Structural Basis for Evolution of Product Diversity in Soybean Glutathione Biosynthesis

Ashley Galant,a Kiani A.J. Arkus,a,b,1 Chloe Zubieta,b Rebecca E. Cahoon,b and Joseph M. Jeza,b,2

a Department of Biology, Washington University, St. Louis, Missouri 63130
b Donald Danforth Plant Science Center, St. Louis, Missouri 63132

The redox active peptide glutathione is ubiquitous in nature, but some plants also synthesize glutathione analogs in response to environmental stresses. To understand the evolution of chemical diversity in the closely related enzymes homoglutathione synthetase (hGS) and glutathione synthetase (GS), we determined the structures of soybean (Glycine max) hGS in three states: apoenzyme, bound to γ-glutamylcysteine (γEC), and with hGS, ADP, and a sulfate ion bound in the active site. Domain movements and rearrangement of active site loops change the structure from an open active site form (apoenzyme and γEC complex) to a closed active site form (hGS-ADP-SO$_4^{2-}$ complex). The structure of hGS shows that two amino acid differences in an active site loop provide extra space to accommodate the longer β-Ala moiety of hGS in comparison to the glycyl group of glutathione. Mutation of either Leu-487 or Pro-488 to an Ala improves catalytic efficiency using Gly, but a double mutation (L487A/P488A) is required to convert the substrate preference of hGS from β-Ala to Gly. These structures, combined with site-directed mutagenesis, reveal the molecular changes that define the substrate preference of hGS, explain the product diversity within evolutionarily related GS-like enzymes, and reinforce the critical role of active site loops in the adaptation and diversification of enzyme function.

INTRODUCTION

The tripeptide glutathione (GSH) is found in nearly all eukaryotes and prokaryotes and functions as a key component in an array of redox-linked cellular systems (Meister, 1995). In plants, GSH maintains cellular redox homeostasis, detoxifies harmful xenobiotics and heavy metals, and can regulate enzyme activity through glutathionylation (May et al., 1998; Noctor and Foyer, 1998; Rouhier et al., 2008). Although GSH is the predominant thiol-containing tripeptide found in plants, various plant species produce glutathione homologs in which the terminal Gly is substituted with a different amino acid (Figure 1A). For example, legumes make GSH, in addition to producing homoglutathione (hGSH), in which β-Ala replaces Gly, in a tissue-specific manner (Klapheck et al., 1995; Matamoros et al., 1999). Synthesis of hGSH maintains redox balance in legume nodules (Moran et al., 2000) and is critical for rhizobia-legume nodulation in roots (Matamoros et al., 2003; Fendo et al., 2005; Loscos et al., 2008). Similarly, many grasses synthesize GSH and hydroxymethylglutathione, with Ser instead of Gly, and exposure to cadmium activates the production of γ-glutamylcysteinylglutamate in maize (Zea mays; Rauser et al., 1986; Klapheck et al., 1994; Meuwly et al., 1995). The molecular details of how these peptides are generated and the biological functions of GSH analogs in plants are poorly understood, but these specialized peptides likely provide for specific responses to various environmental stresses.

Although the biosynthetic routes for the Ser- and Glu-containing peptides are unclear, the two-step pathways leading to GSH and hGSH are similar and better understood at the metabolic level. In the first reaction of the pathway, Glu-Cys ligase catalyzes the formation of γ-glutamylcysteine (γEC) from Glu and Cys (Jez et al., 2004; Hicks et al., 2007). The second step in the synthesis of either GSH or hGSH depends on the specificity of the synthetase for the terminal substrate. In nearly all organisms, glutathione synthetase (GS) catalyzes the addition of Gly to γEC (Meister, 1995; Jez and Cahoon, 2004; Herrera et al., 2007). In legumes, homoglutathione synthetase (hGS) uses β-Ala instead of Gly to form hGSH (Matamoros et al., 1999; Fendo et al., 2001; Iturbe-Ormaetxe et al., 2002). Although GS and hGS share similar reaction mechanisms based on biochemical and structural studies, the molecular basis for the difference in substrate specificity is unclear due to no available structural data for any plant GS or hGS.

Based on sequence similarity, both GS and hGS are members of the ATP-grasp enzyme superfamily (Galperin and Koonin, 1997). All ATP-grasp family members catalyze the ATP-dependent ligation of the carboxyl group carbon of one substrate to the amino- or imino-nitrogen of another substrate. For example, hGS catalyzes the transfer of the γ-phosphate group of ATP to the C-terminal carboxylate of γEC to yield an acylphosphate intermediate (Figure 1B). Subsequent nucleophilic attack on this intermediate by β-Ala leads to formation of hGSH with release of ADP and inorganic phosphate (Figure 1B). The structurally characterized tetrameric GS from Escherichia coli (Yamaguchi...
Protein Expression and Kinetic Analysis of hGS

Soybean hGS was overexpressed in E. coli as a His-tagged fusion protein and purified using Ni²⁺-affinity and size-exclusion chromatographies. Analysis of the protein by SDS-PAGE showed a monomeric molecular mass of 50 kD, which agrees with the predicted mass based on amino acid sequence (see Supplemental Figure 1 online). The protein eluted from the gel filtration column as a 102-kD species corresponding to a dimer (see Supplemental Figure 1 online). Other eukaryotic GS also are dimeric (Polekhina et al., 1999; Gogos and Shapiro, 2002; Jez and Cahoon, 2004). Purified recombinant hGS had a specific activity of 1.2 μmol min⁻¹ mg protein⁻¹ and required Mg²⁺ for activity. Steady state kinetic parameters of hGS for γEC, ATP, and β-Ala were determined (Table 1). In comparison to the GS from Arabidopsis (Jez and Cahoon, 2004; Herrera et al., 2007), hGS displayed a turnover rate (V/E) fivefold lower but with comparable Kₘ values for both ATP and γEC. In contrast with GS, which shows no activity if Gly is substituted with β-Ala, Ser, or Glu (Jez and Cahoon, 2004), hGS exhibited a 700-fold preference for β-Ala over Gly as the terminal substrate. Estimates of the turnover rate and Kₘ values of hGS with Gly should be considered as approximate because higher concentrations of Gly, and higher amounts of protein were required to observe activity. hGS did not accept either Ser or Glu as a substrate.

Results

Protein Expression and Kinetic Analysis of hGS

Electrophoretic analysis of the fusion protein and purified using Ni²⁺-affinity and size-exclusion chromatographies. Analysis of the protein by SDS-PAGE
between the closed and open active site structures (Figures 3B and 3C). In addition, a second loop (residues 479 to 491) forms part of the active site. For consistency with the human and yeast GS structures (Polekhina et al., 1999; Gogos and Shapiro, 2002), this second loop is referred to as the Ala-rich loop, even though the corresponding Ala residues are replaced by a Leu and a Pro in hGS. Dimerization of hGS occurs through a pseudo-twofold action between the closed and open active site structures (Figures 3A and 3C). With the exception of residues 410 to 416, the lid domain becomes ordered, with residues in the Gly-rich loop providing multiple interactions with the nucleotide. Likewise, the Ala-rich loop shifts to position residues for contact with both ADP and hGSH. As noted for the bacterial and eukaryotic GS (Polekhina et al., 1999; Gogos and Shapiro, 2002; Jez and Cahoon, 2004), nucleotide binding triggers movement of the lid domain and Ala-rich loop through multiple protein–ligand interactions. Based on the proposed reaction mechanism for GS (Herrera et al., 2007), enclosure of the active site likely prevents hydrolysis of the reactive acyl-phosphate intermediate (Figure 1B).

### γ-Glutamylcysteine Binding Site in the Open Form

In the reactions catalyzed by hGS and GS, γEC is a common substrate, and its binding site is highly conserved in both sequence and structure between hGS and the GS from human, yeast, and Arabidopsis. In the γEC binding site of hGS, Ser-176, Arg-295, Glu-241, and Gin-238 interact with the glutamyl portion of the molecule (Figure 4A). Of these, the charge–charge interaction between the Arg and the carboxylate group is critical for the reaction mechanism for GS (Herrera et al., 2007), suggesting an analogous role for this interaction in hGS. Similar to interactions observed in the structure of yeast GS complexed with γEC and an ATP analog (Gogos and Shapiro, 2002), Tyr-298 forms a hydrogen bond to the carbonyl of the glutamyl group and there is a bidentate charge–charge interaction between Arg-153 and the

### Domain Movements: Open and Closed Active Site Forms

The apoenzyme and γEC-bound structures are nearly identical, with an r.m.s. deviation of 0.5 Å. Disordered regions include most of the Gly-rich loop (residues 391 to 396) and other portions of the lid domain (residues 410 to 420) in each monomer of the dimer. For both open form structures, the Ala-rich loop is ordered in one monomer but disordered (residues 480 to 489) in the second monomer. In the open form, both the lid domain and Ala-rich loop are positioned away from the active site to reveal the binding sites for γEC and ATP and allow for substrate binding (Figures 3A and 3B).
carboxylate of the cysteinyl moiety (Figure 4A). Nearly all of these interactions are conserved when hGSH is bound in the site.

Active Site and Ligand Binding in the Closed Form

To define the active site, hGS was cocrystallized in the presence of reaction products ADP and hGSH (Figure 4B). In addition to the reaction products, a sulfate ion and three magnesium ions were identified in the active site of the closed form structure. Clear tetragonal density for the sulfate, which mimics binding of the inorganic phosphate product, was observed (Polekhina et al., 1999). Based on the positional similarity with the yeast and human GS structures, coordination, and strong electron density (4\(\sigma\)), three atoms were modeled as Mg\(^{2+}\).

As with the \(\gamma\)EC binding site, the residues forming the nucleotide binding site between the lid domain and Ala-rich loop of hGS (Figures 3B and 4B) are structurally conserved with those in the structures of human and yeast GS (Polekhina et al., 1999; Gogos and Shapiro, 2002). The adenosine ring forms main-chain contacts with Ile-427 and Gin-425 and a hydrogen bond with Lys-388. The ribose hydroxyl groups interact with Lys-477 and Glu-450, respectively. A series of polar interactions occur between the diphosphate tail and Lys-334, Asn-397, and two Mg\(^{2+}\) ions. The \(\alpha\)- and \(\beta\)-phosphate groups of the nucleotide and Glu-169 coordinate one Mg\(^{2+}\) with a second ion bound by the \(\beta\)-phosphate group, the sulfate, Glu-169, Asn-171, and Glu-392. Based on mechanistic studies of Arabidopsis GS, the magnesium ions and their coordinating residues play critical roles in stabilizing charges during catalysis (Herrera et al., 2007). The functional role of the third Mg\(^{2+}\) is unclear, as it does not interact with any of the bound ligands. This ion is coordinated by interactions with Glu-392 and main-chain contacts with Met-170 and Gly-332 that appear to help orient residues coordinated to the other Mg\(^{2+}\) ions.
Within the peptide binding site, all the interactions of the glutamyl portion of hGSH are identical to those observed in the γEC complex with minor differences in interactions with the cysteinyl group (Figure 4). Ser-176 is observed in alternate conformations. The side chains of Tyr-298 and Arg-153 are repositioned in the closed form complex. Tyr-298 rotates away from the tripeptide, and Arg-153 now interacts with the cysteinyl carbonyl group and the sulfate. The Arg is essential for catalyzing formation of the acylphosphate intermediate in the first part of the catalytic mechanism and in guiding nucleophilic attack in the second half of the reaction to yield the tripeptide product (Herrera et al., 2007). The carboxylate of the β-Ala moiety of hGSH forms a hydrogen bond with the backbone amide of Val-486 and an ionic interaction with the guanido group of Arg-475. Additional van der Waals contacts between the β-Ala–derived portion of hGSH are made with Leu-487 and Pro-488 in the Ala-rich loop. Interestingly, these two residues differ in hGS compared with GS.

Determination of Substrate Specificity and Product Diversity

In the active site of hGS, Leu-487 and Pro-488 are the only residues that differ from the characterized eukaryotic GS sequences (Figure 5). In GS, these residues are sequential Ala residues, which help give the Ala-rich loop its name. Structural comparison of hGS and human GS shows that the Ala-rich loop in hGS is shifted ∼3 Å away from the corresponding position of the loop in the GS structure to accommodate the larger β-Ala moiety (Figure 5).

To test the functional significance of Leu-487 and Pro-488 in determining the specificity of hGS for β-Ala over Gly, we generated Ala substitutions at each position (L487A and P488A) and the corresponding double mutant (L487A/P488A). Each mutant protein was expressed, purified, and assayed to determine steady state kinetic parameters for β-Ala and Gly as substrates (Table 3). Wild-type hGS displays a specificity ratio 708:1 in preference of β-Ala. Each point mutation altered substrate preference to different degrees. Although the L487A mutant shows a 3.4-fold reduction in catalytic efficiency with β-Ala and a 46-fold improvement using Gly as a substrate, this enzyme still prefers the hGS substrate by nearly fivefold. The P488A mutation yields an enzyme with almost equal preference for either substrate, resulting from a 274-fold increase in efficiency with Gly and a minor 2.3-fold reduction in \( k_{cat}/K_m \) with β-Ala. The L487A/P488A mutant retains activity with β-Ala at a 10-fold reduction compared with the hGS, but this mutant is as effective with Gly as the parent enzyme is with β-Ala. The combination of substitutions in the double L487A/P488A mutant converts hGS into a GS with a 950-fold increase in \( k_{cat}/K_m \) with Gly.
Functional diversity across enzyme families with shared threedimensional structures and reaction chemistry is a hallmark in the evolution of metabolic pathways. Nearly all eukaryotes and prokaryotes synthesize the multifunctional peptide GSH (Meister, 1995); however, some plants also synthesize GSH analogs with substitutions of the terminal Gly (Figure 1A) (Rauser et al., 1986; Klapheck et al., 1994; Klapheck et al., 1995; Meuwly et al., 1995; Matamoros et al., 1999). In particular, many legumes produce hGSH for root nodulation (Matamoros et al., 2003; Frendo et al., 2005; Loscos et al., 2008). As hGSH likely evolved from GS, we examined the structural basis for adaptation of product diversity in hGS. Crystallographic analysis of soybean hGSH provides insight on structural changes during the catalytic cycle of both hGS and GS and, combined with site-directed mutagenesis, defines active site differences that govern substrate preference. This work reinforces the critical role of flexible loops in the adaptation and diversification of enzyme function.

Catalysis in hGS and GS requires the orchestration of binding multiple substrates and the rearrangement of active site features, including the lid domain, Gly-rich loop, and Ala-rich loop. Together with studies of the kinetic and chemical mechanisms of GS (Jez and Cahoon, 2004; Herrera et al., 2007; Gogos and Shapiro, 2002; Gunasekaran et al., 2003). The structural conservation between the active sites of hGS and GS is substrate specificity for β-Ala and Gly, respectively. In each enzyme, residues in the Ala-rich loop contact the terminal residue of the tripeptide product (Figure 5). A Leu and Pro in the hGS from soybean and other legumes replaces the invariant double Ala sequence of the eukaryotic GS (Moran et al., 2000; Frendo et al., 2001; Iturbe-Ormaetxe et al., 2002; Skipsey et al., 2005). Structurally, the Ala-rich loop of hGS shifts relative to the same loop in GS to allow space for binding of the larger hGSH product and define active site differences that govern substrate preference. This work reinforces the critical role of flexible loops in the adaptation and diversification of enzyme function.

**DISCUSSION**

Functional diversity across enzyme families with shared threedimensional structures and reaction chemistry is a hallmark in the evolution of metabolic pathways. Nearly all eukaryotes and prokaryotes synthesize the multifunctional peptide GSH (Meister, 1995); however, some plants also synthesize GSH analogs with substitutions of the terminal Gly (Figure 1A) (Rauser et al., 1986; Klapheck et al., 1994; Klapheck et al., 1995; Meuwly et al., 1995; Matamoros et al., 1999). In particular, many legumes produce hGSH for root nodulation (Matamoros et al., 2003; Frendo et al., 2005; Loscos et al., 2008). As hGSH likely evolved from GS, we examined the structural basis for adaptation of product diversity in hGS. Crystallographic analysis of soybean hGSH provides insight on structural changes during the catalytic cycle of both hGS and GS and, combined with site-directed mutagenesis, defines active site differences that govern substrate preference. This work reinforces the critical role of flexible loops in the adaptation and diversification of enzyme function. 

Catalysis in hGS and GS requires the orchestration of binding multiple substrates and the rearrangement of active site features, including the lid domain, Gly-rich loop, and Ala-rich loop. Together with studies of the kinetic and chemical mechanisms of GS (Jez and Cahoon, 2004; Herrera et al., 2007), crystal structures of hGS (Figures 2 and 3) and GS (Polekhina et al., 1999; Gogos and Shapiro, 2002) now provide views of the progression through the catalytic cycle from apoenzyme (hGS and yeast GS) to second substrate complex (hGS-Gly complex) to second substrate complex (yeast GS in complex with Gly and an ATP analog) to product complex (hGS and human GS). Kinetic analysis of Arabidopsis GS indicates a mechanism in which γEC is the preferred first substrate followed by ATP (Jez and Cahoon, 2004). Within the active site, hGS shares common structural and chemical features with GS. In both enzymes, the γEC binding site is structurally static, whereas the ATP and β-Ala/Gly binding sites are dynamic. The structure of the hGS-γEC complex in the open active site conformation (Figures 3B and 4A) provides direct evidence for formation of this complex in agreement with the predicted mechanism for GS and hGS. Binding of ATP, which makes extensive contacts with residues in the lid domain, Gly-rich loop, and Ala-rich loop (Figures 3C and 4B), likely triggers transformation to the closed active site structure (Gogos and Shapiro, 2002; Gunasekaran et al., 2003). The closed active site conformation protects the reactive acylphosphate reaction intermediate from hydrolysis (Figure 1B) and orders the Ala-rich loop to form the binding site for either Gly or β-Ala (Figure 4B). Functionally, these conformational changes provide a cooperative linkage through the reaction cycle as binding of one substrate enhances binding of the next substrate, as suggested by the interaction factors in the kinetic mechanism (Jez and Cahoon, 2004). Structural/functional analysis of hGS and GS also suggests that the dynamic nature of the active site is important for catalysis and substrate recognition.

The structural conservation between the active sites of hGS and GS implies a shared reaction mechanism (Herrera et al., 2007). In the first half of the hGS reaction, formation of the acylphosphate intermediate occurs by transfer of the γ-phosphate of ATP to γ-glutamylcysteine. For this step, the Mg2+ ion in the active site orient the phosphate group and Arg-153 likely stabilizes the transition state. In the second half of the reaction, nucleophilic attack of the β-Ala amino group on the acylphosphate intermediate releases phosphate and yields hGSH. Positioning of Arg-153 and the Mg2+ bound by Glu-169 and Asn-171 would stabilize the transition state with the Ala-rich loop and Arg-475 orienting β-Ala for attack on the reaction intermediate to yield hGSH.

The major difference between hGS and GS is substrate specificity for β-Ala and Gly, respectively. In each enzyme, residues in the Ala-rich loop contact the terminal residue of the tripeptide product (Figure 5). A Leu and Pro in the hGS from soybean and other legumes replaces the invariant double Ala sequence of the eukaryotic GS (Moran et al., 2000; Frendo et al., 2001; Iturbe-Ormaetxe et al., 2002; Skipsey et al., 2005). Structurally, the Ala-rich loop of hGS shifts relative to the same loop in GS to allow space for binding of the larger hGSH product and β-Ala substrate (Figure 5A). Site-directed mutagenesis of Leu-487 and Pro-488 demonstrates that changes at both positions are necessary to convert hGS (Kcat/Km,β-Ala = 708 M⁻¹ s⁻¹) to a GS with comparable catalytic efficiency (Kcat/Km,β-Ala = 950 M⁻¹ s⁻¹) (Table 3). Interestingly, the L487A/P488A mutant retains limited activity with β-Ala (Kcat/Km,β-Ala = 77 M⁻¹ s⁻¹). This suggests that additional changes in the Ala-rich loop, or more subtle allosteric mutations, may be required to completely shift substrate preference and product specificity. The mobility of active site features in both hGS and GS (i.e., the lid domain and Ala-rich loop) likely plays a role in determining the rate of catalysis and in allowing evolutionary changes in these enzymes.

In both hGS and GS, structuring of the lid domain and Ala-rich loop appears linked to binding of ATP and the terminal substrate (i.e., β-Ala or Gly). Although the rate constants for each step in the catalytic cycle of either enzyme are unknown, the crystal structures of these enzymes suggest that dynamic active site structures may limit catalysis and explain the different turnover rates of GS (kcat ~ 12 s⁻¹) and hGS (kcat ~ 2 s⁻¹) (Table 1) (Gunasekaran et al., 2003; Tokuriki and Tawfik, 2009). Based on these results, it is possible that the nucleophilic attack of the terminal substrate is a limiting step in the reaction mechanism. Presumably, GS is a highly evolved enzyme in eukaryotes.

### Table 3. Substrate Specificity of Wild-type and Mutant Soybean hGS

<table>
<thead>
<tr>
<th>Substrate</th>
<th>V/Ei (s⁻¹)</th>
<th>Km (mM)</th>
<th>Kcat/Km (M⁻¹ s⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>β-Ala</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hGS</td>
<td>2.4 ± 0.1</td>
<td>3.4 ± 0.1</td>
<td>708</td>
</tr>
<tr>
<td>L487A</td>
<td>0.8 ± 0.1</td>
<td>3.8 ± 0.6</td>
<td>211</td>
</tr>
<tr>
<td>P488A</td>
<td>2.5 ± 0.2</td>
<td>8.0 ± 1.4</td>
<td>313</td>
</tr>
<tr>
<td>LP/AAα</td>
<td>1.9 ± 0.2</td>
<td>24.8 ± 2.3</td>
<td>77</td>
</tr>
</tbody>
</table>

Values are expressed as a mean ± SE for n = 3.

αDenotes the corresponding double mutant (L487A/P488A).
because of the central role that glutathione plays in regulating intracellular redox state (Meister, 1995). By contrast, hGS likely evolved by gene duplication and subsequent mutation (Todt et al., 1999; Penning and Jez, 2001; Gunasekaran et al., 2003; Tokuriki and Tawfik, 2009). The flexible and mutable nature of loops allows for the sampling of the new sequences and localized structures that generate shifts in substrate specificity or new catalytic activity. Frendo et al. (2001) originally proposed that legumes evolved hGS from gene duplication of GS after the divergence of the order Fabales, which includes the legumes, and Tawfik, 2009), and additional sequence changes in the lid domain and/or Ala-rich loop may be needed to optimize interactions with substrates and the movement of active site features.

Active site loops are central in the evolution of enzyme functionality (Todd et al., 1999; Penning and Jez, 2001; Gunasekaran et al., 2003; Tokuriki and Tawfik, 2009). The flexible and mutable nature of loops allows for the sampling of the new sequences and localized structures that generate shifts in substrate specificity or new catalytic activity. Frendo et al. (2001) originally proposed that legumes evolved hGS from gene duplication of GS after the divergence of the order Fabales, which includes the legumes, and Tawfik, 2009), and additional sequence changes in the lid domain and/or Ala-rich loop may be needed to optimize interactions with substrates and the movement of active site features.

While this work helps illuminate the molecular basis for hGS evolution from an ancestral GS, many questions remain as to the role hGSH and other GS analogs in plants. Although the interplay between genomes, protein function, and a plant’s environment shapes the evolution of new metabolism, it is unclear why legumes required evolution of hGS and hGSH production in nodules. Aside from the shared localization of hGS in nodules (Moran et al., 2000; Frendo et al., 2001; Iturbe-Ormaetxe et al., 2002; Skipsey et al., 2005), there appears to be no correlation between the presence of hGS in a legume species and the position of that species in the legume phylogeny (Wojciechowski et al., 2004). Nonetheless, given the conservation of hGS in the legumes examined so far, it seems likely that environmental factors, such as nodulation and/or habitat, contributed to the diversification of GSH metabolism. In addition, as suggested by the presence of Ser- and Glu-containing GSH analogs in other plants (Rauser et al., 1986; Klapheck et al., 1994; Meuwly et al., 1995), the adaptation of GSH biosynthesis for production of specialized tripeptides in response to environmental stresses may be more widespread. Continued genomic and biochemical explorations of legumes, and other plants, promise new insights on how these plants evolved more specialized environmental response systems.

**METHODS**

**Materials**

All oligonucleotides were synthesized by Integrated DNA Technologies. Np3-nitrotriacetic acid (NTA) was from Qiagen. Benzamidine-sepharose and the HiLoad 26/60 Superdex-200 FPLC column were purchased from GE/Amersham Health Sciences. The QuikChange site-directed mutagenesis kit was from Stratagene. hGSH was from Bachem. All other reagents were of ACS grade or better and were purchased from Sigma-Aldrich.

**Protein Expression, Purification, and Mutagenesis**

Soybean (Glycine max) hGS was PCR-amplified from a soybean seed cDNA library using 5′-dTTCGCATGGCATGGCTCAACCGTTGACC-ACC-3′ as the forward primer (the Ncol site is underlined, and the start codon is in bold) and 5′-dTGGCCGGCCGCTCAAAGTTAGGTATACAGTATCTACAC-3′ as the reverse primer (the Ncol site is underlined, and the stop codon is in bold). The resulting PCR product was digested with Ncol and NotI and then subcloned into pHIS8 (Jez et al., 2000) for expression of an N-terminally octahistidine-tagged protein. Automated nucleotide sequencing confirmed the fidelity of the bacterial expression construct (Washington University Sequencing Facility).

Transformed Escherichia coli BL21(DE3) cells were grown at 37°C in Terrific broth containing 50 μg mL−1 kanamycin until A600 < 0.8. After induction with 1 mM isopropyl 1-thio-β-D-galactopyranoside, the cultures were grown at 20°C for 4 to 8 h. Cells were pelleted by centrifugation (10,000g; 10 min) and resuspended in 50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 20 mM imidazole, 5 mM MgCl₂, 10% (v/v) glycerol, and 1% (v/v) Tween 20. Sonication was used to lyse cells. Following centrifugation (45,000g; 45 min), the supernatant was passed through a Ni²⁺-NTA column. The column was then washed with the same buffer minus Tween 20. His-tagged protein was eluted with 50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 250 mM imidazole, 5 mM MgCl₂, and 10% (v/v) glycerol. Incubation with thrombin (1/1000th the amount of hGS by weight) during overnight dialysis at 4°C against washbuffer removed the His tag. Diahydrated protein was reloaded on a mixed benzamidine-sepharose/Ni²⁺-NTA column. The flow-through of this step was loaded onto a HiLoad 26/60 Superdex-200 FPLC column equilibrated with 25 mM HEPES, pH 7.5, 5 mM MgCl₂, and 100 mM NaCl. Fractions containing purified protein were pooled, concentrated to 10 to 12 mg mL⁻¹, and stored at −80°C. Protein concentration was determined by the Bradford method (Protein Assay, Bio-Rad) with BSA as standard.

Site-directed mutants of hGS (L487A, P488A, and L487A/P488A) were generated using oligonucleotides containing the desired mutations (see Supplemental Table 1 online) and the QuikChange PCR method with the pHIS8-hGS vector as template. Introduction of the desired mutation was confirmed by sequencing of the constructs. Expression and purification of each mutant protein was performed as described for the wild-type protein.

**Enzyme Assays**

The activity of hGS was determined spectrophotometrically at 25°C by measuring the rate of formation of ADP using a coupled assay with pyruvate kinase and lactate dehydrogenase. A standard reaction mixture (0.5 mL) contained 100 mM HEPES, pH 7.5, 150 mM NaCl, 20 mM MgCl₂, 2.5 mM γ-EC, 10 mM β-Ala (or Gly), 2.5 mM disodium ATP, 2 mM sodium phosphonoxyvurate, 0.2 mM NADH, 5 units of type III rabbit muscle pyruvate kinase, and 10 units of type II rabbit muscle lactate dehydrogenase. The rate of decrease in A340 (ε = 6270 M⁻¹ cm⁻¹) was observed using a Beckman DU800 UV/Vis spectrophotometer. Steady state kinetic parameters were determined by initial velocity experiments in which concentrations for two substrates were fixed at saturating levels and the third substrate concentration varied (0.2 to 10 times the Kᵦₔ value). Untransformed data was fit to the Michaelis-Menten equation, v = kcat[S]/(Kᵦₔ + [S]), using Kaleidagraph (Synergy Software).

**Protein Crystallization and Structure Determination**

Crystals of hGS were obtained by the vapor diffusion method in 25°C by measuring the rate of formation of ADP using a coupled assay with pyruvate kinase and lactate dehydrogenase. A standard reaction mixture (0.5 mL) contained 100 mM HEPES, pH 7.5, 150 mM NaCl, 20 mM MgCl₂, 2.5 mM γ-EC, 10 mM β-Ala (or Gly), 2.5 mM disodium ATP, 2 mM sodium phosphonoxyvurate, 0.2 mM NADH, 5 units of type III rabbit muscle pyruvate kinase, and 10 units of type II rabbit muscle lactate dehydrogenase. The rate of decrease in A₃₄₀ (ε = 6270 M⁻¹ cm⁻¹) was observed using a Beckman DU800 UV/Vis spectrophotometer. Steady state kinetic parameters were determined by initial velocity experiments in which concentrations for two substrates were fixed at saturating levels and the third substrate concentration varied (0.2 to 10 times the Kᵦₔ value). Untransformed data was fit to the Michaelis-Menten equation, v = kcat[S]/(Kᵦₔ + [S]), using Kaleidagraph (Synergy Software).
Radiation Facility (SSRL) on monochromatic beamline 9-1. Diffraction data was integrated and reduced using XDS (Kabsch, 1993) and scaled with XSCALE (Kabsch, 1993). The structure of closed-form hGS in complex with ADP and hGSH was solved by molecular replacement performed with PHASER (McCoy et al., 2007) using a homology model of the soybean enzyme generated with SWISS-MODEL (Kopp and Schwede, 2003) from the structure of human GS (PDB: 2HGS; Polekhina et al., 1999). Model building was performed in O (Jones et al., 1993), and all refinements were performed with REFMAC (Murshudov et al., 1997). Waters were added using ARP (Lamzin and Wilson, 1993). Quality of the model was evaluated using PROCHECK (Laskowski et al., 1993). Structures of the open form hGS and open form hGS in complex with ADP and hGSH was solved by molecular replacement using a homology model of the soybean enzyme generated with SWISS-MODEL (Kopp and Schwede, 2003) from the structure of human GS (PDB: 2HGS; Polekhina et al., 1999). Model building was performed in O (Jones et al., 1993), and all refinements were performed with REFMAC (Murshudov et al., 1997). Waters were added using ARP (Lamzin and Wilson, 1993). Quality of the model was evaluated using PROCHECK (Laskowski et al., 1993). Structures of the open form hGS and open form hGS in complex with γEC were solved by molecular replacement using the final closed form hGS structure. Modeling building, refinement, and assessment were performed as above. Crystal parameters, data collection statistics, and refinement statistics for the three structures are summarized in Table 2. Atomic coordinates and structure factors have been deposited in the Protein Data Bank (www.rcsb.org). All structural figures were generated with PyMol (http://www.pymol.org).

Supplemental Data

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

Supplemental Figure 1. Protein Expression and Purification Analysis.

Supplemental Table 1. Oligonucleotide Primers Used for Site-Directed Mutagenesis.

ACKNOWLEDGMENTS

This work was funded by an American Chemical Society Petroleum Research Fund grant (ACS-PRF-43012-AC4) and a National Science Foundation grant (MCB-0824492) to J.M.J. K.A.J.A was a recipient of an ACS-PRF Summer Underrepresented Minority Research Fellowship. The SSRL Biotechnology Program is supported by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program, and the Department of Energy, Office of Biological and Environmental Research.

Received September 5, 2009; revised October 9, 2009; accepted November 5, 2009; published November 30, 2009.

REFERENCES


