermann[1], Caleb R. Schlachter[1], Ricardo A. Hernandez[1], Shannon Henry[1], Rahul Prakash[1], Nicky Wybouw[2,3], Wannes Dermauw[2,3], Linda S. Shimizu[1], Thomas Van Leeuwen[2,3], Thomas Makris[1,4], Vojislava Grbic[5], Miodrag Grbic[5,6], Maksymilian Chruszcz[1]*

[1] Department of Chemistry and Biochemistry, University of South Carolina, Columbia, SC 29208, USA

[2]Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Science Park 904, 1098 XH, Amsterdam, the Netherlands

[3]Department of Plants and Crops, Ghent University, Ghent, B-9000, Belgium

[4]Department of Molecular and Structural Biochemistry, NC State University, Raleigh, NC 27607, USA

[5]Department of Biology, Western University, London, Ontario, N6A 5B7, Canada

[6]University of La Rioja, Logrono, Spain

*To whom correspondence should be addressed: Maksymilian Chruszcz, Department of Chemistry and Biochemistry, University of South Carolina, Columbia, SC 29208 USA; chruszcz@mailbox.sc.edu; Tel: (803) 777-7399, FAX: (803) 777-9521.

tetur10g01570 **Nucleotide Sequence:**

ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTAC
TTCCAATCCATGACCGAAAGCACCGTGGACCGCATTAACGGCATTACCCCGAGCGC
ACTGGACCTGATTGGCAACACCCCGCTGATCGCACTGGACCGTCTGTGGCCGGGTCC
GGGTCGTCTGCTGGCAAAATGCGAATTTCTGAACCCGACCGCCTCCCTGAAGGACCG
TAGCTCTTATTACATGATTGCAAAAGCTAAGGAAAGCGGTCAGCTGAAAGATGGCG
AATCTGTCATTGAAGTGACCAGTGGTAACCAAGGCGGTGGCATCGCATGTGTTACGG
CTGTCATGGGTCATCCGTTCACCGTTACGATGTCGAAAGGCAATAGCCCGCAGCGTG
CAATTATGATGAACGCGCTGGGCGCCAATGTGATCCTGGTTGATCAAGTCACCGGCA
AACCGGGTAACGTGACGGCTGATGACGTTGCGGCCGCAGAAGAAACCGCAATGAAG
ATCCGCGAAGAAACGAACGCTTATTACGTGGACCAGTTTAACAATCCGACCAATTGC
CTGGCGCATTATGAAACCACGGGTCCGGAAATTTGGCGTCAAACCAATGGCCGCAT
CGATGCCTTTCTGGTTGGTTGCGGCACCGGTGGCTGTTTTGTCGGCACGTCGAAATTC
CTGAAAGAAAAGAACCCGAATGTTCGTTGCTTCGTGGTTGAACCGGAAGGTTGCCA
GCCGATTGCAGGCTGTACCATCACGAAACCGCTGCACCTGCTGCAAGGTAGTGGCT
ATGGTTGTGTGCCGACCCTGTTCGATAAAAAGGTCTACAACGACAGTATTTCCGTGT
CAGATGAAGAAGCCATCGAATACCGCAAACTGCTGGGCCAGAAGGAAGGTCTGTTT
TGTGGCTTCACCACGGGTGGCAATATTGCTGCGGCCATCAAACTGCTGAAGTCTGGC
CAGCTGCCGAAAGACGCCTGGGTTGTGACCATCCTGTGTGACTCTGGCCTGAAGTAT
CCGGAATAA

Table S1. Primers used for TuCAS cloning.

Primer Name	Primer Sequence
K52A_F	5’ CGCCTCCCTGGCGGACCGTAGC 3’
K52A_R	5’ GTCGGGTTCAGAAATTCGC 3’
K207A, E208A, K209A_F	5’ AGCGAACCCGAATGTTCGTTGC 3’
K207A, E208A, K209A_R	5’ GCTGCCAGGAATTTCGACGTGCC 3’
E151A, E152A_F	5’ GAAGATCCGCGCAGCAACGAACGCTT 3’
E151A, E152A_R	5’ ATTGCGGTTTCTTCTGCG 3’
Q275A, K276A, E277A_F	5’GGCAGGTCTGTTTTGTGGCTTCACC 3’
Q275A, K276A, E277A_R	5’GCCGCGCCCAGCAGTTTGCGGTA 3’

Table S2. Data collection, refinement, and model validation statistics. Values for the highest resolution shells are shown in parenthesis.

PDB accession code	6PMU	6XO2	7MFJ
Data collection			
Diffraction source	APS, 22ID	APS, 22ID	APS, 22ID
Wavelength (Å)	1.000	1.000	1.000
Space group	$P4_32_12$	$P4_12_12$	$P4_32_12$
a, c (Å)	91.4, 137.5	64.4, 143.5	91.8, 140.1
Resolution range (Å)	40.0-2.10 (2.14-2.10)	40.0-1.60 (1.63-1.60)	40.0-2.35 (2.39-2.35)
No. of unique reflections	30969 (1700)	40447 (1971)	25330 (1249)
Completeness (%)	88.6 (99.0)	99.4 (99.4)	99.1 (100.0)
Redundancy	4.6 (5.4)	6.3 (4.8)	11.5 (9.8)
$<I/\sigma(I)>$	39.0 (3.7)	33.5 (2.1)	17.5 (5.0)
R_{meas}	0.111 (0.476)	0.087 (0.763)	0.127 (0.472)
$R_{p.i.m}$	0.051 (0.202)	0.034 (0.335)	0.037 (0.147)
CC1/2	(0.880)	(0.757)	0.941
Refinement			
Resolution range (Å)	40.00-2.1 (2.15-2.10)	38.47-1.60 (1.64-1.60)	39.40-2.35 (2.41-2.35)
Completeness (%)	88.6 (99.2)	99.3 (99.0)	96.60 (93.3)
No. of reflections, working set	29271 (2381)	38452 (2983)	24758 (1646)
No. of reflections, test set	1488 (115)	1918 (112)	1245 (92)
Final R_{cryst}	0.180 (0.183)	0.164 (0.234)	0.170 (0.260)
Final R_{free}	0.223 (0.219)	0.189 (0.256)	0.223 (0.301)
Rmsd Bonds (Å)	0.008	0.015	0.002
Rmsd Angles (°)	1.5	1.9	1.4
Ramachandran Plot			
Allowed (%)	100	100	100
Favored (%)	97	98	97

Table S3. Summary of enzymatic activity parameters for TuCAS catalyzed reactions.

	Cysteine		**KCN**		**O-acetyl-Serine**		**Na_2S**	
	K_M (mM)	k_{cat} (min^{-1})	K_M (mM)	k_{cat} (min^{-1})	K_M (mM)	k_{cat} (min^{-1})	K_M (mM)	k_{cat} (min^{-1})
TuCAS	1.60±0.20	31.7	2.30±0.20	31.0	4.30±0.30	3.60	4.5±0.4	5.1
PrBSAS1[1]	0.42±0.04	449.4	7.7±1.5	615				
PrBSAS2[1]	0.62±0.01	101.4	0.27±0.08	25.8				
CysC1-a[2]	2.54		0.06		39.88		8.24	
CYSL-2[3]	1.2±0.20	4140	0.8±1.0	4140	1.9±0.3	768	1.4±1.0	768
Arabidopsis thaliana CAS[4]	0.14	159.6	0.02	0.13	8.03	120.0	0.04	90
Glycine max CAS[5]	0.81±0.25	2334.0			8.87±1.31	109.2		
Glycine max CYS[5]	0.30±0.01	12.6			3.60±0.40	3450		
TuCAS[6]	0.31	128.1*			3.17	38.76*		
BmorCYS[7]			2.66		13.04		5.88	
SfruCYS1[7]			0.54		4.74		13.04	
SfruCYS2[7]			1.87		4.81		13.74	

[1] (van Ohlen et al., 2016)

[2] (Hatzfeld et al., 2000

[3] (Vozdek et al., 2013)

[4] (Yamaguchi et al., 2000)

[5] (Yi et al., 2012)

[6] (Wybouw et al., 2014)

[7] (Li et al., 2021)

*data as V_{max} in μmol. $min^{-1}mg^{-1}$

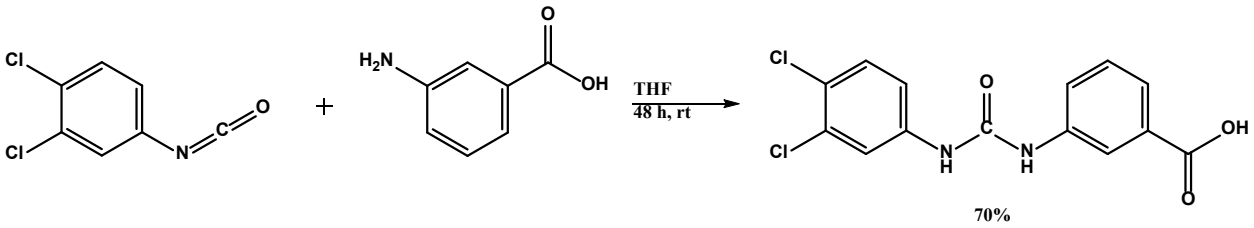


Figure S1. Synthesis of 3-(3-(3,4-dichlorophenyl)ureido)benzoic acid (Compound 1). 3-amino benzoic acid (100 mg, 0.73 mmol) was stirred in dry THF (10 mL) in an oven-dried Schlenk tube under an N_2 atmosphere and cooled to 0 °C. A solution of the 3,4-dichlorophenyl isocyanate (137 mg, 0.73 mmol) in THF (2 mL) was added dropwise. The reaction was gradually warmed to room temperature and stirred for 48 h. Then the solvent was reduced in vacuo, and 1N HCl (20 mL) was added. The precipitate that formed was collected by filtration and washed with cold Et_2O (20 mL) to afford the product as a white solid (166 mg, 70%). ^{1}H NMR (300 MHz, DMSO) δ 12.95 (s, 1H), 9.04 (s, 2H), 8.13 (s, 1H), 7.89 (d, J = 2.4 Hz, 1H), 7.63 (d, J = 8.6 Hz, 1H), 7.57 (d, J = 7.8 Hz, 1H), 7.53 (d, J = 8.8 Hz, 1H), 7.41 (t, J = 7.8 Hz, 1H), 7.34 (dd, J = 8.8, 2.4 Hz, 1H). ESI-MS *m/z* expected 325.01, found 325.0142 $(M+H)^+$.

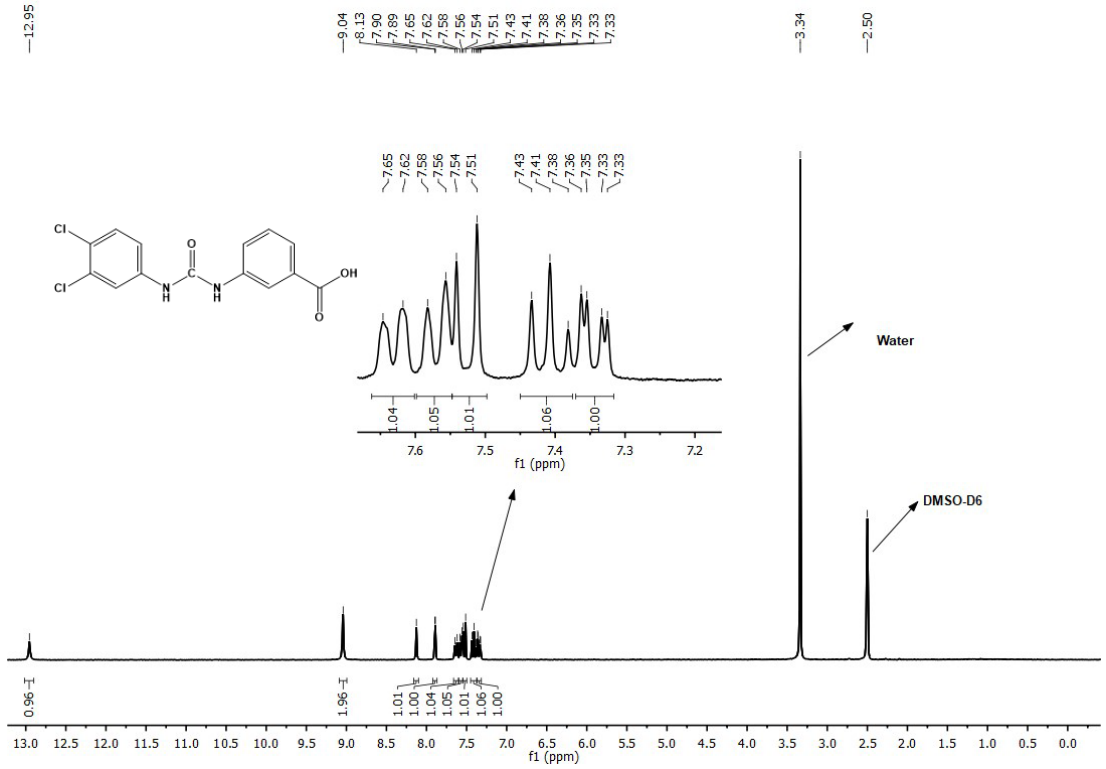


Figure S2. ^{1}H NMR (300 MHz, DMSO) of 3-(3-(3,4-dichlorophenyl)ureido)benzoic acid.

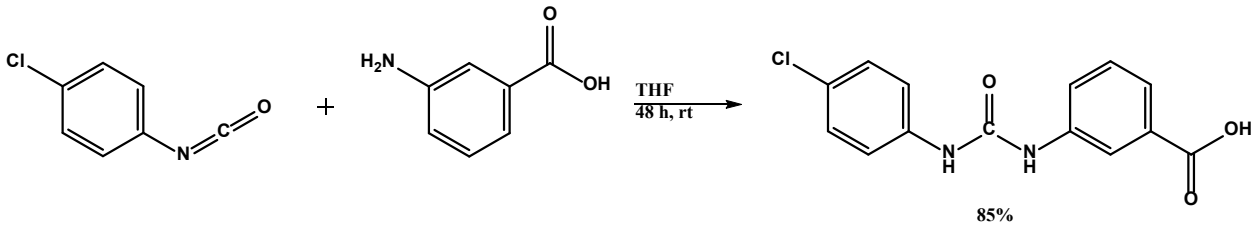


Figure S3. Synthesis of 3-(3-(4-chlorophenyl)ureido)benzoic acid (Compound 2). 3-amino benzoic acid (150 mg, 1.09 mmol) was stirred in dry THF (10 mL) in an oven-dried Schlenk tube under an N_2 atmosphere and cooled to 0 °C. A solution of the 1-chloro-4-isocyanatobenzene (168 mg, 1.09 mmol) in THF (2 mL) was added dropwise, and the solution gradually warmed to room temperature and stirred for 48 h. The white precipitate was collected by filtration and washed with cold Et_2O (20 mL) to afford the product as a white solid (270 mg, 85%). 1H NMR (300 MHz, DMSO) δ 12.93 (s, 1H), 8.94 (s, 1H), 8.85 (s, 1H), 8.12 (s, 1H), 7.63 (d, J = 8.0 Hz, 1H), 7.55 (d, J = 7.8 Hz, 1H), 7.50 (d, J = 8.9 Hz, 2H), 7.40 (t, J = 7.9 Hz, 1H), 7.33 (d, J = 8.9 Hz, 2H). ESI-MS expected *m/z* 291.05, found *m/z* 291.0529 $(M+H)^+$.

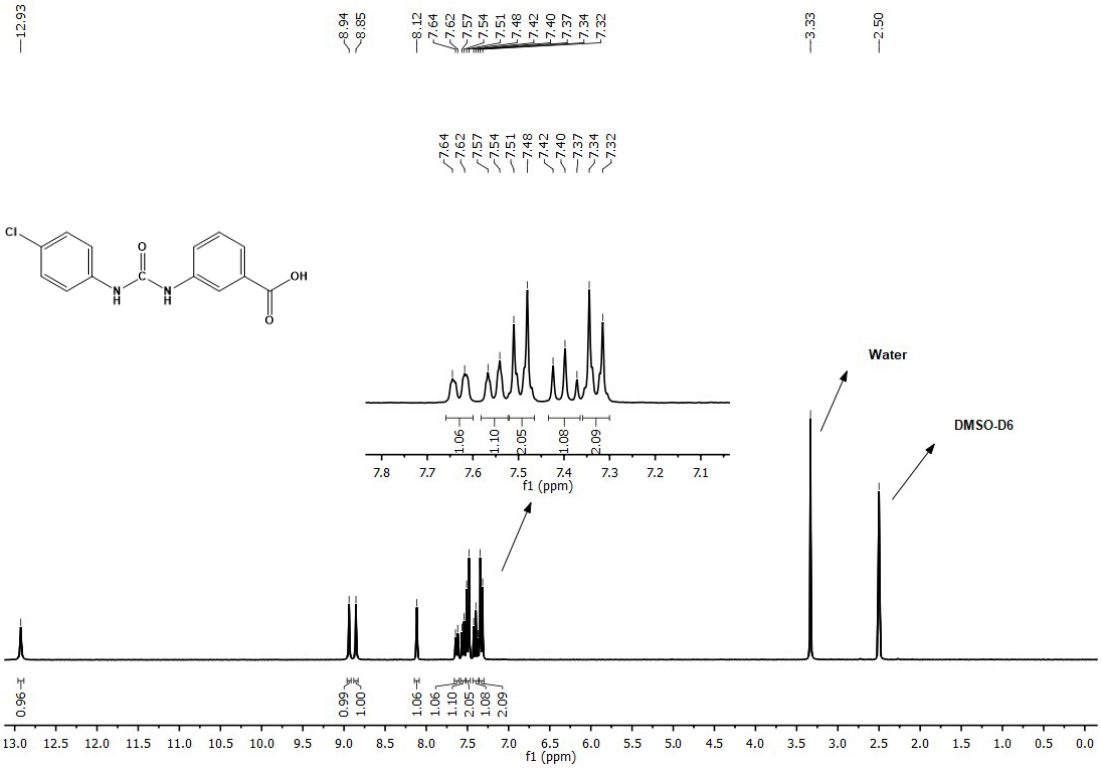


Figure S4. ^{1}H NMR (300 MHz, DMSO) of 3-(3-(4-chlorophenyl)ureido)benzoic acid.

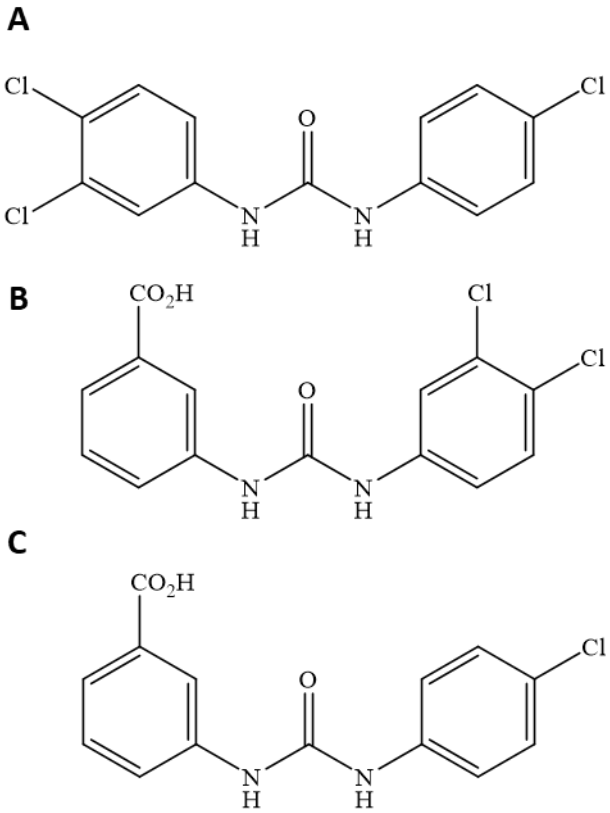


Figure S5. A) Triclocarban, B) 3-(3-(3,4-dichlorophenyl)ureido)benzoic acid (Compound 1), and C) 3-(3-(4-chlorophenyl)ureido)benzoic acid (Compound 2). The three compounds above were tested as inhibitors for TuCAS. Only compound 1 showed mixed inhibition for TuCAS.

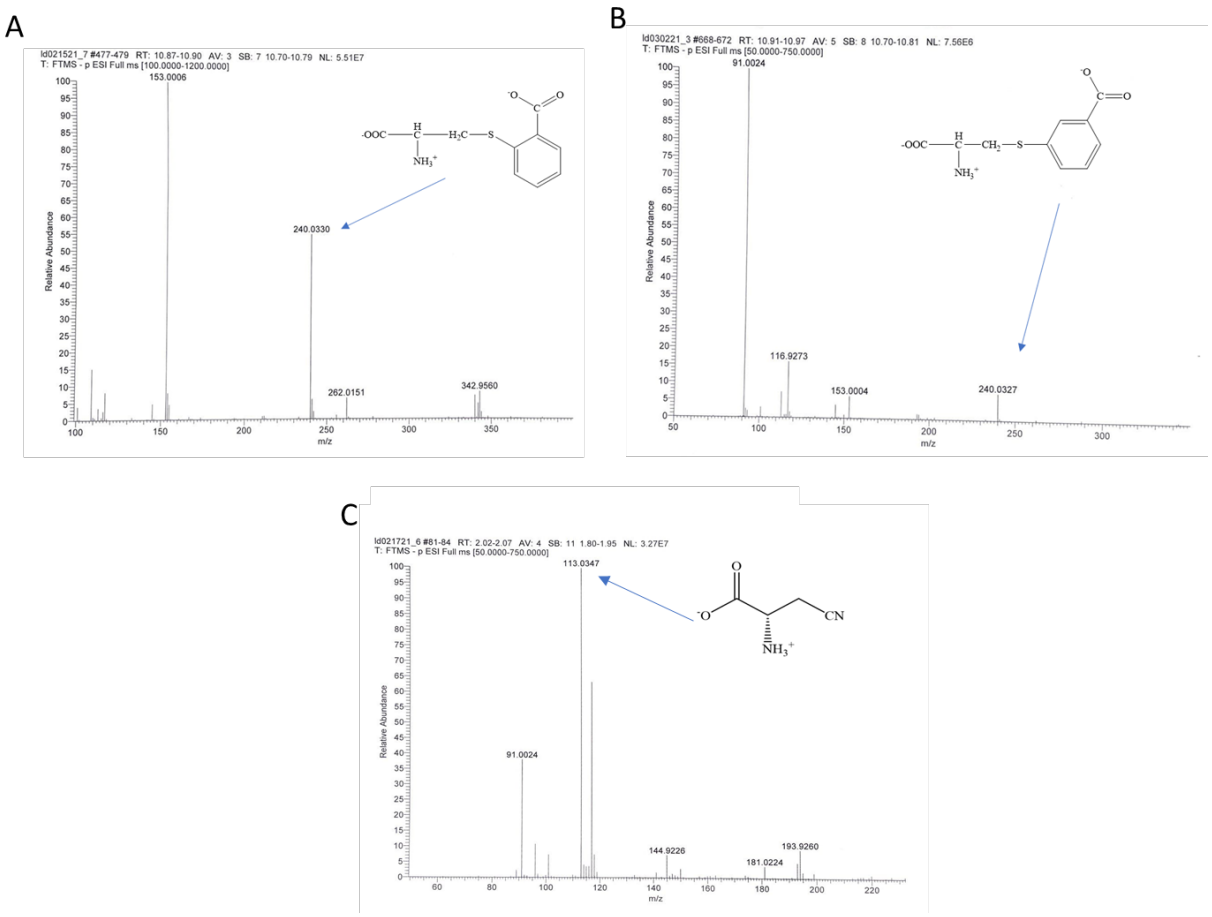


Figure S6. Mass spectroscopy results. A) 10 mM OAS and 10 mM thiosalicylic acid were used as the substrates with the product at 240 Da. B) 10 mM OAS and 10 mM 3-mercaptobenzoic acid were used as the substrate. The product which is a cysteine derivative 240 Da was observed. C) 10 mM OAS and 5 mM KCN were used as substrates. Formation of β-cyanoalanine synthase at 113 Da was detected.

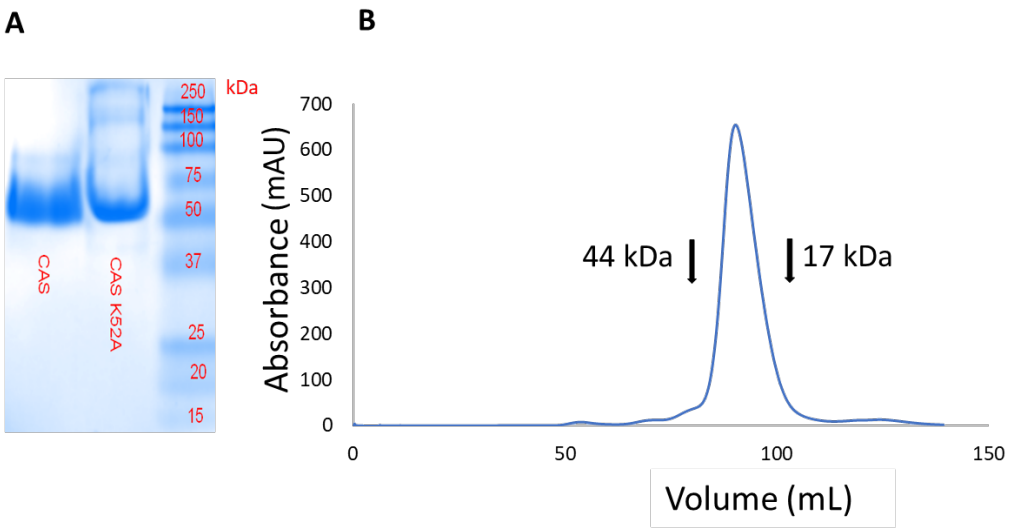


Figure S7. Results of native electrophoresis (A) and size exclusion chromatography (SEC) (B) for wild type and mutants of TuCAS. The protein runs as a monomer in SEC (arrows indicate position of peaks of two different proteins used to calibrate the column), however the native gel shows that TuCAS forms a dimer in the native state. The size exclusion chromatography results for TuCAS mutant were identical to the wild type.

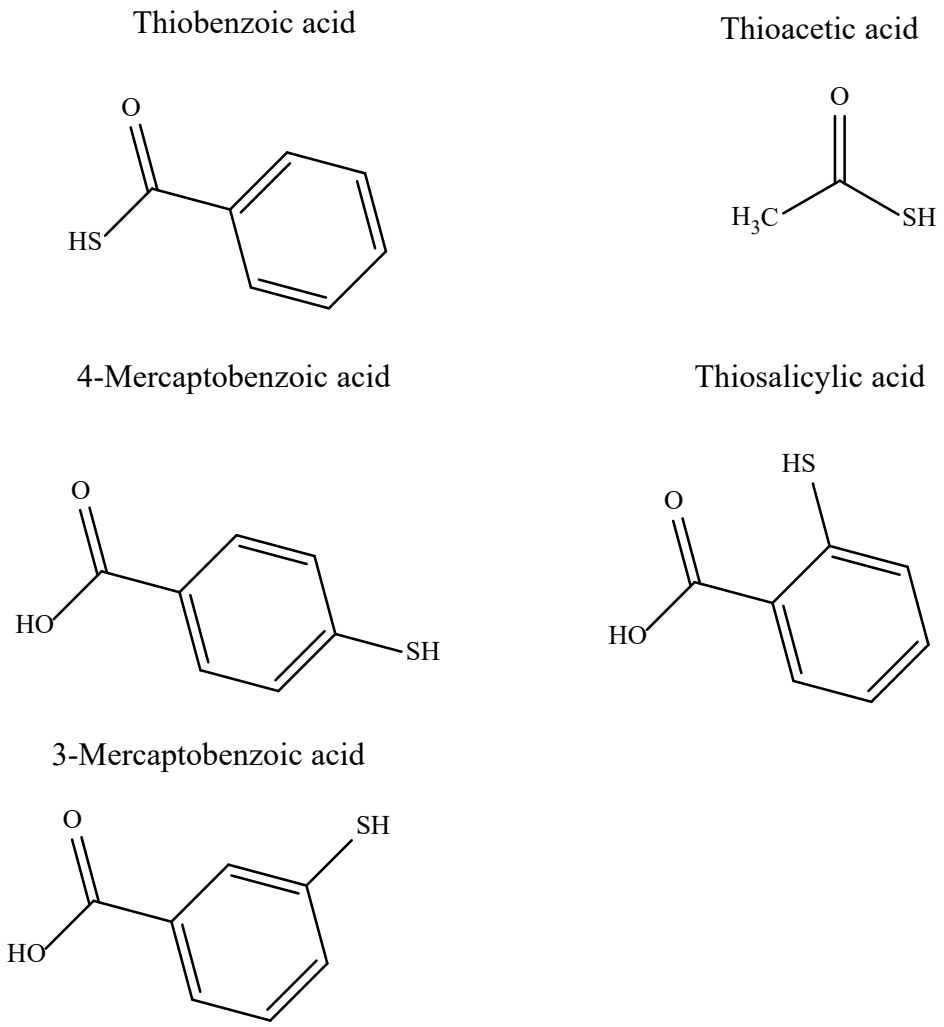
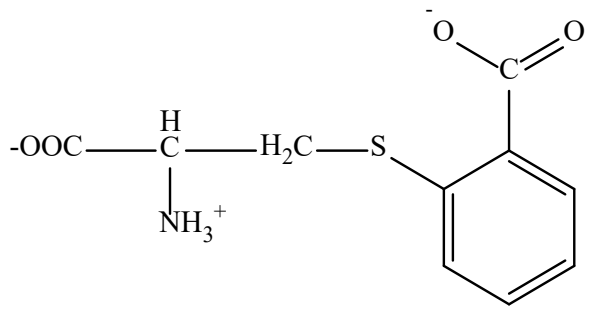


Figure S8. Tested substrates in substitution of Na_2S. Enzymatic assays were performed using the TuCAS incubated with OAS and substrates above. Products of the reactions were investigated using mass spectroscopy and Ninhydrin assay. Thiosalicylic acid and 3 mercaptobenzoic acid are confirmed to be substrates for TuCAS.

A



B

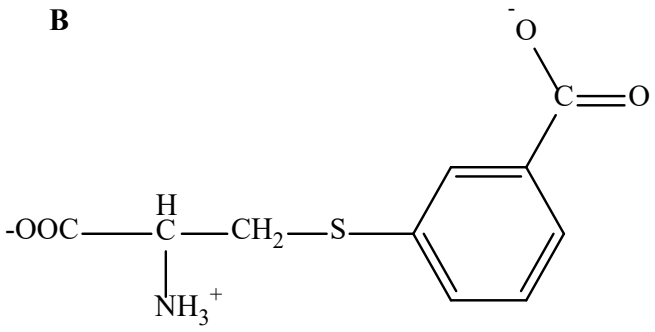


Figure S9. Products for tested substrates in substitution of Na_2S.Enzymatic assays were performed using the TuCAS incubated with OAS and substrates in Figure S10. Products of the reactions were investigated using mass spectroscopy. A) Thiosalicylic acid was used as the second substrate. B) 3-mercaptobenzoic acid was used as the second substrate.

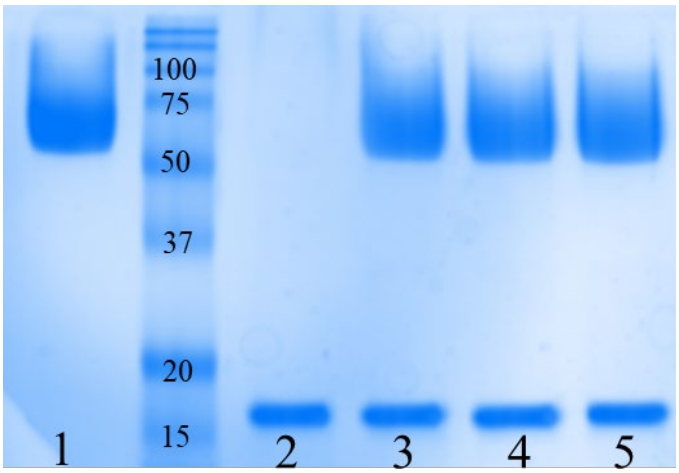
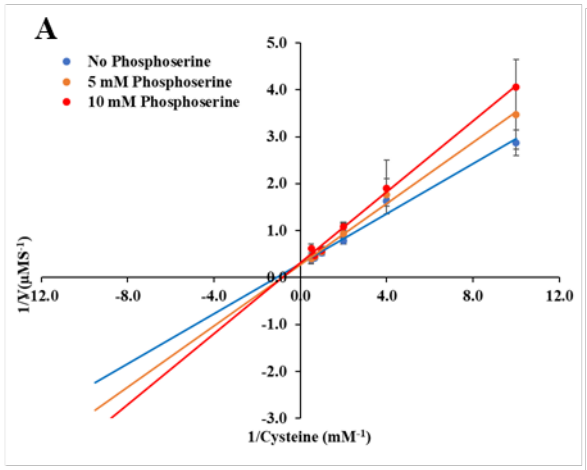
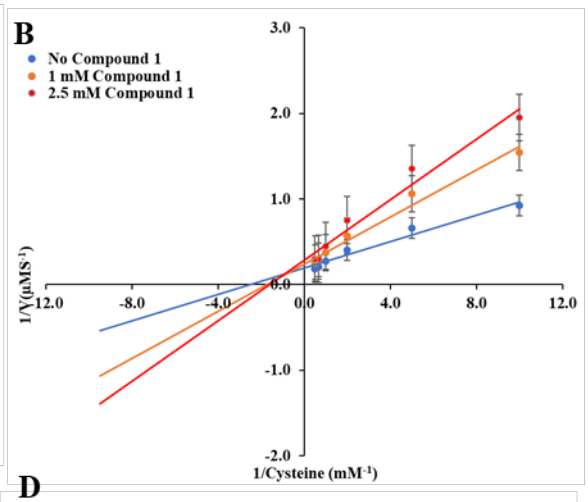
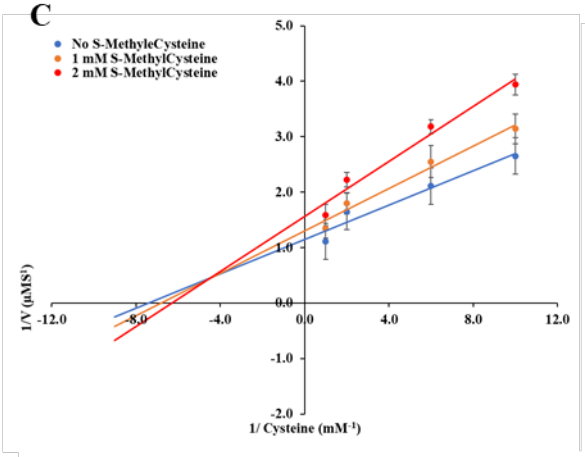
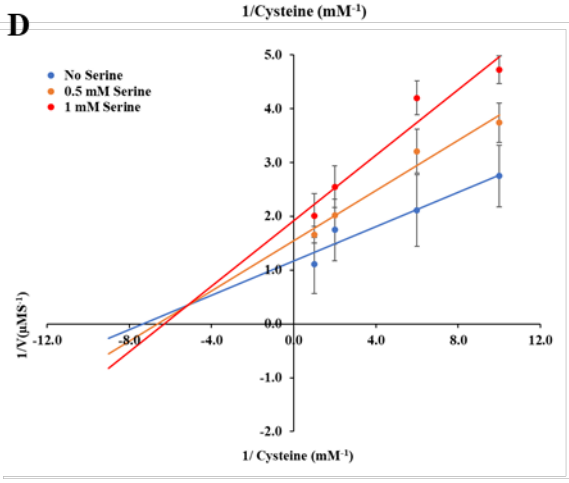


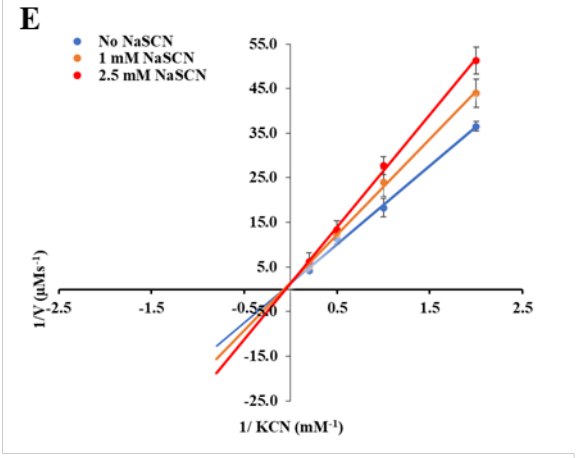
Figure S10. Native gel electrophoresis results for TuCAS and Tetur36g00900. (1) TuCAS alone; (2) Tetur36g00900 alone; (3) mixture of TuCAS and Tetur36g00900; (4) mixture of TuCAS, Tetur36g00900, and OAS; (5) mixture of TuCAS, Tetur36g00900, and cysteine. All the mixtures were incubated with PLP for 16 hours at 4°C. Native gel does not show any interactions between the proteins.

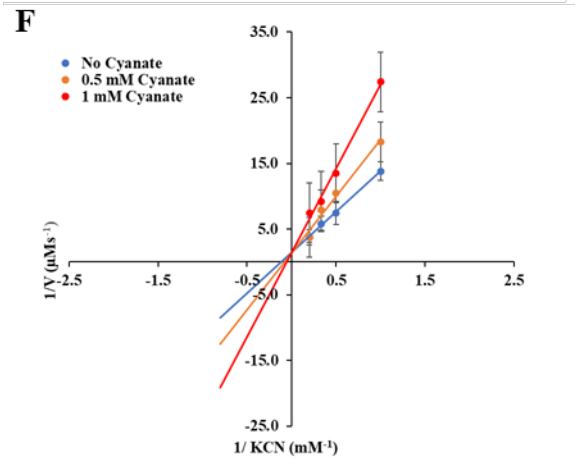

A
No Phosphoserine
5 mM Phosphoserine
10 mM Phosphoserine
1/V(μMS⁻¹)
1/Cysteine (mM⁻¹)


B
No Compound 1
1 mM Compound 1
2.5 mM Compound 1
1/V(μMS⁻¹)
1/Cysteine (mM⁻¹)


C
No S-MethyleCysteine
1 mM S-MethylCysteine
2 mM S-MethylCysteine
1/V (μMS⁻¹)
1/ Cysteine (mM⁻¹)


D
No Serine
0.5 mM Serine
1 mM Serine
1/V(μMS⁻¹)
1/ Cysteine (mM⁻¹)


E
No NaSCN
1 mM NaSCN
2.5 mM NaSCN
1/V (μMs⁻¹)
1/ KCN (mM⁻¹)


F
No Cyanate
0.5 mM Cyanate
1 mM Cyanate
1/V (μMs⁻¹)
1/ KCN (mM⁻¹)

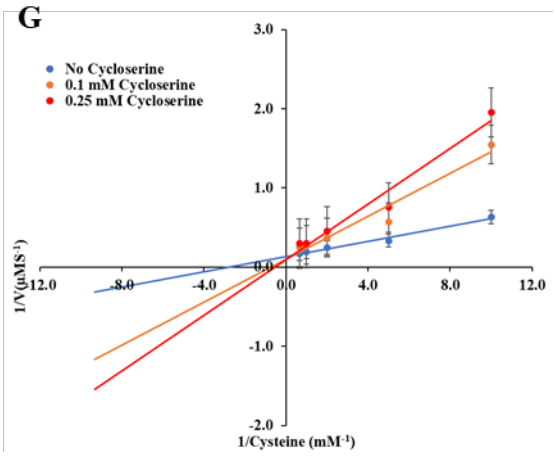


Figure S11. Inhibitors of β-cyanoalanine synthesis. A) O-phospho-L-serine (competitive inhibitor of L-cysteine, B) Compound 1 (mixed-inhibition of L-cysteine), C) S-methyl-L-cysteine (competitive inhibitor of L-cysteine), D) L-serine (mixed non-competitive inhibitor of L-cysteine), E) Sodium thiocyanate (competitive inhibitor of potassium cyanide). F) Cyanate (competitive inhibitor of potassium cyanide). G) L-cycloserine (mixed-inhibition for L-cysteine)

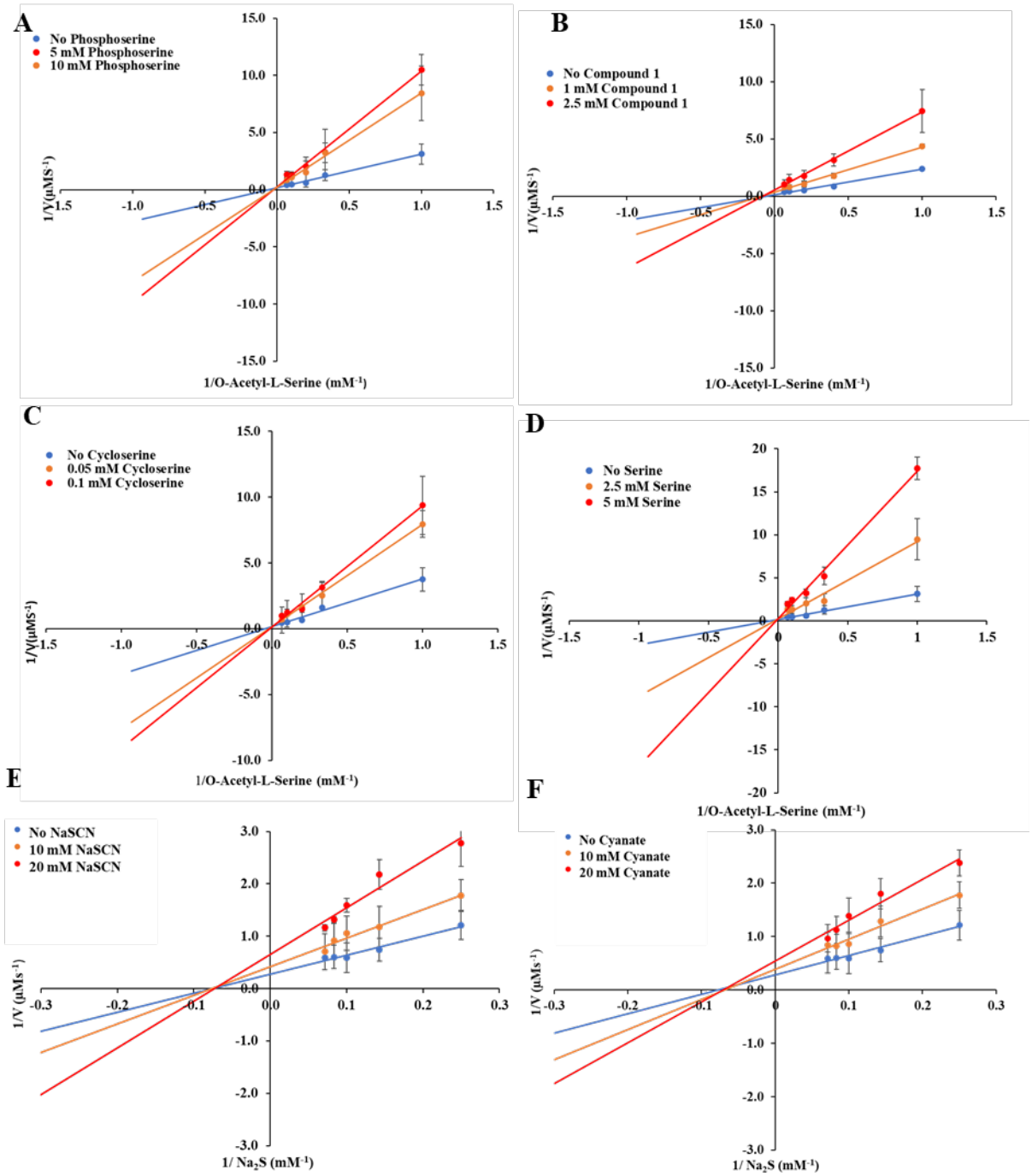


Figure S12. Inhibitors of cysteine synthesis. A) O-phospho-L-serine (competitive inhibitor of O-acetyl-L-serine), B) Compound 1 (mixed-inhibition of O-acetyl-L-serine). C) L-cycloserine (competitive inhibitor of O-acetyl-L-serine). D) L-serine (competitive inhibitor of O-acetyl-L-serine), E) Sodium thiocyanate (noncompetitive inhibitor of sodium sulfide). F) Sodium cyanate (noncompetitive inhibitor of sodium sulfide).