Crystal structure of γ-glutamylcysteine synthetase: Insights into the mechanism of catalysis by a key enzyme for glutathione homeostasis

Takao Hibi†, Hiroshi Nii†, Toru Nakatsu‡, Akira Kimura§, Hiroaki Kato‡, Jun Hiratake¶, and Jun’ichi Oda†

1Department of Bioscience, Fukui Prefectural University, Fukui 910-1195, Japan; 2Department of Structural Biology, Graduate School of Pharmaceutical Sciences, Kyoto University, Kyoto 606-8501, Japan; 3Greenbio Co. Ltd., 15 ShimogamoMorimoto-cho, Sakyo-ku, Kyoto 606-6805, Japan; and 4Institute for Chemical Research, Kyoto University, Uji, Kyoto 611-0011, Japan

Edited by Gregory A. Petsko, Brandeis University, Waltham, MA, and approved September 7, 2004 (received for review May 12, 2004)

γ-Glutamylcysteine synthetase (γGCS), a rate-limiting enzyme in glutathione biosynthesis, plays a central role in glutathione homeostasis and is a target for development of potential therapeutic agents against parasites and cancer. We have determined the crystal structures of Escherichia coli γGCS unliganded and complexed with a sulfoximine-based transition-state analog inhibitor at resolutions of 2.5 and 2.1 Å, respectively. In the crystal structure of the complex, the bound inhibitor is phosphorylated at the sulfoximido nitrogen and is coordinated to three Mg2+ ions. The cysteine-binding site was identified; it is formed inductively at the transition state. In the unliganded structure, an open space exists around the representative cysteine-binding site and is probably responsible for the competitive binding of glutathione. Upon inhibitor binding, the side chains of Tyr-241 and Tyr-300 turn, forming a hydrogen-bonding triad with the carbonyl group of the inhibitor’s cysteine moiety, allowing this moiety to fit tightly into the cysteine-binding site with concomitant accommodation of its side chain into a shallow pocket. This movement is caused by a conformational change of a switch loop (residues 240–249). Based on this crystal structure, the cysteine-binding sites of mammalian and parasitic γGCSs were predicted by multiple sequence alignment, although no significant sequence identity exists between the E. coli γGCS and its eukaryotic homologues. The identification of this cysteine-binding site provides important information for the rational design of novel γGCS inhibitors.

Glutathione homeostasis is critically important for maintaining both intracellular redox balance and defense against oxidative or chemical stress (1–3). γ-Glutamylcysteine synthetase (L-glutamate:L-cysteine γ-ligase, EC 6.3.2.2; γGCS) catalyzes the first and rate-limiting step of glutathione biosynthesis. The enzyme’s activity is precisely controlled by nonallosteric feedback inhibition by glutathione, the limited availability of cellular L-Cys, the physiologically indispensable component of glutathione, and parasitic organisms against parasites and cancer. We have determined the crystal structures of Escherichia coli γGCS unliganded and complexed with a sulfoximine-based transition-state analog inhibitor at resolutions of 2.5 and 2.1 Å, respectively. In the crystal structure of the complex, the bound inhibitor is phosphorylated at the sulfoximido nitrogen and is coordinated to three Mg2+ ions. The cysteine-binding site was identified; it is formed inductively at the transition state. In the unliganded structure, an open space exists around the representative cysteine-binding site and is probably responsible for the competitive binding of glutathione. Upon inhibitor binding, the side chains of Tyr-241 and Tyr-300 turn, forming a hydrogen-bonding triad with the carbonyl group of the inhibitor’s cysteine moiety, allowing this moiety to fit tightly into the cysteine-binding site with concomitant accommodation of its side chain into a shallow pocket. This movement is caused by a conformational change of a switch loop (residues 240–249). Based on this crystal structure, the cysteine-binding sites of mammalian and parasitic γGCSs were predicted by multiple sequence alignment, although no significant sequence identity exists between the E. coli γGCS and its eukaryotic homologues. The identification of this cysteine-binding site provides important information for the rational design of novel γGCS inhibitors.

Experimental Procedures

Crystallization and Data Collection. The unliganded γGCS used in this study (518 amino acid residues; molecular mass, 58,205 Da) was mutated, expressed, purified, and crystallized as described (16). Selenomethionine (Se-Met)-substituted protein was obtained by the pathway inhibition method (17), and was purified and crystallized by using the same protocol as for the native protein. Both the unliganded and derivative crystals were diazed against mother liquor containing 20% glucose for flash cooling.

In preparing the complexed crystals, the purified enzyme was incubated with 20 mM of I in the presence of 5 mM ATP and 50 mM MgCl2 at 25°C until its activity was completely abolished, as described (14). The complex was crystallized by using the sitting-drop vapor diffusion method against 15% (w/vol) polyethylene glycol (PEG) 2000 MME and 250 mM MgCl2 in 20 mM

This paper was submitted directly (Track II) to the PNAS office.

Freely available online through the PNAS open access option.

Abbreviations: γGCS, γ-glutamylcysteine synthetase; BSO, L-buthionine sulfoximine; rmsd, rms distance.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 1V4G (unliganded model) and 1VA6 (complexed model)).

†To whom correspondence should be addressed at: Department of Bioscience, Fukui Prefectural University, 4-1-1 Kenjojouma, Matsuoka-cho, Yoshida-gun, Fukui 910-1195, Japan. E-mail: oda@fpus.ac.jp.

© 2004 by The National Academy of Sciences of the USA
Data were processed, merged, and scaled with MOSFLM and the anomalous diffraction (MAD) data were collected. Intensity BL41XU. The XAFS spectrum of the selenomethionine derivative showed a clear peak and edge, and multiwavelength anomalous diffraction (MAD) data were collected. Intensity were processed, merged, and scaled with MOSFLM and the CCP4 program suite (18) or CRYSTALCLEAR software (Rigaku/MSC, Tokyo). Data collection statistics are given in Table 1.

**Structure Solution and Refinement.** Multiwavelength anomalous diffraction data in the resolution range of 40 to 2.6 Å were of sufficient quality to identify the positions of all 52 Se atoms by using SOLVE (19). The program RESOLVE (20) was used to generate the initial solvent-flattened electron density map for the polypeptide chains with distinct solvent boundaries. The initial model was built with the program O (21) and refined with CNS (22). After an initial round of refinement, the model was refined against the 2.5-Å data of the unliganded enzyme. The space group of this crystal was P2₁, and two copies of each model were searched for in the asymmetric unit. The rotation and translation functions were determined over the resolution range of 10 to 3 Å. After several rounds of iterative maximum likelihood positional and TLS refinement, the electron density from the Fo – Fe map depicted a clear density for MgADP and the sulfoximine. The structures of the two enzyme molecules in the asymmetric unit were essentially the same, because their Ca atoms exhibited a low rmsd of 0.30 Å. The quality of the final structures was assessed with the PROCHECK (26) and WHATCHECK (27) programs. Statistics for refinement are given in Table 1.

**Results and Discussion**

**Overall Structure of the γGCS-Inhibitor Complex.** Fig. 1 shows that the crystal structure comprises two structural domains: a catalytic domain (residues 18–387 and 442–518) and a small domain (residues 1–16 and 388–441). These domains are linked by swapping their terminal short chains (α1, α12, and α13) with each other and through a disulfide bridge (Cys–372–Cys–395), which is quite rare in cytosolic proteins. The catalytic domain contains six anti-parallel β-strands that form a curved partial barrel with a funnel-shaped internal cavity. The electron-density peak corresponding to the inhibitor appears near the tip of the funnel’s cone, and that representing the nucleotide is found in the cone’s interior. Two arm-like structures cover the open side of each monomer.

Table 1. Data collection and refinement statistics

<table>
<thead>
<tr>
<th>Data collection statistics</th>
<th>γGCS</th>
<th>Se-Met-γGCS</th>
<th>γGCS-sulfoximine 2-MgADP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>R3</td>
<td>R3</td>
<td>P2₁</td>
</tr>
<tr>
<td>Unit cell parameters</td>
<td>a = b = 326.82 Å</td>
<td>a = b = 327.03 Å</td>
<td>a = 70.45 Å, b = 97.36 Å</td>
</tr>
<tr>
<td></td>
<td>c = 104.73 Å</td>
<td>c = 104.45 Å</td>
<td>c = 102.18 Å, β = 109.63°</td>
</tr>
<tr>
<td>No. of monomers/asymmetric unit</td>
<td>4</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Wavelength, Å</td>
<td>0.9726</td>
<td>0.9792 (peak)</td>
<td>0.9794 (edge) 0.9839 (remote) 1.0000</td>
</tr>
<tr>
<td>Resolution, Å</td>
<td>47.2–2.50 (2.59–2.50)</td>
<td>55.0–2.60 (2.74–2.70)</td>
<td>48.8–2.10 (2.21–2.10)</td>
</tr>
<tr>
<td>Total no. of reflections</td>
<td>776,878</td>
<td>724,749</td>
<td>711,447</td>
</tr>
<tr>
<td>No. of unique reflections</td>
<td>144,279</td>
<td>127,496</td>
<td>127,456</td>
</tr>
<tr>
<td>Completeness, %</td>
<td>100.0 (100.0)</td>
<td>100.0 (100.0)</td>
<td>100.0 (100.0)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>5.4 (5.3)</td>
<td>5.7 (5.4)</td>
<td>5.6 (5.3)</td>
</tr>
<tr>
<td>l/σ</td>
<td>9.0 (3.4)</td>
<td>9.1 (2.1)</td>
<td>9.1 (1.7)</td>
</tr>
<tr>
<td>Rmerge (Σ</td>
<td></td>
<td>h00</td>
<td></td>
</tr>
</tbody>
</table>

| Refinement statistics      | | | |
| Resolution, Å              | 40.0–2.5 | 40.0–2.6 | 40.0–2.1 |
| Rcryst                     | 0.206 (0.288) | 0.199 (0.283) | 0.225 (0.298) |
| Rfree                      | 0.236 (0.331) | 0.257 (0.381) | 0.260 (0.395) |
| No. of protein atoms       | 15,901 | 7,998 |
| No. of solvent molecules   | 316   | 346 |
| rmsd bond length, Å        | 0.021 | 0.017 |
| rmsd angles, °             | 1.8   | 1.4 |
| Ramachandran plot, %       | Most favored/allowed regions | 91.1/100.0 | 91.0/99.8 |

Values in parentheses are for the highest-resolution shell.
of the β-barrel in the active site. The amino acid sequence of these arms varies widely among γGCS family members, and sequence homology analysis revealed no regions corresponding to these sequences among glutamine synthetase family members (12, 28). Thus, we termed these structures the N-terminal variable arm (residues 105–144) and the central variable arm (residues 240–298), respectively. The inhibitor model obtained is anchored by Tyr-131 and Leu-135 on the N terminus of the N-terminal variable arm and by Tyr-241 at the N terminus of the central variable arm, respectively. Yellow and orange letters show the region corresponding to the switch loop and the serine residues substituted for crystallization (16), respectively. (C) SIGMAA-weighted Fo − Fc omit electron density maps for the inhibitor molecule 2, three Mg2+ ions, and ADP (blue) and for only the Mg2+ ions (red). These maps are contoured at 3 and 8 σ, respectively. Inhibitor and ADP are drawn as stick models, and the protein is drawn as a line model. Mg2+ ions and water molecules are shown as green and red spheres, respectively.

The Nucleotide- and Metal-Binding Sites. The γ-phosphate of ATP has already been transferred to the NS sulfoximine nitrogen atom of the inhibitor 1 to form the N-phosphoryl sulfoximine 2. The phosphorylation of nitrogen, as noted first by Griffith, using BSO (29), mimics the phosphorylation of the γ-carboxyl group of L-Glu. The observed R-configuration around the sulfur atom in 2 is consistent with the previous finding that the sulfoximine-based compound 1, with its R-sulfur atom, is an extremely potent time- and ATP-dependent inhibitor of E. coli γGCS, whereas the corresponding (S)-sulfoximine- or sulfone-based compounds were weak and simple reversible inhibitors (14). The same relative configuration as that of the sulfur atom in 2 is thought to occur in a tetrahedral adduct, suggesting that L-Cys approaches from the si-face of the carbonyl of the γ-glutamylphosphophate intermediate.

A constellation of three important Mg2+ ions is found around the phosphate-binding pocket; these ions are located at sites termed n1, n2, and n3 (12). The n1 Mg2+ is coordinated to Glu-29 OE2, Asp-60 OD1, Glu-67 OE1, a water molecule, and atoms of both the NS nitrogen and OP1 oxygen of the N-phosphoryl sulfoximido group of 2 (Fig. 2). The NS and OP1
atoms of 2 are located at equatorial positions in the octahedral coordination group, so that the OS oxygen atom of the sulfoximine is oriented toward Arg-330 to accept symmetrical bifurcated hydrogen bonds from its guanidino group. In addition to being coordinated with the n1 Mg$^{2+}$, the carbonyl group of the γ-glutamylphosphate intermediate is also probably anchored by Arg-330, thus orienting its si-face toward the Cys-binding site (see below). Furthermore, the guanidino group of Arg-330 can polarize the carbonyl group and then stabilize the oxyanion formed in the transition state to promote the reaction.

The other two Mg$^{2+}$ ions form a bridge between 2 and ADP. The n2 Mg$^{2+}$-binding site consists of Glu-27 OE1 and OE2, His-150 ND1, Glu-328 OE2, the oxygen atom O3B of the γ-phosphate group of ADP, and the oxygen atom OP3 of the transferred phosphate group. The n3 site is constructed from Glu-27 OE1, Glu-67 OE2, a water molecule, the oxygen atoms O2A and O3B, respectively, of the α- and β-phosphate groups of ADP, and the oxygen atom OP1 of the transferred phosphate. Although the presence of a third Mg$^{2+}$ ion has not yet been reported for GCS, this pattern of Mg$^{2+}$ binding has been observed in pyruvate kinase (30), glutathione synthetase (31), D-Ala-D-Ala ligase (32), and D-Ala-D-lactate ligase (33). Therefore, the bridging coordination of the n2 and n3 Mg$^{2+}$ ions may be a common catalytic motif found in the family of ATP-dependent ligases (ADP-forming), and is probably responsible for increasing the reactivity of the γ-phosphate group of ATP and for coordinating, and thus stabilizing, the γ-glutamylphosphate intermediate (34, 35). These Mg$^{2+}$ ions also play a key role as Lewis acids in facilitating the elimination of ADP and phosphate.

The Glutamylate-Binding Site. In contrast to its relatively loose fit of the cysteine-binding site, γGCSs bind L-Glu and its analogs with quite high specificity (11, 36). The glutamate moiety of 2 occurs in an extended conformation in the interior of the hydrophilic stem of the active-site funnel, which also contains part of a polyethylene glycol 400 molecule that was added as a cryoprotectant. The α-amino group of 2 is bound to the main-chain carbonyl group from Ile-146 located at the C terminus of the short α helix o4 in the N-terminal variable arm. Meanwhile, the α-carboxyl group of 2 is bound to both the guanidino group of Arg-235 and the imidazolyl group of His-150. Only small differences are found near this site between the unliganded and the complexed structures, except for the side chain of Tyr-241, which covers the glutamate moiety in the hydrophilic stem in the inhibitor-bound structure, as discussed below.

The Cysteine-Binding Site. Fig. 3 illustrates the residues around the moiety of 2 that mimic L-α-aminobutyrate (L-Abu), a substitute for L-Cys. Studies of the substrate specificity of the E. coli and mammalian enzymes have indicated that, whereas the E. coli enzyme exhibits high specificity for L-Cys and accepts only a few unnatural amino acids, whose side chains closely approximate the size and polarity of that of L-Cys (36), homologs from mammals such as rat and bovine accept a variety of natural and unnatural amino acids (11). The ethyl (-C2H5) side chain of 2 fills up a shallow pocket that is open to the cavity in the funnel cone and is built from the hydrophobic residues Phe-61, Tyr-131, and Leu-135, which make up its wall, and the polar side chain of
Arg-132, which forms its floor. The terminal methyl of the \(-\text{C}_2\text{H}_5\) side chain is 4.2 Å from the guanidino group of Arg-132; no significant polar interactions are expected between these groups. However, the side chain of L-Cys, which is larger, would be \(\approx 0.3\) Å closer to the bottom; polar interaction with this Arg residue would increase its binding affinity. In fact, the \(K_m\) value for L-Cys (0.10 mM) is 40 times smaller than that for L-Abu (3.9 mM), and \(\beta\)-chloro- and \(\beta\)-cyanoalanine are good substrates for \(E.\ coli\) \(\gamma\)GCS (36).

The \(O_2\) oxygen atom of the terminal carboxyl group of 2 is stabilized by hydrogen bonds from Tyr-131 and from Tyr-300, which in turn is involved in a hydrogen-bonding triad with Tyr-241 and Gln-144. This hydrogen-bond network accounts for a significant contribution (3.8 kcal/mol\(^{-1}\)) of the carboxyl group of inhibitor 2 to the overall binding energy, and is evidenced by the lower affinity of BSO, a carbonyl-free version of 2 (15).

There are no specific residues around the -CH\(_2\)- next to the sulfur, implying that the enzyme does not have a basic residue to abstract a proton from the \(\alpha\)-NH\(_3\) group of L-Cys. The lack of a catalytic base for this purpose has also been noted in the mechanistically related ATP-dependent peptide ligases, such as glutathione synthetase (37) and \(d\)-Ala-d-Ala ligase (32).

**Induced Fit Essential for Intrinsic Binding of Cysteine at the Transition State.** Comparison of the complexed and unliganded crystal structures reveals major conformational changes around the central variable arm induced by the binding of transition-state analog 2 (Fig. 4). The mechanisms by which \(\gamma\)GCS properly recognizes its essential substrate, L-Cys, and the competitive inhibitor, glutathione (38), which regulates the synthesis of \(\gamma\)GCS, are unknown. In the unliganded structure, the side chains of Tyr-241 and Tyr-300 are located out of the active site and are stacked loosely on top of each other. In this state, an open space remains around the representative Cys-binding site. This space is large enough to accommodate the cysteinylglycine moiety of glutathione and is probably responsible for the feedback inhibition by glutathione. Upon binding of the inhibitor 2, an altered path of a switch loop (residues 240–249; Fig. 5) is facilitated, and this conformational change of the loop permits an inward shift of a switch loop (residues 240–249; Fig. 5) is facilitated, and a significant contribution (3.8 kcal/mol\(^{-1}\)) of the carboxyl group of inhibitor 2 to the overall binding energy, and is evidenced by the lower affinity of BSO, a carbonyl-free version of 2 (15).

As a result, the C-terminal carboxyl group of 2, which corresponds to the carboxyl group of L-Cys, is recognized by the hydrogen-bonding network involving Tyr-131, Gln-144, Tyr-241, and Tyr-300. This hydrogen-bonding network probably contributes to the orientation of L-Cys for nucleophilic attack toward the carbonyl of the \(\gamma\)-glutamylphosphate intermediate from its si-face. The cooperative binding of both the carboxyl group and the side chain of L-Cys by these induced-fit mechanisms (39) thus establishes the tetrahedral geometry at the transition state, which the inhibitor 2 mimics.

These major conformational changes may be responsible for the slow-binding kinetics commonly observed for the inhibition of \(\gamma\)GCS by sulfoximine-based inhibitors such as BSO and compound 1. This speculation is consistent with the view that the observed slow-binding kinetics are not due to a chemical process, such as the phosphorylation of the NS nitrogen, because they

![Fig. 5. Superimposition of residues 238–251, including the switch loop. The loop’s hinge residues, Gly-240 and Leu-249, are labeled.](Image)

![Fig. 6. Multiple sequence alignment of N-terminal and central variable arm sequences for prokaryotic and eukaryotic \(\gamma\)GCSs. Ec, \(E.\ coli\) K strain; Vc, \(Vibrio\ cholerae\); Pa, \(Pseudomonas\ aeruginosa\); Ba, \(Buchnera\ aphidicola\); Hu, human \(\gamma\)GCS heavy chain; Ra, rat \(\gamma\)GCS heavy chain; Tc, \(Trypanosoma\ cruzi\); Tb, \(T.\ brucei\); Pf, \(P.\ falciparum\). The first and last residue numbers of each sequence are indicated, and the total length of each protein is shown at the end. Residues that are inserted or that occur between the arms are not displayed, and the number of omitted residues is in parentheses. Residues important to the binding of the sulfoximine inhibitor in the \(E.\ coli\) enzyme are highlighted in red, and the hinge residues of the switch loop are in blue. The substitution of His-370 in the human \(\gamma\)GCS (green), which has been found to cause low red-blood cell glutathione levels and hereditary hemolytic anemia (41), is proposed to be included in the central variable arm.](Image)
persisted when chemically synthesized N-phosphoryl sulfoximine 2 was used as an inhibitor (14).

**Proposed Cysteine-Binding Sites of Other γGCS Members.** In contrast with the highly conserved sequences around the MgATP- and Glu-binding sites, the sequences of the N-terminal and central variable arms of the *E. coli* γGCS are quite different from those of other γGCSs. This result has made it difficult to predict the Cys-binding site by homology search experiments. Using the crystal structure shown here, we obtained a set of predicted multiple alignments of the sequences of the variable arms from a number of eukaryotic and prokaryotic enzymes. The alignment predicts several conserved residues that form part of the putative Cys-binding site (Fig. 6). Tyr-241 is conserved in all of the species shown, whereas Leu-135 and Gln-144 are replaced by the functionally equivalent amino acids Ile and Asn, respectively. Among the eukaryotic enzymes, Tyr-131 and Tyr-300, which are replaced with Thr and Trp, respectively. It is not clear to what extent these substitutions would preserve or weaken the hydrogen-bonding network seen in the *E. coli* enzyme, and the resulting impact on binding affinity. Interestingly, a well-conserved Arg-132 in the side-chain pocket of the cysteine-binding site is replaced by Asn and Gin in the enzymes from *T. brucei* and *P. falciparum*, respectively. This result has made it difficult to predict the Cys-binding site shown here provides hints of the Cys-binding site is important in determining the enzyme's specificity for the side chain of cysteine substrates and are inhibited potently by BSO and its higher homologs (29), which lack the terminal carboxyl group. On the other hand, *E. coli* γGCS is only weakly inhibited by BSO (14, 36), whereas the sulfoximine I affects it potently (14, 15). Therefore, the structure of the Cys-binding site is important in determining the enzyme's sensitivity to inhibitors and specificity for its substrates. In this regard, the predicted Cys-binding site shown here provides hints for the rational design of novel γGCS inhibitors to control the cellular levels of glutathione in biological and medical contexts.

Synchrotron radiation experiments were conducted at SPring-8 with the approval of the Japan Synchrotron Radiation Research Institutes (JASRI) (Proposal No. 2001A0084-NL-np, 2002A0440-NL1-np, and 2002B0341-NL1-np). This work was supported in part by grants-in-aid for scientific research from the Ministry of Education, Science, and Culture and the National Project on Protein Structural and Functional Analysis. This work was also supported in part by the Protein 3000 Project from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan. J.O. and T.H. were supported by a grant-in-aid for scientific research from Fukushima University.