Structure of Soybean β-Cyanoalanine Synthase and the Molecular Basis for Cyanide Detoxification in Plants

Hankui Yi, Matthew Juergens, and Joseph M. Jez
Department of Biology, Washington University, St. Louis, Missouri 63130

Plants produce cyanide (CN⁻) during ethylene biosynthesis in the mitochondria and require β-cyanoalanine synthase (CAS) for CN⁻ detoxification. Recent studies show that CAS is a member of the β-substituted alanine synthase (BSAS) family, which also includes the Cys biosynthesis enzyme O-acetylserine sulfhydrylase (OASS), but how the BSAS evolved distinct metabolic functions is not understood. Here we show that soybean (Glycine max) CAS and OASS form α-aminoacylare reaction intermediates from Cys and O-acetylserine, respectively. To understand the molecular evolution of CAS and OASS in the BSAS enzyme family, the crystal structures of Gm-CAS and the Gm-CAS K95A mutant with a linked pyridoxal phosphate (PLP)-Cys molecule in the active site were determined. These structures establish a common fold for the plant BSAS family and reveal a substrate-induced conformational change that encloses the active site for catalysis. Comparison of CAS and OASS identified residues that covary in the PLP binding site. The Gm-OASS T81M, S181M, and T185S mutants altered the ratio of OASS:CAS activity but did not convert substrate preference to that of a CAS. Generation of a triple mutant Gm-OASS successfully switched reaction chemistry to that of a CAS. This study provides new molecular insight into the evolution of diverse enzyme functions across the BSAS family in plants.

INTRODUCTION

In plants, cyanide (CN⁻) is a naturally occurring molecule with important metabolic roles connected to plant protection. The toxicity of cyanide, through its inhibition of cytochrome c oxidase in the mitochondrial electron transport chain (Palmer, 1993), makes it a potent chemical deterrent. A variety of cyanogenic glucosides are present in many plant species, providing a chemical reservoir to release HCN in response to herbivores and pathogens that cause tissue damage. In addition, these same molecules are useful for nitrogen transport within the plant and nitrogen storage for primary metabolism (Moller, 2010). Biosynthetically, the production of cyanide in plants is directly coupled to ethylene synthesis (Bleecker and Kende, 2000). Conversion of 1-amino-cyclopropane-1-carboxylic acid to ethylene by 1-amino-cyclopropane-1-carboxylic acid oxidase results in the release of cyanoformic acid, which spontaneously decarboxylates to release CN⁻ (Peiser et al., 1984). The generation of CN⁻ in ethylene synthesis necessitates its detoxification and is linked to activation of β-cyanoalanine synthase (CAS; Blumenthal et al., 1968; Miller and Conn, 1980; Yip and Yang, 1988; Goudey et al., 1989). CAS catalyzes the addition of CN⁻ to Cys to yield β-cyanoalanine (Figure 1A), which can then be used for recycling of reduced nitrogen through metabolism to Asn (Castric et al., 1972). The activity of this enzyme is widespread in plant tissues (Wurtele et al., 1984, 1985) and localized to mitochondria (Wurtele et al., 1985; Hatzfeld et al., 2000; Maruyama et al., 2000; Warrilow and Hawkesford, 2000, 2002; Lai et al., 2009).

Thus, CAS activity is induced and localized to effectively protect the electron transport chain at the site of ethylene production (Bleecker and Kende, 2000). Recent work by García et al. (2010) shows that CAS is essential for root hair development in Arabidopsis thaliana and suggests that its activity helps prevent the accumulation of HCN, which represses expression of genes encoding enzymes involved in cell wall maintenance and formation of the root hair tip.

Biochemically, CAS activity has been described in a variety of plant species, including blue lupine (Lupinus angustifolius) (Akopyan et al., 1975); spinach (Spinacia oleracea) (Ikegami et al., 1988b), pea (Lathyrus latifolius) (Ikegami et al., 1988a), cassava (Manihot esculenta) (Elias et al., 1997), cocklebur (Xanthium pennsylvanicum) (Maruyama et al., 1998), potato (Solanum tuberosum) (Maruyama et al., 2000, 2001), apple (Malus domestica) (Yip and Yang, 1988), Arabidopsis (Warrilow and Hawkesford, 2000; Yamaguchi et al., 2000), and rice (Oryza sativa) (Lai et al., 2009). These same studies also identify CAS as a 50- to 70-kD dimeric pyridoxal phosphate (PLP)-dependent enzyme. Molecular cloning of CAS from Arabidopsis, spinach, and soybean (Glycine max) has established this enzyme as a member of the β-substituted alanine synthase (BSAS) family (Hatzfeld et al., 2000; Yamaguchi et al., 2000; Watanabe et al., 2008; Yi et al., 2010a, 2010b). The plant BSAS family also includes O-acetylserine sulfhydrylase (OASS), which catalyzes the synthesis of Cys from O-acetylserine and sulfide using a similar overall reaction as CAS (Figure 1A) (Bonnet et al., 2005; Ravilious and Jez, 2012). Structural and functional studies of OASS indicate that the enzyme uses a two-step reaction mechanism in which O-acetylserine reacts with the PLP cofactor to yield a reactive α-aminoacylare intermediate in the first half step. Binding of sulfide in the second step leads to reaction with the...
Table 1. Kinetic Comparison of Gm-CAS and Gm-OASS Substrate Specificities

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>CAS Reaction Substrates</th>
<th>OASS Reaction Substrates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>k&lt;sub&gt;cat&lt;/sub&gt; (s&lt;sup&gt;-1&lt;/sup&gt;)</td>
<td>K&lt;sub&gt;m&lt;/sub&gt; (mM)</td>
</tr>
<tr>
<td>Gm-CAS</td>
<td>38.9 ± 4.5</td>
<td>0.81 ± 0.25</td>
</tr>
<tr>
<td>Gm-OASS</td>
<td>0.21 ± 0.09</td>
<td>0.30 ± 0.01</td>
</tr>
<tr>
<td>KCN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gm-CAS</td>
<td>39.2 ± 0.03</td>
<td>0.26 ± 0.03</td>
</tr>
<tr>
<td>Gm-OASS</td>
<td>0.23 ± 0.03</td>
<td>0.08 ± 0.02</td>
</tr>
</tbody>
</table>

All values are expressed as a mean ± se (n = 3).
To determine whether Gm-CAS and Gm-OASS share common reaction mechanisms involving an α-aminoacrylate intermediate, each protein was incubated with Cys, O-acetylserine, and β-cyanoalanine and then analyzed by UV/visible absorption spectroscopy (Figure 2). Addition of Cys or O-acetylserine to Gm-CAS shifted the maximum absorbance (A<sub>max</sub>) of PLP from 412 nm to 470 nm (Figure 2A), which is consistent with formation of an α-aminoacrylate intermediate from either molecule (Bonner et al., 2005). As reported with Arabidopsis OASS (Bonner et al., 2005), signal for the α-aminoacrylate intermediate (A<sub>max</sub> = 470 nm) was observed after addition of O-acetylserine to Gm-OASS (Figure 2B). Incubation of Gm-CAS with β-cyanoalanine and Gm-OASS with either Cys or β-cyanoalanine yielded a shift to the external aldime (A<sub>max</sub> = 418 nm), which is nonreactive. These experiments demonstrate that both CAS and OASS stably form a common α-aminoacrylate intermediate in their respective reactions (Figure 2C). To better understand the molecular function of CAS, we determined the x-ray crystal structure of the soybean enzyme.

### Table 2. Crystallographic Statistics

<table>
<thead>
<tr>
<th>Crystal</th>
<th>Gm-CAS</th>
<th>Gm-CAS-K95A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>C2</td>
<td>P22,2&lt;sub&gt;1&lt;/sub&gt;</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td>a = 190.32 Å, b = 154.58 Å, c = 70.80 Å; β = 101.9°</td>
<td>a = 71.09 Å, b = 154.33 Å, c = 198.33 Å</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.979</td>
<td>0.979</td>
</tr>
<tr>
<td>Resolution range (Å) (highest shell resolution)</td>
<td>49.2 to 2.50 (2.54 to 2.50)</td>
<td>47.2 to 1.77 (1.80 to 1.77)</td>
</tr>
<tr>
<td>Reflections (total/unique)</td>
<td>327,779/68,674</td>
<td>805,528/201,382</td>
</tr>
<tr>
<td>Completeness (highest shell)</td>
<td>99.8% (100.0%)</td>
<td>96.4% (98.2%)</td>
</tr>
<tr>
<td>&lt;I&gt;/&lt;s&gt; (highest shell)</td>
<td>29.8 (4.1)</td>
<td>25.6 (1.9)</td>
</tr>
<tr>
<td>R&lt;sub&gt;sym&lt;/sub&gt; (highest shell)</td>
<td>8.5% (57.7%)</td>
<td>10.8% (62.6%)</td>
</tr>
<tr>
<td>Model and refinement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R&lt;sub&gt;cryst&lt;/sub&gt; / R&lt;sub&gt;free&lt;/sub&gt;</td>
<td>16.7%/21.0%</td>
<td>17.0%/20.2%</td>
</tr>
<tr>
<td>No. of protein atoms</td>
<td>9,699</td>
<td>14,670</td>
</tr>
<tr>
<td>No. of water molecules</td>
<td>233</td>
<td>1,731</td>
</tr>
<tr>
<td>No. of ligand atoms</td>
<td>60</td>
<td>132</td>
</tr>
<tr>
<td>Root mean square deviation, bond lengths (Å)</td>
<td>0.008</td>
<td>0.006</td>
</tr>
<tr>
<td>Root mean square deviation, bond angles (°)</td>
<td>1.14</td>
<td>1.05</td>
</tr>
<tr>
<td>Average B-factor (Å²)</td>
<td>51.1, 46.7, 41.1</td>
<td>20.5, 33.9, 17.1</td>
</tr>
<tr>
<td>Stereochemistry: most favored, allowed</td>
<td>96.4, 3.6, 0.0%</td>
<td>97.4, 2.6, 0.0%</td>
</tr>
<tr>
<td>generously allowed</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>R<sub>sym</sub> = Σ|I<sub>o</sub> - <I>/ΣI<sub>o</sub>, where <I> is the average intensity over symmetry.

<sup>b</sup>R<sub>cryst</sub> = Σ|F<sub>o</sub> - <F>/ΣF<sub>o</sub>, where summation is more than the data used for refinement.

<sup>c</sup>R<sub>free</sub> is defined the same as R<sub>cryst</sub> but was calculated using 5% of data excluded from refinement.
Three-Dimensional Structure of Gm-CAS

The x-ray crystal structure of Gm-CAS was determined by molecular replacement, with four molecules forming two dimers in the asymmetric unit (Table 2). Each monomer consists of two mixed α/β domains with a PLP cofactor at the active site separating the two halves (Figure 3A). The N-terminal domain (residues 94 to 197) is formed by a central β-sheet (β3 to β5) surrounded by four α-helices (α1 to α4). In between the two domains, α5 (residues 195 to 2019) spans both halves of the structure. The C-terminal domain consists of a short β-sheet (β1 to β2; residues 51 to 93) and a larger β-sheet (β6 to β8) surrounded by three α-helices (α6 to α8). An additional α-helix (α9) forms part of the dimer interface by packing against α2 and α3 of the adjoining monomer. The structure of Gm-CAS shares the same fold as Arabidopsis OASS (Bonner et al., 2005) with a 0.9 Å² root mean square deviation for the aligned 316 Cα atoms. This common scaffold is also found in other PLP-dependent enzymes, including cystathionine β-synthase, Thr dehydratase, and Trp synthase (Meier et al., 2001; Simanshu et al., 2006; Ngo et al., 2007).

The location of PLP in the structure identifies the active site. Clear electron density shows that the cofactor forms a Schiff base linkage with Lys-95, which is part of α1 (Figure 3B). Multiple interactions are formed to anchor PLP in the active site (Figure 3C; see Supplemental Figure 1A online). Asn-126 and Ser-318 form hydrogen bonds with the oxygen and nitrogen atoms, respectively, in the pyridine ring of PLP. Extensive hydrogen bond interactions with the phosphate moiety are made by the backbone nitrogens of Gly-230, Ser-231, and Gly-232 and the side chain hydroxyl groups of Ser-231 and Thr-234. Two additional contacts with the phosphate group are mediated by water molecules that hydrogen bond to the backbone nitrogen of Val-235 and the side chains of Asn-202 and His-206. Across the plant BSAS enzyme family, the PLP-attachment site corresponding to Lys-95 is invariant (Hatzfeld et al., 2000; Yamaguchi et al., 1996).

Figure 3. Crystal Structure of Gm-CAS.

(A) Ribbon diagram showing the overall structure of the Gm-CAS dimer. The left-hand monomer shows secondary structure features with α-helices and β-strands colored pink and blue, respectively. The location of Lys-95 and PLP are shown as stick models (gold).

(B) Electron density of the PLP-Lys-95 Schiff base. A 2Fo-Fc omit map (1.5 σ) is shown.

(C) The Gm-CAS active site. Residues interacting with PLP are shown as stick models. Two water molecules are shown as red spheres marked with a “W”.

(D) Sequence comparison of soybean CAS (GmCAS), Arabidopsis CAS (AtCAS); soybean OASS (GmOASS), and Arabidopsis OASS (AtOASS). The mitochondrial localization sequence of Gm-CAS spans the first 50 amino acids. Conserved residues are highlighted in orange. Residues in white indicate a variation from the highly conserved sequence. The Schiff base Lys is highlighted in red. Residues in the PLP and amino acid binding sites are highlighted in green and blue, respectively.
et al., 2000; Watanabe et al., 2008; Yi et al., 2010a, 2010b). In addition to this critical Lys, the residues forming the PLP binding site are highly conserved in the CAS and OASS from soybean and Arabidopsis (Figure 3D).

### Ligand-Induced Conformational Change in Gm-CAS

To understand recognition of amino acid substrates by Gm-CAS, the K95A mutant was generated, expressed, purified, and used for crystallization. In the structure of Arabidopsis OASS, mutation of the Schiff base Lys results in retention of PLP in the active site (Bonner et al., 2005). This same approach was used to generate a trapped substrate–intermediate complex of Gm-CAS. The 1.77-Å resolution structure of the Gm-CAS K95A mutant was solved by molecular replacement using the wild-type Gm-CAS as a search model. Within the active site of the K95A mutant, unambiguous electron density showed that PLP was trapped in an external aldime linkage with a Cys molecule, which is the natural substrate of Gm-CAS (Figure 4A).

Comparison of the two Gm-CAS structures showed that the N-terminal domain undergoes a rigid body shift in position by ~15° toward the PLP-Cys bound in the active site (Figure 4B). In addition, α9, which forms part of the dimerization interface, is prevented from packing against residues 123-126 of the adjacent monomer, also coordinate changes position. The conformational change of the N-terminal domain alters the size of the active site entrance with the structural transition from an open conformation (wild-type structure) to a closed conformation (K95A mutant) (Figure 4C). In the open form, the entrance is wide enough to allow for binding of Cys while in the first half reaction. In comparison, the active site cleft is solvent inaccessible in the closed form, although the reactive side chain of the amino acid attached to PLP is oriented toward a narrow opening at the surface (Figure 4D).

Within the K95A active site, the binding interactions between the protein and PLP are similar to those observed in the wild-type structure (see Supplemental Figures 1A and 1B online), although the pyridine ring rotates ~15° toward the protein interior (see Supplemental Figure 1C online). Extensive hydrogen bonds are formed between the Cys carboxylate group and side chain and main chain atoms of Thr-123, Ser-124, Gly-125, and Asn-126 in the substrate interaction or TSGN loop (residues 123-126) to anchor the ligand in the active site (Figure 4D; see Supplemental Figure 1B online). Additional hydrogen bonds from the backbone nitrogen of Met-127 and the side chain amide of Gln-196 contribute to Cys binding. In addition, two water molecules mediate contacts between PLP and Arg-154 and Gln-196 (see Supplemental Figure 1B online). This intricate interaction network between the substrate and the TSGN loop drives the shift in position of the N-terminal domain to enclose the active site entrance during catalysis.

### Specificity of CAS versus OASS Activity

The three-dimensional structures and reaction chemistry of CAS and OASS are conserved; however, each protein is specialized for metabolic function (i.e., CN− detoxification and Cys biosynthesis, respectively). Structural and sequence comparison of the residues around the PLP-Cys external aldime in the K95A mutant structure revealed only a few differences between CAS and OASS from a variety of species (Figures 5A and 5B). Of the residues interacting with the Cys moiety of PLP-Cys, Met-127 is conserved in CAS from different plant species but is a Thr residue in OASS. This residue is positioned at the end of the TSGN substrate interaction loop and may alter how the loop interacts with different amino acid ligands. Near the PLP phosphate group, Met-227 in β6 and Ser-231 in the β6/α6 loop of Gm-CAS also covary as a Ser and Thr, respectively, in the plant OASS proteins.

To test whether any of these three amino acid differences alter CAS versus OASS activity, site-directed mutants of Gm-OASS were generated with the goal of switching its reaction sequence to prefer Cys and function as a CAS. Spectroscopic analysis of Gm-CAS (Figure 2A) showed that the α-aminoacrylate was chemically accessible for both Cys and O-acetylsereine, whereas Gm-OASS at comparable concentrations of these substrates formed the reactive intermediate only with O-acetylsereine (Figure 2B). Because stable formation of this intermediate from Cys is required for CAS activity, we targeted Gm-OASS for mutagenesis to introduce formation of the α-aminoacrylate intermediate from Cys.
Each position in the Gm-OASS active site was mutated to the corresponding residue from Gm-CAS, and the resulting Gm-OASS T81M, S181M, and T185S mutants were expressed and purified to homogeneity. Kinetic analysis of the three Gm-OASS point mutants showed that each change decreased catalytic efficiency ($V_{\text{max}}/K_m$) with O-acetylserine by less than threefold (Figure 5C, Table 3). With Cys as a substrate, the T81M mutation did not alter substrate preference but modestly improved the spectroscopic signal associated with formation of the $\alpha$-aminoacrylate intermediate compared with Gm-OASS (Figure 2B; see Supplemental Figure 2A online). The S181M and T185S mutants showed twofold to fourfold improvements in both enzymes use the same chemistry to generate a common reaction intermediate from different substrates. Comparison of CAS and OASS also reveal how subtle active site variations lead to specialization of function between these two BSAS family members and suggest how the biochemical functions of other BSAS proteins in plants evolved.

### DISCUSSION

CAS plays a central role in the detoxification of CN$^-$ produced as a byproduct of ethylene biosynthesis (Blumenthal et al., 1968; Miller and Conn, 1980; Yip and Yang, 1988; Goudey et al., 1989), but how this enzyme performs its critical function was unclear. Although parallels between the reactions catalyzed by CAS in CN$^-$ detoxification and OASS in Cys synthesis have been suggested (Figure 1A), the structural and functional studies presented here provide the first molecular insight into the mechanism of CN$^-$ detoxification in plants and evidence that both enzymes use the same chemistry to generate a common reaction intermediate from different substrates. Comparison of CAS and OASS also reveal how subtle active site variations lead to specialization of function between these two BSAS family members and suggest how the biochemical functions of other BSAS proteins in plants evolved.

The structure of Gm-CAS (Figures 3 and 4) shares a common three-dimensional fold with other PLP-dependent enzymes, including OASS from Arabidopsis and bacteria (Burkhard et al., 1999; Meier et al., 2001; Bonner et al., 2005; Claus et al., 2005; Francois et al., 2006; Simanshu et al., 2006; Ngo et al., 2007). The conserved structure is important for maintaining the PLP binding site across the BSAS family in plants and for providing core reaction chemistry for different metabolic functions. Although Gm-CAS adopts a similar structure as the OASS from Arabidopsis (Bonner et al., 2005), kinetic analysis of Gm-CAS and Gm-OASS confirms that each enzyme prefers distinct substrates (Table 1), as observed in the BSAS from other plant
species (Hatzfeld et al., 2000; Yamaguchi et al., 2000; Watanabe et al., 2008; Yi et al., 2010a, 2010b). The ability of CAS and OASS to perform the other’s reaction, albeit with a lower catalytic efficiency, implies a common chemical mechanism. Spectroscopic analysis of Gm-CAS demonstrates formation of the \( \alpha \)-aminoacylate intermediate from either Cys or O-acetylseryne (Figure 2A). By contrast, Gm-OASS forms the reaction intermediate with O-acetylseryne but not with Cys (Figure 2B), ensuring efficient production of Cys without significant diversion to the unproductive external aldimine form of PLP. These results show that even though each enzyme can bind both ligands, there are differences in their active sites that guide the reaction chemistry along two possible paths that result in formation of either useful reactive \( \alpha \)-aminoacylate intermediate that leads to product formation or a nonreactive external aldimine (Figure 2C).

The reaction catalyzed by CAS begins with binding of Cys to the active site, which is in the open conformation (Figure 3). This positions the carboxylate group of Cys near the substrate interaction or TSGN loop to trigger the N-terminal domain movement that leads to the closed form of the enzyme (Figure 4). In the first half reaction catalyzed by CAS (Figure 6A), the \( \alpha \)-amine of the Cys reacts with the C4’ of the PLP-Lys-95 Schiff base to release the Lys residue. Formation of PLP-Cys allows Lys-95 to act as a general base in the reaction, which allows for trapping of the PLP-Cys external aldimine (Figure 4). This result presented in our study confirms the critical catalytic role for this residue in the CAS reaction mechanism. At the same time, trapping of the transient PLP intermediate bound to the natural substrate of CAS indicates a catalytic role for Lys-95 in the second half reaction. Based on the structure of Gm-CAS with PLP-Cys in the active site (Figure 4), the amino acid side chain is oriented toward the active site entrance and is ideally positioned for nucleophilic attack by \( \text{CN}^- \) in the second half reaction. Subsequent protonation of the \( \alpha \)-carbon of \( \beta \)-cyanoalanine by Lys-95 leads to product release and regeneration of the enzyme-bound Schiff base form of PLP in the second half reaction (Figure 6A), which resets the active site for the next round of catalysis.

Table 3. Kinetic Comparison of Wild-Type and Mutant Gm-OASS with Amino Acid Substrates

<table>
<thead>
<tr>
<th>Type of Gm-OASS</th>
<th>Cys</th>
<th>O-acetylseryne</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( k_{\text{cat}} ) (s(^{-1}))</td>
<td>( K_m ) (mM)</td>
</tr>
<tr>
<td>Gm-OASS</td>
<td>0.21 ( \pm ) 0.09</td>
<td>0.30 ( \pm ) 0.01</td>
</tr>
<tr>
<td>T81M</td>
<td>0.45 ( \pm ) 0.11</td>
<td>0.60 ( \pm ) 0.10</td>
</tr>
<tr>
<td>S181M</td>
<td>1.06 ( \pm ) 0.34</td>
<td>0.73 ( \pm ) 0.19</td>
</tr>
<tr>
<td>T185S</td>
<td>0.77 ( \pm ) 0.09</td>
<td>0.31 ( \pm ) 0.04</td>
</tr>
<tr>
<td>Triple</td>
<td>0.59 ( \pm ) 0.13</td>
<td>0.13 ( \pm ) 0.01</td>
</tr>
</tbody>
</table>

All values are expressed as a mean \( \pm \) se (n = 3).

Given the chemical parallels between CAS and OASS, what governs the ability of each enzyme to preferentially form the same reactive intermediate from either Cys or O-acetylseryne? Comparison of the Gm-CAS active site structure and sequence with other CAS and OASS proteins from a variety of plant species identified three amino acids near the PLP-Cys ligand that covaried between the two enzymes (Figures 5A and 5B). Single point mutations of these residues in Gm-OASS (T81M, S181M, and T185S) were introduced to test their potential effect on enhancing the CAS activity of this enzyme (Figure 5C, Table 3). Although the Gm-OASS mutants displayed altered ratios of...

Figure 6. Summary of Reactions Catalyzed in the BSAS Enzyme Family of Plants.

(A) Reaction mechanism of CAS.
(B) Overview of enzymatic activities in the BSAS family, showing first half reaction substrates leading to the \( \alpha \)-aminoacylate intermediate and second half reaction substrates and products.
CAS:OASS activity, the individual changes failed to convert the substrate preference of OASS into that of a CAS. Nonetheless, the T185S mutant was sufficient to generate a more stable α-aminoacyl intermediate when incubated with Cys, suggesting this residue in OASS sets the reaction equilibrium so that it is preferable for Cys biosynthesis (see Supplemental Figure 2 online). Ultimately, combining all three changes into Gm-OASS successfully altered the reaction chemistry to that of a CAS and allowed formation of a stable α-aminoacyl intermediate from Cys (Figure 5C, Table 3). These results imply a dynamic interplay between active site structure, the substrate, and the subtle conformational changes that occur during catalysis (Austin et al., 2008). The effect of combining the three mutations suggests that multiple structural changes likely alter the environment around the PLP molecule, which in turn affects how the cofactor can move during the reaction sequence to shape the chemical landscape and allow Cys to form the reactive α-aminoacyl intermediate instead of the nonreactive external aldimine (Eliot and Kirsch, 2004). Thus, active site structure in CAS and OASS guides the chemical reaction sequence to tailor their respective biological activities.

Although the Gm-OASS T81M/S181M/T185S mutant functions as a CAS, its catalytic efficiency does not match that of the authentic Gm-CAS (Table 1). This suggests that additional changes, inside and outside the active sites, are required for optimization of function between CAS and OASS. Considering the dynamic features of CAS, such as the ligand-induced movement of the N-terminal domain and likely localized structural changes of PLP during the reaction sequence, multiple amino acid differences will contribute to the evolutionary changes that distinguish CAS and the efficient detoxification of CN− from OASS and its role in Cys biosynthesis.

The structure of Gm-CAS also has implications for how the plant BSAS family evolved multiple enzymatic activities (Figure 6B). In addition to CAS and OASS, other members of the BSAS enzyme family catalyze yet other specialized reactions. Cys desulphhydrase (DES) is a minor cytosolic isoform that regulates Cys homeostasis by catalyzing the elimination of sulfide from Cys to yield pyruvate in late developmental stages and under biotic/abiotic stress conditions (Alvarez et al., 2010; Álvarez et al., 2012). S-sulffocysteine synthase (SSCS) is a chloroplast-localized enzyme with a role in long-day light-dependent redox control in this organelle and catalyzes the addition of thiosulfate to O-acetylserylase (Bermúdez et al., 2010). At the biochemical level, all four types of BSAS (CAS, OASS, DES, and SSCS) share similar substrates and reaction chemistry. CAS and DES both use Cys as a substrate and release sulfide to form the α-aminoacyl intermediate. These enzymes differ in the second half reaction with CAS using CN− to complete the reaction sequence and DES preventing any nucleophilic addition. Similarly, both OASS and SSCS accept O-acetylserylase and sulfide, but larger nucleophiles (i.e., thiosulfate) are used by SSCS for the second half reaction.

The evolution of BSAS enzyme function needs to satisfy three requirements: binding of the first substrate, formation of a reactive intermediate from that substrate, and nucleophilic attack by the second substrate (Figure 6B). Using the structure of Gm-CAS as a guide, we have demonstrated how combination of three mutations in the active site of an OASS leads to formation of the reactive α-aminoacyl intermediate from Cys and improvement in CAS activity (Figure 5, Table 3). The DES from Arabidopsis also uses Cys as a substrate (Alvarez et al., 2010), but differs from both CAS and OASS at the three active site residues identified here (Met, Met, and Ser in CAS; Thr, Ser, and Thr in OASS; and Ile, Ala, and Thr in DES) (Figure 5). It remains to be determined whether these changes alter Cys versus O-acetylserylase preference in DES. The second half reaction also provides an added level of substrate discrimination, which leads to biochemical diversity in the BSAS enzyme family. Changes in the size of the active site entrance in the closed form of the enzyme would provide a steric gate to control nucleophile access to the α-aminoacyl intermediate. The biosynthesis of Cys in bacteria by two OASS isoforms (CysK and CysM) provides an example of this. CysK is specific for sulfide, whereas CysM accepts sulfide, thiosulfate, and a diverse range of larger nucleophiles (Maier, 2003; Claus et al., 2005). In this case, differences in surface loops were proposed to determine second substrate access to the active site; however, additional functional studies are required to demonstrate this plausible model in plant BSAS enzymes that display different reaction product specificity. The evolutionary diversification of biochemical function in the plant BSAS family to span essential reactions for CN− detoxification (CAS), Cys synthesis (OASS) and homeostasis (DES), and metabolism of other thiol-containing compounds (SSCS) likely results from the accumulation of multiple changes that tailor enzymatic properties.

METHODS

Protein Expression, Purification, and Mutagenesis

The pET-28a-Gm-CAS (Glyma09g93390; GmBSAS3;1) bacterial expression construct encodes for an N-terminally hexahistidine-tagged protein lacking the first 51 amino acids of a putative mitochondrial localization sequence. The Gm-OASS (Glyma11g00810; GmBSAS1;1) expression construct was previously described (Chronis and Krishnan, 2003). Site-directed mutagenesis to generate the Gm-CAS K95A mutant used the QuikChange PCR mutagenesis method (Agilent Technologies) (H. Yi and J.M. Jez, unpublished data).

Expression and purification of Gm-OASS used protocols similar to those used for Arabidopsis thaliana OASS (Bonner et al., 2005; Francois et al., 2006; Kumaran and Jez, 2007; Kumaran et al., 2009). For bacterial expression of wild-type and mutant Gm-CAS, Escherichia coli BL21(DE3) cells were transformed with each expression construct and grown in Terrific broth containing 50 μg L−1 kanamycin at 37°C, until A600 nm reached ~1.0. Addition of isopropyl-β-D-1-thiogalactopyranoside (1 mM final) was used to induce expression of N-terminal His-tagged protein, with the cells grown at 20°C for 18 h. Cells were harvested by centrifugation (3000g for 15 min at 4°C), and then resuspended in lysis buffer (50 mM Tris-HCl, pH 8.0, 20 mM imidazole, 500 mM NaCl, 1% glycerol, and 1% Tween-20). Sonication was used for cell lysis. Cell debris was removed by centrifugation (12,000g for 30 min at 4°C), and the supernatant was loaded on a Ni2+-nitrilotriacetic acid agarose (Qiagen). After extensive washing (50 mM Tris-HCl, pH 8.0, 20 mM imidazole, 500 mM NaCl, and 10% glycerol), the His-tagged proteins were eluted (50 mM Tris-HCl, pH 8.0, 250 mM imidazole, 500 mM NaCl, and 10% glycerol). Eluted proteins were further purified by size-exclusion chromatography on a HiLoad 26/60 Superdex-200 fast protein liquid chromatography column (GE Healthcare Life Sciences) using a buffer
 containing 25 mM HEPES (pH 7.5) and 100 mM NaCl. Fractions of purified protein with the expected size were pooled and stored at −80°C for enzyme assay after addition of final 10% glycerol or were concentrated to ~10 to 12 mg mL⁻¹ for protein crystallization. Protein concentration was determined by the Bradford method (Protein Assay; Bio-Rad) with BSA as standard.

**Enzyme Assays**

Assays of OASS activity were determined by measuring the amount of Cys produced, as previously described (Bonner et al., 2005). Analysis of CAS activity was performed in 0.5 mL reaction mix containing 100 mM Tris-HCl, pH 9.0, Cys (0 to 1 mM), and KCN (0 to 2 mM) (Han et al., 2008). All reactions were initiated by addition of protein (~200 ng for CAS or ~1 μg for OASS). The CAS reaction was terminated after 10 min of incubation at 25°C by addition of 50 μL of 30 mM FeCl₃ in 1.2 N HCl and 50 μL of 20 mM N,N-dimethyl-p-phenylenediamine dihydrochloride in 7.2 N HCl. The amount of methylene blue produced from one molecule of sulfide and two molecules of N,N-dimethyl-p-phenylenediamine dihydrochloride was colorimetrically determined at 670 nm. Steady state kinetic parameters were determined by nonlinear fitting of initial velocity versus substrate concentration to the Michaelis-Menten equation in SigmaPlot.

**Spectroscopic Analysis**

Binding of Cys, O-acetylserine, and β-cyanoalanine to Gm-CAS and Gm-OASS was monitored by recording the change in PLP signal using a Beckman DU800 UV/visible spectrophotometer (Bonner et al., 2005). Wavelength spectra (350 to 600 nm) of protein (200 to 500 μg) in the absence or presence of different ligands were obtained in 25 mM HEPES, pH 7.5, 100 mM NaCl.

**Protein Crystallization and Structure Determination**

Crystals of Gm-CAS were obtained by the vapor diffusion method in 4 μL hanging drops of a 1:1 mixture of protein and crystallization buffer (10% polyethylene glycol 8000, 0.2 M NaCl, and 0.1 sodium/potassium phosphate, pH 7.0, and 0.2 M MgSO₄) at 4°C over a 0.7-mL reservoir. The K95A mutant was crystallized in similar drops from a solution containing 20% polyethylene glycol 3350, 0.2 M potassium citrate, pH 7.5. All crystals were stabilized in cryoprotectant (crystallization solution with 30% glycerol) before flash freezing in liquid nitrogen. Diffraction data were collected at beamline 19ID of the Argonne National Lab Advanced Photon Source, with indexing, integration, and scaling performed with HKL3000 (Minor et al., 2006). The structure of Gm-CAS was solved by molecular replacement implemented in PHASER (McCoy et al., 2007) using the structure of Arabidopsis OASS (Boerner et al., 2005; Protein Data Bank [PDB]: 1Z7W) as a search model. Four PLPs and 233 water molecules. The model includes residues 51 to 367, 44 to 372, and 51 to 367.

**Accession Numbers**

Sequence data from this article can be found in the GenBank/EMBL databases under the following accession numbers: XP_003534555 (soybean CAS), ABF13209.1 (apple CAS1), ABF13210.1 (apple CAS2), NP_191703.1 (Arabidopsis CAS), CAA07177.1 (pinch CAS), AAAG8452.1 (rice CAS), AAL66921.1 (soybean OASS1), NP_001238392.1 (soybean OASS2), XP_003554689.1 (soybean OASS3), Q00834 (soybean OASS), and CAA56593 (Arabidopsis OASS). Coordinates and structure factors for wild-type Gm-CAS (PDB: 3VBE) and the Gm-CAS K95A mutant (PDB: 3VC3) have been deposited in the Research Collaboratory for Structural Bioinformatics PDB.

**Supplemental Data**

The following materials are available in the online version of this article.

**Supplemental Figure 1.** Ligand Interactions in Gm-CAS.

**Supplemental Figure 2.** Spectroscopic Analysis of Gm-OASS Mutants.

**ACKNOWLEDGMENTS**

This article was funded by a grant from the U.S. Department of Agriculture (NRI-2005-02518) to J.M.J. M.J. received an American Society of Plant Biologists Summer Undergraduate Research Fellowship. Portions of this research were carried out at the Argonne National Laboratory Structural Biology Center of the Advanced Photon Source, a national user facility operated by the University of Chicago for the Department of Energy Office of Biological and Environmental Research (DE-AC02-06CH11357).

**AUTHOR CONTRIBUTIONS**

H.Y., M.J., and J.M.J. designed the research; H.Y., M.J., and J.M.J. performed research; H.Y., M.J., and J.M.J. analyzed data; and H.Y. and J.M.J. wrote the article.

Received April 3, 2012; revised June 4, 2012; accepted June 11, 2012; published June 26, 2012.

**REFERENCES**


