Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Mutation of Gly51 to serine in the P-loop of *Lactobacillus casei* folylpolyglutamate synthetase abolishes activity by altering the conformation of two adjacent loops

Clyde A. Smith,[a]* Jennifer A. Cross,[b]‡ Andrew L. Bognar[c] and Xiaolin Sun[b]§

[a]Stanford Synchrotron Radiation Laboratory, Stanford University, Menlo Park, CA, USA, [b]School of Biological Sciences, University of Auckland, Auckland, New Zealand, and [c]Department of Medical Genetics and Microbiology, University of Toronto, Toronto, Canada

‡ Present address: ESRF, 6 Rue Jules Horowitz, 38043 Grenoble, France.
§ Present address: HortResearch, Palmerston North, New Zealand.

Correspondence e-mail: csmith@slac.stanford.edu

Received 18 February 2006
Accepted 15 March 2006

PDB References: apoFPGS, 2gca, r2gcasf; G51S mutant, 2gc5, r2gc5sf; G51S/S52T double mutant, 2gcb, r2gcbsf; S73A mutant, 2gc6, r2gc6sf.

Based upon the three-dimensional structure of *Lactobacillus casei* folylpolyglutamate synthetase (FPGS), site-directed mutagenesis studies were performed on three residues associated with the ATPase site: Gly51, Ser52 and Ser73. Gly51 and Ser52 are at the end of the P-loop, which is involved in triphosphate binding. A G51S mutant enzyme and a G51S/S52T double-mutant enzyme were made in order to alter the FPGS P-loop to more closely resemble the sequences found in other ATPase and GTPase enzymes. Ser73 is on a neighboring loop (the Ω-loop) and precedes a proline residue found to be in a *cis* conformation. The carbonyl O atom of Ser73 is one of the protein ligands for the essential Mg^{2+} ion involved in ATP binding and hydrolysis and the Ω-loop is involved in binding the folate substrate 5,10-methylenetetrahydrofolate. The serine residue was mutated to alanine and this is the only one of the three mutants which retains some FPGS activity. The structures of the G51S, G51S/S52T and S73A mutant proteins have been solved to high resolution, along with the structure of the apo wild-type FPGS. The P-loop in both the G51S and G51S/S52T mutant proteins remains unaltered, yet both structures show a large conformational rearrangement of the Ω-loop in which a *cis*-Pro residue has switched conformation to a *trans*-peptide. The structure of the Ω-loop is severely disrupted and as a consequence structural rearrangements are observed in the peptide linker joining the two domains of the enzyme. Magnesium binding in the active site is also disrupted by the presence of the serine side chain at position 51 and by the repositioning of the carbonyl O atom of Ser73 and a water molecule is bound in place of the Mg^{2+} ion. The S73A mutant protein retains the *cis*-Pro configuration in the Ω-loop and the Mg^{2+} site remains intact. The *cis*-Pro is also observed in the structure of the substrate-free form of FPGS (apoFPGS), maintained in the absence of Mg^{2+} by a hydrogen-bonding network involving water molecules in the active site. It is only in the complete absence of water or Mg^{2+} in the binding site that the *cis*-Pro switches to the *trans* conformation.

1. Introduction

Bacteria and eukaryotes require folate coenzymes in a number of important metabolic cycles including the biosynthesis of methionine and thymidylate and the *de novo* synthesis of purines, where they serve as carriers of one-carbon (C1) units (Shane, 1989), typically in the form of 5,10-methylene-tetrahydrofolate (5,10-mTHF) or 10-formyl-tetrahydrofolate. Endogenous folates in the cell are predominantly in the form of polyglutamate derivatives (formed by the addition of L-glutamate groups at the γ-carboxy of the terminal gluta-

mate) that are the normal substrates for enzymes involved in C1 metabolism. It has been shown that many folate-dependent enzymes have a higher affinity for folylpolyglutamates (McGuire & Coward, 1984; Shane, 1989) and antifolates. The MgATP-dependent polyglutamylation of folate coenzymes is catalysed by the enzyme folylpolyglutamate synthetase (FPGS; EC 6.3.2.17), involving a complex mechanism which requires the formation of a quaternary complex with MgATP, a folate derivative and L-glutamate, followed by the ordered release of the products, MgADP, the glutamylated folate and phosphate (Bognar & Shane, 1983; Cichowicz & Shane, 1987).

Three crystal structures of the wild-type *Lactobacillus casei* FPGS have been determined: the MgATP complex (FPGS–MgATP) to 2.4 Å resolution (Sun *et al.*, 1998), the binary MgAMPPCP complex (FPGS–MgAMPPCP) to 1.9 Å resolution and a ternary complex containing 5,10-mTHF and AMPPCP (FPGS–MgAMPPCP–5,10-mTHF) to 1.85 Å resolution (Sun *et al.*, 2001). More recently, the structures of FPGS from *Thermotoga maritima* (PDB code 1o5z) and *Escherichia coli* (Mathieu *et al.*, 2005) have been reported. All of the FPGS structures are essentially identical and show that the molecule has a modular two-domain architecture (Sun *et al.*, 1998). The active site is located at the interface between the two domains and the ATP-binding site is associated with a glycine-rich P-loop (Fig. 1) linking a β-strand and an α-helix, similar to the P-loop found in a number of other mononucleotide-binding enzymes including *ras*-P21, adenylate kinase, the G proteins and myosin (Schulz, 1993; Smith & Rayment, 1996*a*). In these enzymes, a hydroxyl residue at the end of the P-loop is one of the ligands for the 'classical' Mg^{2+} associated with ATP binding and hydrolysis. The other ligands are normally a second hydroxyl residue from a neighboring loop, some water molecules and the ATP itself (Smith & Rayment, 1996*a*). In FPGS, however, there is a glycine residue at the equivalent position in the P-loop (Fig. 1) and the 'classical' Mg^{2+} site is formed by a highly conserved glutamate residue from a neighboring loop (Glu143; Fig. 1) and the carbonyl O atom of Ser73 (Sun *et al.*, 1998, 2001; Sheng *et al.*, 2002).

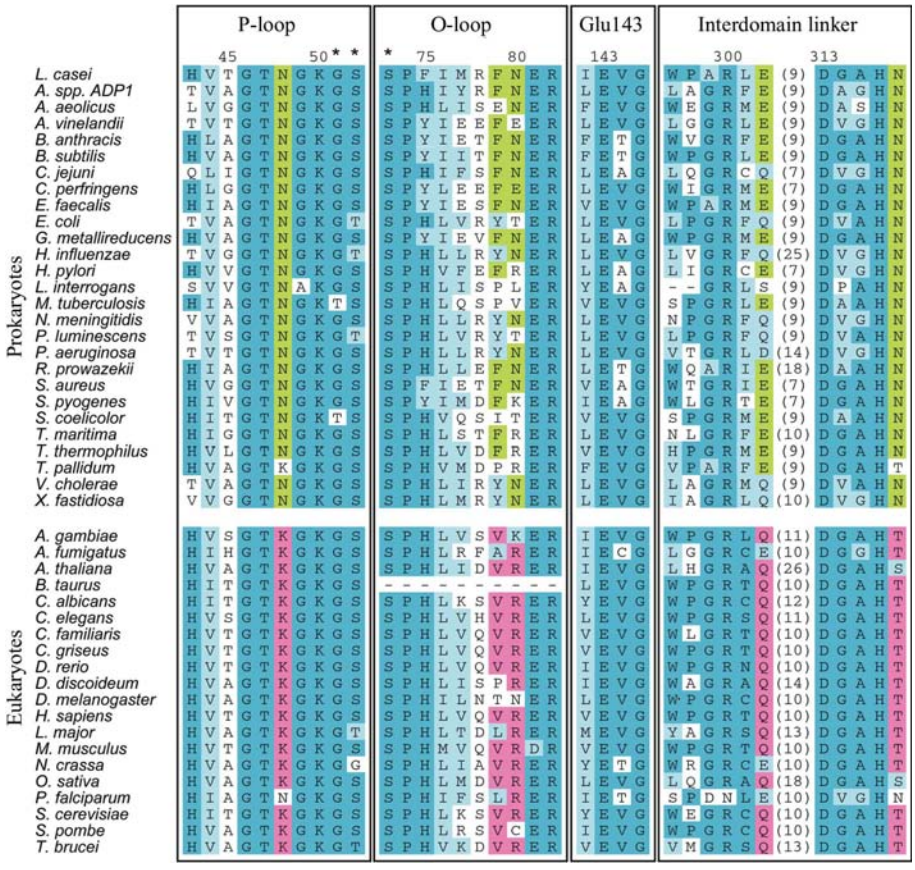

Figure 1
Partial sequence alignment of the known prokaryotic and eukaryotic FPGS enzymes. Four regions have been chosen corresponding to the P-loop, the Ω-loop, the loop following strand B4 (the Glu143 loop) and the linker connecting the N-terminal and C-terminal domains. Residues conserved across the prokaryotes and eukaryotes are shaded dark blue, while residues with a lower degree of conservation across both kingdoms are shaded in light blue. Residues conserved primarily in the prokaryotes are shaded light green and those predominantly observed in the eukaryotes are colored pink. The three residues mutated in this study are indicated with an asterisk. The numbering scheme relates to the *Lactobacillus casei* enzyme. The sequences used in this alignment are, in order, *L. casei*, *Acinetobacter* spp. ADP1, *Aquifex aeolicus*, *Azotobacter vinelandii*, *Bacillus anthracis*, *B. subtilis*, *Campylobacter jejuni*, *Clostridium perfringens*, *Enterococcus faecalis*, *Escherichia coli*, *Geobacter metallireducens*, *Haemophilus influenzae*, *Helicobacter pylori*, *Leptospira interrogans*, *Mycobacterium tuberculosis*, *Neisseria meningitidis*, *Photorhabdus luminescens*, *Pseudomonas aeruginosa*, *Rickettsia prowazekii*, *Staphylococcus aureus*, *Streptococcus pyogenes*, *Streptomyces coelicolor*, *Thermotoga maritima*, *Thermus thermophilus*, *Trepanoma pallidum*, *Vibrio cholerae*, *Xylella fastidiosa*, *Anopheles gambiae* (mosquito), *Aspergillis fumigatus*, *Arabidopsis thaliana*, *Bos taurus* (cow), *Candida albicans*, *Caenorhabditis elegans*, *Canis familaris* (dog), *Cricetulus griseus* (chinese hamster), *Danio rerio* (zebrafish), *Dictyostelium discoideum* (slime mold), *Drosophila melanogaster*, *Homo sapiens*, *Leishmania major*, *Mus musculus* (mouse), *Neurospora crassa*, *Oryza sativa*, *Plasmodium falciparum*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Trypanosoma brucei*.

It was originally proposed that the loop that contained Ser73 (the Ω-loop) was involved in shaping the active site and although its role was not initially understood, it was likely that the unusual structure of the loop would be retained in all FPGS enzymes (Sun *et al.*, 1998). An Ω-loop, although not as well defined as other loops or turns such as the β-turn or the γ-turn, is a common feature in most protein structures and generally links two adjacent secondary-structure elements such that the shape of the loop resembles the Greek letter Ω. The active-site Ω-loop in FPGS comprises ten residues and joins two antiparallel β-strands (*B*2 and *B*3). The X-ray structure of the folate complex of *L. casei* FPGS showed that this loop was involved in binding the 5,10-mTHF

Table 1
Data-collection and refinement statistics.

Values in parentheses relate to the highest resolution shell: 2.48–2.4 Å for apoFPGS, 1.90–1.85 Å for G51S, 2.38–2.30 Å for G51S/S52T and 1.97–1.90 Å for S73A.

	apoFPGS	G51S	G51S/S52T	S73A
Data collection				
Temperature (K)	293	100	110	100
Maximum resolution (d_{min}) (Å)	2.4	1.85	2.3	1.9
Observed reflections	52017	273891	84254	165321
Unique reflections to d_{min}	13391	31920	16532	29157
R_{merge}† (%)	9.6 (37.5)	7.3 (39.0)	10.0 (37.3)	6.0 (29.2)
$I/\sigma(I)$	7.3 (2.1)	18.9 (1.9)	10.1 (2.9)	21.0 (3.8)
Completeness (%)	84.2 (88.4)	95.0 (72.9)	93.6 (95.0)	96.1 (93.1)
Unit-cell parameters				
a (Å)	54.1	53.2	53.6	53.1
b (Å)	45.9	45.8	45.3	45.9
c (Å)	85.4	84.3	84.8	84.3
β (°)	107.9	107.7	107.9	107.0
Refinement				
Resolution range (Å)	20.0–2.4	20.0–1.85	20.0–2.3	20.0–1.9
R_{work}/R_{free}‡ (%)	17.4/24.8§	20.4/24.7¶	19.9/25.9§	19.1/23.5¶
Total atoms: protein/solvent	3128/201	3097/255	3108/281	3133/274
R.m.s. deviation from ideality				
Bonds (Å)	0.006	0.006	0.006	0.006
Angles (°)	1.20	1.28	1.26	1.37

† $R_{merge} = \sum |I - \langle I \rangle| / \sum I \times 100$, where I is the the observed intensity and $\langle I \rangle$ is the mean intensity. ‡ $R = \sum ||F_o| - k|F_c|| / \sum |F_o| \times 100$. R_{free} was calculated with 5% of the reflections. § *CNS* refinement. ¶ *REFMAC* refinement.

substrate (Sun *et al.*, 2001). An interesting feature of this Ω-loop is the presence of a *cis*-proline at its N-terminus (Pro74) and it is this *cis* configuration that directs the carbonyl of Ser73 towards the 'classical' Mg^{2+}. This Ω-loop is highly conserved in all FPGS sequences known to date (Fig. 1), with the *cis*-Pro motif forming a sharp turn at the N-terminus of the loop, while the conserved acidic residue near the C-terminal end is involved in stabilizing interactions with the backbone of the loop.

In order to fully understand the relevance of the P-loop and the Ω-loop to enzyme activity, we have undertaken mutagenesis studies on residues in these two loops (Toy & Bognar, 1994; Sheng *et al.*, 2000). Here, we report the high-resolution structures of the substrate-free form of *L. casei* FPGS (apoFPGS) and three site-directed mutant proteins: G51S and G51S/S52T both in the P-loop and S73A from the Ω-loop.

2. Experimental procedures

2.1. Protein purification

The expression and purification of apoFPGS has been described previously (Cody *et al.*, 1992). The G51S/S52T double mutant was prepared, expressed and purified as described by Toy & Bognar (1994). For expression and purification of G51S and S73A, PCR mutagenesis of the *L. casei* FPGS was performed directly in the expression plasmid pCYB1 (New England Biolabs) and the plasmids were then introduced into *E. coli* strain BL21. The mutant proteins were expressed as a fusion product with an intein-linked chitin-binding domain and purified with the IMPACT one-step system (New England Biolabs). The cultures were induced with 1 m*M* IPTG at 303 K for 6 h. The cells were then lysed by sonication or using a French press and the crude extracts were loaded onto a chitin column and washed with a buffer solution containing 10 m*M* Tris–HCl pH 8.0, 500 m*M* NaCl and 0.1 m*M* EDTA. On-column cleavage of the mutant protein was achieved by quickly flushing the column with 50 m*M* DTT in 10 m*M* Tris–HCl, 50 m*M* NaCl, 0.1 m*M* EDTA and leaving this to stand for 24 h at room temperature. The DDT induced the self-cleavage activity of the intein and the mutant proteins were eluted with cleavage buffer, followed by dialysis against 10 m*M* Tris–HCl pH 7.5, 200 m*M* KCl. In all cases the mutant proteins were dialyzed against 50 m*M* HEPES pH 7.5, 20 m*M* KCl, 100 m*M* $(NH_4)_2SO_4$ prior to crystallization.

2.2. Crystallization

The crystals of the G51S and S73A mutant proteins were grown by microseeding using small crystals initially grown by the hanging-drop method from 50 m*M* acetate buffer pH 5.3 and 6% PEG 4000. Equal volumes of the enzyme and a solution containing 50 m*M* acetate pH 5.3 and 16–24% PEG 4000 were mixed and left to incubate at 277 K for 4 h. Precipitated protein was removed by centrifugation and the enzyme/PEG solution placed in 20 µl microbridges. The supersaturated solution was then streak-seeded with a fine whisker and left at 291 K. Crystals generally appeared overnight and grew to full size within 3–5 d.

Crystals of apoFPGS and the G51S/S52T mutant protein were grown using a procedure similar to that described for wild-type *L. casei* FPGS (Sun *et al.*, 1998). The apoFPGS crystals were transferred to 50 m*M* HEPES buffer pH 7.5 containing 5% PEG 5000 and soaked overnight prior to mounting in a thin-walled quartz capillary for data collection, while the G51S/S52T crystals were transferred into cacodylate buffer pH 8.5 containing 10% PEG 4000 and 25% ethylene glycol just prior to flash-freezing.

2.3. X-ray data collection

X-ray diffraction data were collected from a single frozen G51S crystal using a MAR 345 image-plate detector mounted on a Rigaku RU-H3R generator running at 5 kW. A total of 220 images were collected with an oscillation range of 1° and processed using the *HKL* program package (Otwinowski & Minor, 1997). The final data set containing 31 920 reflections had an R_{merge} of 0.073 to 1.85 Å resolution. For the S73A mutant, 165 images were collected at 110 K and scaled to give a final data set containing 29 157 reflections with an R_{merge} of 0.060. For the G51S/S52T mutant, data were collected from a frozen crystal using an R-AXIS IIc image-plate detector mounted on a Rigaku RU-200 generator. A total of 115 images were collected and processed using the *HKL* program package (Otwinowski & Minor, 1997), giving a final data set containing 16 532 reflections with an R_{merge} of 0.10. Diffraction data were collected from a single apoFPGS crystal at room temperature using a Rigaku R-AXIS IIc image-plate detector mounted on an RU-200 rotating-anode source. 75 images were

collected using an oscillation range of 1.5°, with data extending to approximately 2.3 Å resolution. The images were processed with the *R-AXIS* software and scaled using the *CCP*4 suite of programs (Collaborative Computational Project, Number 4, 1994). Additional statistics for all four data sets are given in Table 1.

2.4. Structure solution and refinement

All four crystals were essentially isomorphous with the wild-type FPGS–MgATP crystals (unit-cell parameters $a = 54.0$, $b = 46.1$, $c = 84.9$ Å, $\beta = 107.3°$). In the case of the three frozen mutant FPGS crystals, there was a slight decrease in all three unit-cell parameters (the average decrease was about 0.6 Å, with an average decrease in unit-cell volume of just under 3%). Initial *R* factors calculated between the wild-type FPGS atomic coordinates and the three mutant protein data sets gave values ranging from 43.4 to 49.9%. While this was most probably a consequence of the unit-cell shrinkage and a translation or rotation of the molecule in the unit cell, a change in the relative orientations of the N-terminal and C-terminal domains arising from the mutations could not be discounted. Therefore, the molecules were positioned in the asymmetric unit by molecular replacement using the program *AMoRe* (Navaza, 1994). The structures of the mutants were subsequently refined with either *CNS* (Brünger *et al.*, 1998) or *REFMAC* (Murshudov *et al.*, 1999) and relevant refinement statistics are given in Table 1.

Molecular replacement was also used to position the apoFPGS molecule in the asymmetric unit using the wild-type MgATP complex (FPGS–MgATP) as the starting model, with the Mg^{2+}, pyrophosphate and all water molecules removed. The correlation coefficient and *R* factor following rigid-body fitting were 79.7 and 26.3%, respectively, implying an extremely good fit of the model to the diffraction data. The apoFPGS structure was initially refined using the least-squares procedures implemented in *TNT* (Tronrud *et al.*, 1987), where the initial *R* factor was 27.9% ($R_{free} = 35.5\%$) for data from 6.0 to 2.5 Å resolution. Refinement was completed with *CNS* (Brünger *et al.*, 1998).

The atomic coordinates and the structure factors for the apoFPGS, G51S mutant, G51S/G52T double mutant and S73A mutant structures were deposited in the Protein Data Bank (Berman *et al.*, 2000) with PDB codes 2gca, 2gc5, 2gca and 2gc6, respectively. All superpositions were performed using *LSQKAB* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). Figures were generated using *MOLSCRIPT* (Kraulis, 1991), *BOBSCRIPT* (Esnouf, 1997) and *RASTER3D* (Merritt & Murphy, 1994).

3. Results

3.1. The apoFPGS structure

The apoFPGS structure was refined to a crystallographic *R* factor of 0.174 ($R_{free} = 0.248$) using data between 20.0 and 2.4 Å resolution. The final model (Fig. 2) comprises residues 1–343, 348–374 and 386–425, along with 201 water molecules, and has the same overall structure as observed for the wild-type FPGS–MgATP (Sun *et al.*, 1998). The molecule is divided into two domains, the N-terminal domain (residues 1–295) and the C-terminal domain (301–425), joined by an extended six-residue linker peptide (296–300). The N-terminal domain is composed of a central seven-stranded β-sheet flanked by seven α-helices, with an additional small three-stranded β-sheet packing against the side of the domain. The C-terminal domain has an α/β fold resembling dihydrofolate reductase (DHFR): a central six-stranded β-sheet flanked by four α-helices. The active site of the enzyme is located in a large cleft between the two domains.

Superposition of the whole apoFPGS model onto wild-type FPGS–MgATP (Sun *et al.*, 1998) gives an r.m.s. deviation of 0.53 Å for 392 C^{α} positions. Inspection of the two superimposed models on a graphics workstation indicated that the two structures overlap almost exactly. There is clearly no relative movement of the two domains when the MgATP is bound; the r.m.s. deviations for the individual N-terminal and C-terminal domains of 0.52 Å (285 atoms) and 0.43 Å (107 atoms) are virtually identical to the overall r.m.s. deviation.

In the FPGS–MgATP complex, five short pieces of polypeptide were omitted from the model as they had poorly defined electron density. Three of these breaks were in the N-terminal domain (18–19, 146–150 and 170–176) and two in the C-terminal domain (342–347 and 373–284). In the apoFPGS structure, the three missing regions in the N-terminal domain are located in well ordered electron density. Fig. 3(*a*) shows the $2F_o - F_c$ electron density associated with one of these loops: Ile146–Thr150 (between strands *B*4 and *B*5). This loop appears to be

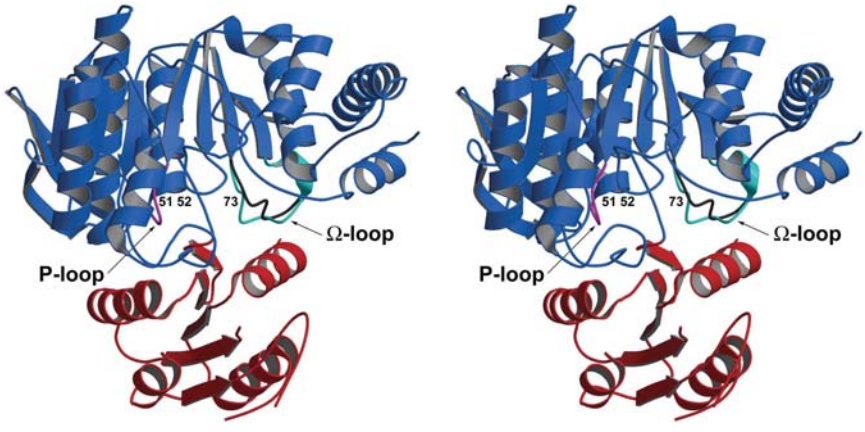

Figure 2
Ribbon representation of the *L. casei* G51S FPGS mutant. The N-terminal domain is shown in blue, while the C-terminal domain is in red. The location of the P-loop (magenta) is indicated and the positions of the three mutations (G51S, S52T and S73A) are given. The Ω-loop conformation seen in G51S is shown in cyan, overlaid with the loop conformation observed in the wild-type FPGS–MgATP structure, shown in black.

stabilized by two new interactions with the side chain of the conserved lysine residue in the neighboring P-loop (Lys50). In FPGS–MgATP, the N^{ζ} atom of Lys50 is oriented towards the phosphate groups of the ATP, adopting a conformation similar to that observed in the nucleotide-bound forms of the G proteins (Pai *et al.*, 1990; Coleman *et al.*, 1994) and myosin (Smith & Rayment, 1996*b*). In the apoFPGS structure, however, the N^{ζ} atom points in the opposite direction and hydrogen bonds to the carbonyl O atoms of Val144 and Gly145 (Fig. 3*a*). Additional interactions with the side chain of Thr45 and several solvent molecules also help to stabilize the loop.

The active site of apoFPGS is similar to that seen in the FPGS–MgATP complex, despite there being no ATP or Mg^{2+} bound. Moreover, the conformation of the Ω-loop is identical, with a *cis*-proline at position 74 (Sun *et al.*, 1998). The loop is stabilized by several external interactions: a salt bridge between Arg78 and Asp92, hydrogen-bonding interactions between the main-chain amide N atom of Ser73 and the side chain of Glu143, the carbonyl O atom of Ile76 and the side chain of Gln421, the side chain of Asn80 and the carbonyl O atom of Ile90 and numerous water-mediated hydrogen bonds. Additional stabilization of the loop comes from three internal hydrogen bonds between the side chain of Glu81, which reaches across the loop and interacts with the main-chain amide N atoms of Ile76, Met77 and Arg88 (Fig. 3*b*). Superposition of the Ω-loop in apoFPGS with the equivalent loop in the FPGS–MgATP complex gives an r.m.s. deviation in main-chain atom positions between the two loops of 0.35 Å (for residues 72–81; 40 atoms), indicating that the structure of this loop is not dependent upon the presence of an Mg^{2+} ion in the cation-binding pocket. In the crystal structures of the two other FPGS enzymes whose structures have recently been determined, *T. maritima* (PDB code 1o5z) and *E. coli* (Mathieu *et al.*, 2005), there is also a *cis*-peptide at this position and the Ω-loops in these enzymes are essentially identical in conformation (Fig. 3*b*). Superposition of the Ω-loops from the wild-type *L. casei*, *T. maritima* and *E. coli* enzymes gives an r.m.s. deviation of between 0.31 and 0.38 Å for 44 main-chain atoms). In the *T. maritima* and *E. coli* structures, the conserved glutamate residue (Glu82 and Glu91, respectively) is in an identical location (Fig. 3*b*) and again stabilizes the loop through three hydrogen-bonding interactions with main-chain amide N atoms.

An examination of the region surrounding the 'classical' Mg^{2+}-binding site in apoFPGS shows that the Glu143 residue is in a similar position as in the FPGS–MgATP structure, although the carboxylate group has rotated by approximately 20°. The two hydrogen bonds between the Glu143 side chain and the Ser73 main chain are retained and the Mg^{2+} ion has been replaced by a water molecule. This water molecule is hydrogen bonded to both the $O^{\varepsilon 1}$ and $O^{\varepsilon 2}$ atoms of Glu143 (2.64 and 2.54 Å, respectively) and to the carbonyl O atom of Ser73 (2.57 Å). This hydrogen-bonding network could mimic the coordination sphere of the Mg^{2+} ion and help retain the Ω-loop in its native conformation.

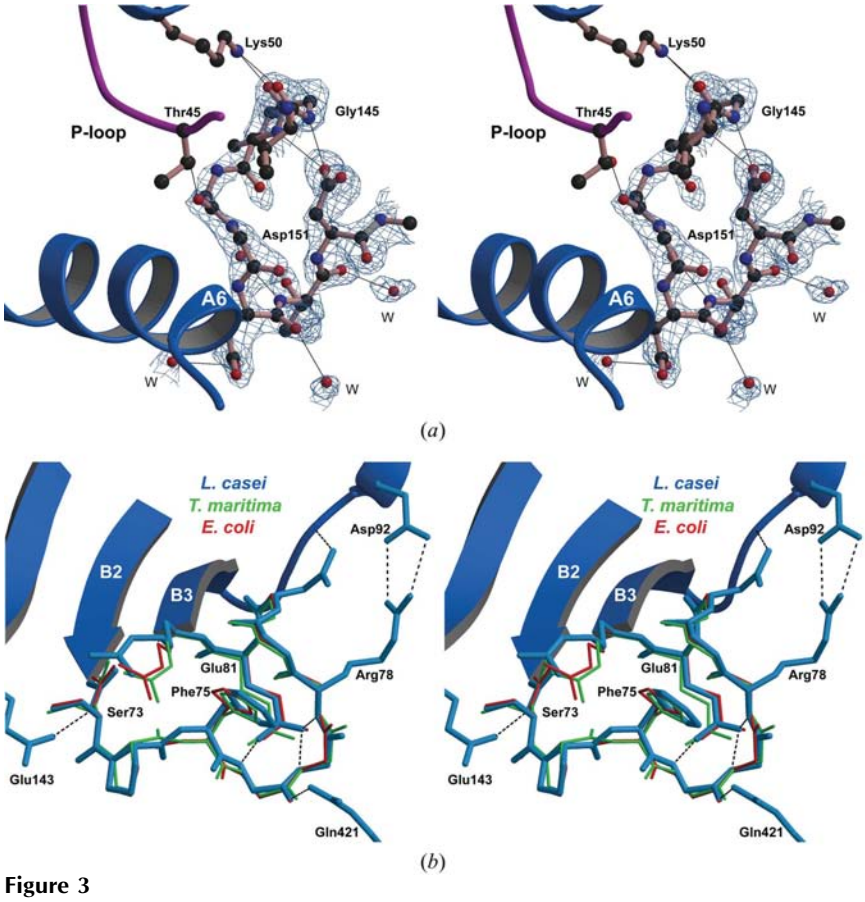

Figure 3
(*a*) Stereoview of the $2F_o - F_c$ electron density (contour level 1.1σ) associated with the 146–150 loop in the G51S mutant structure. The P-loop (magenta) and a small part of helix *A*3 are shown at the top of the figure and helix *A*6 is to the lower left of the 146–150 loop. Some of the residues involved in stabilizing the loop through hydrogen-bonding interactions are indicated. The location of three water molecules which make hydrogen-bonding interactions with the loop are also shown, labeled W. (*b*) Stereoview of the superposition of the Ω-loop from *L. casei* apoFPGS (blue bonds), *T. maritima* FPGS (green bonds) and *E. coli* FPGS (red bonds). The hydrogen-bonding interactions of the conserved glutamate (Glu81 in *L. casei*) with the main-chain amide N atoms of residues Ile76, Met77 and Arg78 are indicated as dashed black lines in the center of the loop. Additional external hydrogen-bonding interactions with Glu143, Asp92 and Gln421 are also shown (see text). The location of the conserved aromatic residue responsible for folate binding in the *L. casei* enzyme (Phe75) is also indicated; the corresponding residue in both the *E. coli* and *T. maritima* structures is a histidine.

3.2. The G51S structure

The structure of the G51S mutant protein was refined at 1.85 Å resolution, giving a final crystallographic *R* factor of 0.204 and an $R_{\rm free}$ of 0.247 using 5% of the reflections randomly selected prior to refinement. The resultant protein model has good geometry,

with no residues in unfavored conformations as determined by a Ramachandran plot (data not shown). Additional refinement statistics are given in Table 1. The final model comprises residues 1–18, 21–342, 248–375 and 386–425, along with 213 solvent molecules and two sulfate ions.

Superposition of the G51S mutant onto the wild-type FPGS–MgATP structure gives an r.m.s. difference of 0.68 Å for all 390 C^{α} positions in the G51S model. When the N- and C-terminal domains are superimposed separately, the r.m.s. differences are 0.64 Å (283 C^{α} atoms) and 0.58 Å (107 C^{α} atoms), respectively. With the N-terminal domain of G51S superimposed onto the N-terminal domain of the wild-type structure, a subsequent superposition of the C-terminal domains gives a value for the relative domain movement of less than 1.5°. This indicates that there is very little difference in the relative orientation of the two domains.

During the course of the refinement and model rebuilding, special attention was paid to the parts of the structure surrounding the active site. It was noted early in the refinement that there was extra electron density associated with residue 51 (at the N-terminus of helix *A*3) consistent with a serine residue in two alternate rotamer conformations, one major (about 80% based upon occupancy refinement) and the other minor. Additional electron density in the active site was modeled as a sulfate anion bound in the oxyanion hole formed by the P-loop and the N-terminus of helix A3. The initial $F_o - F_c$ difference map, calculated after the first cycle of *CNS* refinement, is shown in Fig. 4(*a*). It was also noted in these early stages that the structure of the adjacent Ω-loop (Ser73–Arg82) did not fit the available $2F_o - F_c$ electron density. This loop (72–82) was removed prior to the next cycle of refinement and an $F_o - F_c$ omit map calculated (Fig. 4*b*). The map clearly showed that the conformation of the Ω-loop had altered between residues Ser73 and Met77. The peptide bond between Ser73 and Pro74 has changed to a *trans* conformation and the loop between residues 73 and 78 follows a different path to its counterpart in the wild-type structure (see Fig. 2).

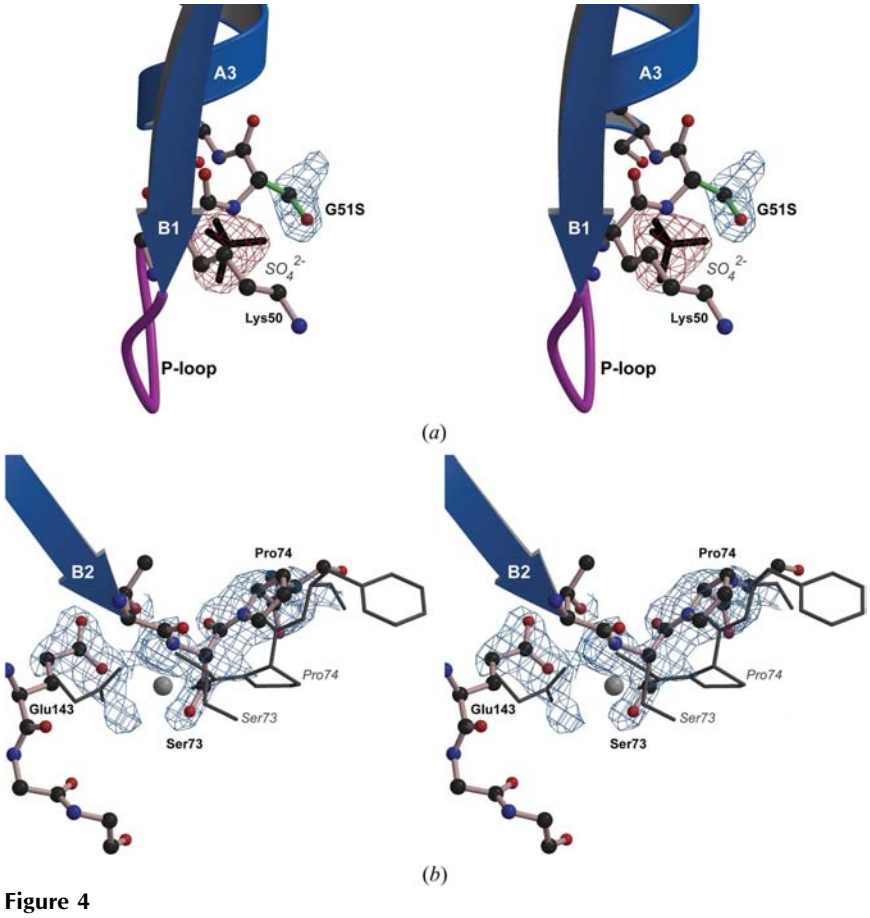

Figure 4
(*a*) Stereoview of the site of the Gly-Ser mutation in the G51S FPGS mutant. The initial $F_o - F_c$ difference map in blue (contour level 2.5σ) shows the location of the two additional atoms in the serine side chain. There are clearly two possible rotamers and the model is shown (green) in the major conformation. The red $F_o - F_c$ density (contour level 2.5σ) shows the location of a sulfate anion (black bonds) bound in the oxyanion hole at the N-terminus of helix *A*3, associating with the backbone atoms of residues Gly49, Lys50 and Ser51. (*b*) Stereoview of the $F_o - F_c$ omit map (blue, contour level 2.0σ) calculated after removal of the Ω-loop (residues Ser73–Arg81) and Glu143. The positions of residues Ser73–Phe75 and Glu143 are shown in gray bonds in the positions they occupy in the wild-type FPGS–MgATP structure (labeled in italics), while the locations these residues adopt in the G51S mutant are shown in ball-and-stick representation (labeled in bold). The position of the 'classical' Mg^{2+} ion in FPGS–MgATP is shown as a light gray sphere in the center of the figure. The G51S mutation is omitted for clarity.

As a result of this *cis* to *trans* conformational change at Pro74, the carbonyl O atom of Ser73 no longer points towards the Mg^{2+}-binding site but rather has shifted away by just over 3 Å (Fig. 5*a*). The Ser73 side chain now occupies roughly the same position as the carbonyl O atom in the wild-type structure, although the O^{γ} atom also points away from the active site and makes only one hydrogen-bonding interaction with a water molecule (Fig. 5*a*). The Ser73 side chain was originally involved in water-mediated hydrogen bonds to the side chain of Arg82 (data not shown), but in this mutant those hydrogen bonds are lost. Another important result of the *cis–trans* conformational change is a major rearrangement of the main chain from Pro74 to residue Met77 (Fig. 5*a*). The side chains of Phe75 and Met77 have both flipped through almost 180°, with the phenyl ring of Phe75 moving about 12 Å and now projecting into the active site. The main-chain amide N atoms of Ile76 and Met77 have moved between 2 and 3 Å away from their original positions and the hydrogen-bonding interactions with Glu81 which served to stabilize the loop in the wild-type enzyme are lost. In the mutant structure, the glutamate side chain now projects away from the Ω-loop (Fig. 5*a*) into the solvent and is hydrogen bonded to three water molecules.

The presence of the bulky phenyl ring of Phe75 in the active site leads to some additional structural rearrangements. The Phe75 side chain now occupies a position previously occupied by an arginine side chain (Arg300) in the wild-type enzyme (Fig. 5*a*). The side chain of Arg300 has moved to a position formerly occupied by a tryptophan side chain (Trp297), while the tryptophan side chain rotates through about 75° and causes some minor rearrangements of the main chain of residues 294, 295 and 296 (Fig. 5*b*). All of these residues are in the highly conserved seven-residue interdomain linker peptide (see Fig. 1). However, as noted earlier, there has been essentially no movement of the C-terminal domain relative to the N-terminal domain and the residues at both ends of this linker are still in structural coincidence in the G51S mutant and the wild-type structures, such that this conformational change is essentially localized to these residues (Fig. 5*b*).

In addition to the more obvious structural changes involving the Ω-loop and the interdomain linker, the new serine side chain at residue 51 also causes the side chain of Glu143 to move approximately 1.5 Å from its original location. This turns out to be a simple rotation of approximately 90° about the $C^{\alpha}-C^{\beta}$ bond, which shifts one of the carboxylate O atoms ($O^{\varepsilon 2}$) more than 4 Å from its wild-type position and leaves the $O^{\varepsilon 1}$ atom in roughly the same place (Fig. 5*a*). This new position of the glutamate side chain is partially stabilized by the formation of a new hydrogen bond between the $O^{\varepsilon 2}$ atom and the $O^{\gamma 1}$ atom of Thr72, which has to adopt a new rotamer conformation in order to make this interaction (Fig. 5*a*). The carboxylate group of Glu143 is now no longer in the correct position to interact with an Mg^{2+} ion and hydrogen bond to the main chain of Ser73, as it does in the wild-type enzyme.

3.3. The G51S/S52T structure

The double-mutant protein was one of the *L. casei* P-loop mutant enzymes generated prior to the determination of the FPGS–MgATP structure in an effort to determine whether this region was essential for enzyme function (Toy & Bognar, 1994). The structure of this mutant protein was determined to 2.3 Å resolution and shows a similar series of conformational changes in the Ω-loop and the interdomain linker as observed in the G51S mutant. The S52T mutation does not appear to affect the conformational changes occurring in the Mg^{2+}-binding site, since this residue is located on the other side of helix *A*3, away from the active site. The additional C atom in the threonine side chain is readily accommodated in a large internal pocket and causes only a minor rearrangement of the side chain of Gln261. In the wild-type structures, the $O^{\varepsilon 2}$ atom of Gln261 accepts a hydrogen bond from the O^{γ} of Ser52. This hydrogen-bonding interaction is maintained in G51S/S52T despite the small movement of the glutamine side chain.

The serine residue at position 51 adopts the same major rotamer conformation as in the single G51S mutant, pointing towards the Ω-loop but making no interactions with the main chain. Once again, Pro74 has switched to a *trans* conformation and the resulting conformation of the loop is identical to that seen in the G51S mutant; the r.m.s. deviation for 40 main-chain atoms between the G51S/S52T and G51S structures is 0.45 Å. Furthermore, the same structural rearrangements in the seven-residue interdomain linker peptide are seen in this double mutant. In addition, the side chain of Glu143 is

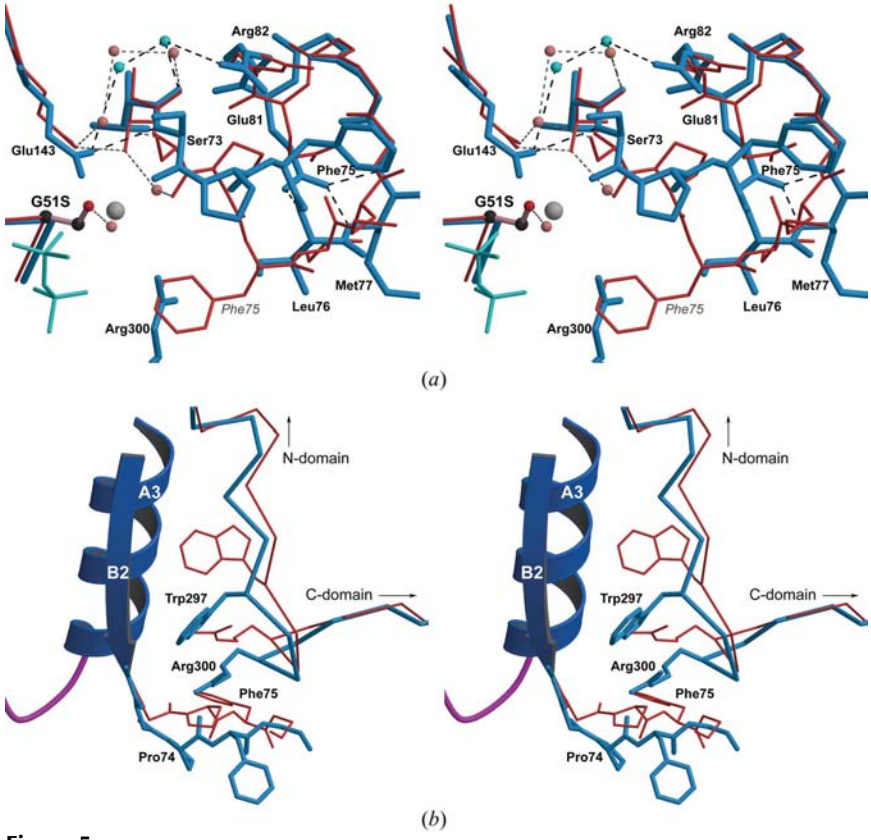

Figure 5
(*a*) Stereoview of the altered 'classical' Mg^{2+}-binding site and the change in Ω-loop conformation in the G51S mutant. Residues shown as light blue bonds are for the wild-type FPGS–MgATP structure and those shown as thinner red bonds represent the mutant structure. The serine side chain at position 51 is shown in ball-and-stick representation at the lower left. Water molecules which make critical hydrogen-bonding contacts in the vicinity of the Mg^{2+} site are indicated in cyan for the wild-type structure and pink for the G51S mutant structure. The location of the Mg^{2+} and pyrophosphate in the wild-type structure are shown at the lower left as a light gray sphere and thin cyan bonds, respectively. The large-scale movement of the Phe75 side chain from its position in the wild-type enzyme (center right) to the position in the mutant (bottom center) can be clearly seen. (*b*) Stereoview of the active site of the G51S mutant, showing the sequential movement of the side chains of Arg300 and Trp297 (center) in the interdomain linker as a result of the conformational change in the Ω-loop (bottom). The wild-type structure is shown as light blue bonds, with part of the P-loop shown in magenta to the left, and the G51S mutant structure is shown in thin red bonds. The location of the N-domain and the C-domain are indicated and it can be seen that the disruption to the interdomain linker caused by the rearrangement of the Ω-loop is localized to only a very small region.

displaced from the position it occupies in the wild-type structures in a similar manner to that in G51S.

3.4. The S73A structure

The S73A mutant protein was generated to test whether the residue at position 73 was important in determining the Ω-loop conformation. Kinetic characterization has shown that of the three mutants in the current study only S73A shows any degree of enzyme activity, estimated at around 5% (Sheng *et al.*, 2000). The structure was solved and refined at high resolution (1.9 Å) to a final crystallographic R factor of 0.191 and an $R_{\rm free}$ of 0.235. The final model consists of 407 residues (1–17, 21–343, 348–374 and 386–425), 274 solvent molecules and one sulfate anion. Once again, the overall conformation of the enzyme remains unchanged following mutagenesis. The r.m.s. deviation between the S73A mutant and the MgATP–FPGS structure (for 389 matching C^{α} atoms) is 0.47 Å and there is no evidence for any relative domain movement. The mutation of Ser73 has no apparent effect on the main-chain conformation of the Ω-loop; the r.m.s. deviation in the positions of the 44 main-chain atoms from Thr72 to Arg82 is only 0.21 Å.

The water-mediated hydrogen-bonding network between the Ser73 side chain and Arg82 has been lost, however, and as a result the Arg82 side chain swings away towards the solvent and makes hydrogen-bonding interactions with two water molecules. The stabilizing hydrogen-bonding interactions made by Glu81 across the loop are intact, along with the external hydrogen bond involving residues Ile76, Arg78 and Asn80. The residues involved in the active site (the P-loop residues and Glu143) are in identical positions to those in the MgATP–FPGS and apoFPGS structures.

4. Discussion

The crystallographic analyses of the G51S and G51S/S52T mutants of *L. casei* FPGS show that the conformation of the Ω-loop plays an important role in the formation of the active site and that a disruption of this loop results in an inactive enzyme. In almost all of the FPGS sequences known to date, this loop is very highly conserved (see Fig. 1) and has the consensus sequence T/S-**SP**-H(Y/F)-*xxxx*-N/R-**ER** (the residues shown in bold are completely conserved in all known FPGS sequences). However, there are some interesting differences when the sequences of the prokaryotes and eukaryotes are compared: in the prokaryotes, the two residues immediately prior to the conserved glutamate are generally a phenylalanine–asparagine pair, whereas in the eukaryotic enzymes a valine–arginine pair predominates (Fig. 1). Furthermore, virtually all eukaryotes have a histidine residue following the conserved proline, while in the prokaryotes this is sometimes replaced by other aromatic residues. In *L. casei*, for example, this is a phenylalanine, which together with Tyr414 from the C-domain forms an aromatic 'clamp' to anchor the pterin ring of the 5,10-mTHF substrate (Sun *et al.*, 2001). In other FPGS enzymes, the aromatic residue equivalent to Phe75 in the Ω-loop may play a similar role. The significance of the other sequence changes in the center of the loop are not yet clearly understood.

In the three bacterial enzymes whose structures are known to date, the conserved glutamate residue at the C-terminus of the Ω-loop plays a key role in maintaining the structural integrity of the loop by hydrogen bonding to the main-chain amide N atoms of residues 76, 77 and 78 (*L. casei* numbering). However, the most critical residue in the Ω-loop is the proline at position 74. In the wild-type structures (the MgATP complex and apoFPGS) this residue adopts a *cis* configuration which orients the carbonyl O atom of the preceding residue (Ser73) towards the P-loop, thus forming part of the 'classical' Mg^{2+}-binding site. Without this conformation at the N-terminus of the loop, the magnesium-binding site would not be formed; this is exactly what happens in the G51S and G51S/S52T mutants.

In these two mutant structures, the peptide between Ser73 and Pro74 has changed to a *trans* conformation such that the Ser73 carbonyl O atom is now no longer directed into the 'classical' Mg^{2+}-binding site. Just what triggers this change in peptide configuration is not entirely clear, but it appears to be related to the presence of a number of stabilizing interactions with the main-chain atoms of Ser73. In the wild-type MgATP complex, the side chain of Glu143 makes hydrogen-bonding interactions with the amide N atom (2.9 Å) and carbonyl O atom (3.2 Å) of Ser73, while the carbonyl O atom is in turn coordinated to the 'classical' Mg^{2+} ion (Sun *et al.*, 1998). The movement of the Glu143 side chain as a result of the G51S mutation, along with the insertion of the serine side chain into the cation-binding pocket, are sufficient to prevent the binding of Mg^{2+} or water in the cation site, leading to the loss of all the stabilizing interactions which were present in the wild-type structures. In apoFPGS, where the conformation of the Ω-loop has remained intact, the interactions with Ser73 are similar to those observed in the wild-type MgATP complex, although a water molecule now replaces the Mg^{2+} ion and is hydrogen bonded to the carbonyl O atom of Ser73. It is interesting to note that in an E143A mutant of *L. casei* FPGS, the structure of the Ω-loop is identical to that observed in the wild-type structures (Sheng *et al.*, 2002), despite the loss of Mg^{2+} binding. In this mutant, the carboxylate group of the glutamate is replaced by two water molecules which make hydrogen-bonding interactions with the carbonyl O atom and the amide N atoms of Ser73, while the Mg^{2+} site appears to be occupied by a third water molecule which is also hydrogen bonded to the Ser73 carbonyl O atom. These are rather intriguing results in that they suggest that it is the presence of these interactions that is the critical factor in the stabilization of the Ω-loop, rather than the identity of the hydrogen-bonding partner or the occupant of the 'classical' cation-binding site.

The G51S mutant was originally engineered in an attempt to produce a P-loop similar to that observed in the G proteins, which have a hydroxyl residue directly following the conserved lysine. The 'classical' P-loop consensus sequence is **G***xxxx***GK**-S/T-S/T, whereas the FPGS P-loop (see Fig. 1) has the general sequence *xx***GT**-N/K-**GKG**-S/T (where the resi-

dues shown in bold are completely conserved in all known FPGS sequences; *M. tuberculosis* and *S. coelicolor* appear to be exceptions, having a threonine directly following the conserved lysine and being more in line with the 'classical' P-loop in this regard). As noted earlier in enzymes with a 'classical' P-loop, the hydroxyl residue immediately following the conserved lysine is one of the ligands for Mg^{2+} coordination. Clearly, incorporating a serine side chain at this position of the FPGS P-loop does not give rise to a Mg^{2+} ligand but rather disrupts the structure of the metal site so profoundly as to completely abolish binding and subsequently enzyme activity.

Two similar mutants (G51T and G51S/S52T) have been studied previously as part of a detailed investigation into the importance of the P-loop in enzyme activity (Toy & Bognar, 1994). It was found that the single G51T mutant had a 22-fold increase in the K_m for glutamate and a ninefold increase for the 5,10-mTHF substrate, while the K_m for ATP was only slightly above that of the wild type. In the case of the double mutant G51S/S52T, the K_m for glutamate showed a 77-fold increase, with a 20-fold increase in the K_m for 5,10-mTHF. While these results implicated Gly51 in glutamate or folate binding, the structures of the MgATP–FPGS complex (Sun *et al.*, 1998) and the AMPPCP- and folate-bound complexes (Sun *et al.*, 2001) clearly implicate this part of the structure in ATP binding. Folate binding is associated primarily the Ω-loop (Sun *et al.*, 2001) and in this case a structural rearrangement of this loop, as observed in the G51S mutants, would appear to be consistent with the 5,10-mTHF kinetic data. The location of the L-glutamate-binding site is not yet known, although structural comparisons between FPGS and the UDP-*N*-acetylmuramoyl:L-alanine:L-glutamate (UMAG) product complex of the bacterial cell-wall ligase MurD (Bertrand *et al.*, 1999) indicated that L-glutamate binding could be associated with a region at the top end of the C-terminal domain, just beyond the linker peptide (Sheng *et al.*, 2000). This region contains a conserved aspartate (Asp313 in *L. casei*) which is involved in hydrogen bonding to the ribose of the ATP and a histidine residue three positions along, followed either by an asparagine in prokaryotes or a threonine in eukaryotes (see Fig. 1). The glutamate moiety of the UMAG product was anchored by interactions with the histidine and two residues from the N-terminus of the final α-helix (which correspond to residues Ser412 and Leu413 in *L. casei*). It is not yet known whether this site also corresponds to the binding site for free L-glutamate, although recent modeling studies using the *L. casei* crystal structure (Tan & Carlson, 2005) also support this region as being involved in binding the incoming free amino acid. If the true binding site were near this region, then the effect that the mutation of Gly51 has on L-glutamate binding is difficult to explain. However, since this putative binding site is near the interdomain linker and very close to the Ω-loop itself, it is conceivable that this site could be affected by the conformational change observed in the Ω-loop and the subsequent structural rearrangements which take place in the linker peptide.

It would appear, based on the structural details of the S73A mutant, that the residue prior to the proline does not play an important role. When changed to an alanine, the Ω-loop still folds in exactly the same way as in the wild type. However, we know that all FPGS sequences to date have a serine at this position and this, along with the observation that the activity of the S73A mutant has decreased to around 5% that of the wild type, implies that something much more subtle is happening. Kinetic analysis of this mutant shows that it has a 20-fold increase in K_m for the folate substrate and an 11-fold decrease in V_{max} (Sheng *et al.*, 2000). Although removal of the hydroxyl group may not effect the structure of the loop, it does alter the hydrogen-bonding network in this region, and because this is adjacent to the magnesium-binding site it may well affect the binding of the essential Mg^{2+} ion and explain the V_{max} decrease.

The increase in folate K_m is somewhat more perplexing, since the binding site for the 5,10-mTHF substrate in the *L. casei* enzyme is associated with the Ω-loop (Sun *et al.*, 2001), which remains intact in the S73A mutant. The recent structure of the *E. coli* FPGS with a phosphorylated dihydropteroate (DHPP) intermediate bound (Mathieu *et al.*, 2005), along with the modeling studies with *L. casei* (Tan & Carlson, 2005), may provide some clues. The *E. coli* enzyme is bifunctional, in that it has both dihydrofolate synthetase (DHFS) activity in which the first glutamate residue is added to dihydropteroate to produce dihydrofolate and also FPGS activity in which the polyglutamate tail is sequentially added. The DHPP intermediate is bound somewhat differently than the 5,10-mTHF in the *L. casei* structure. The pterin ring is in a

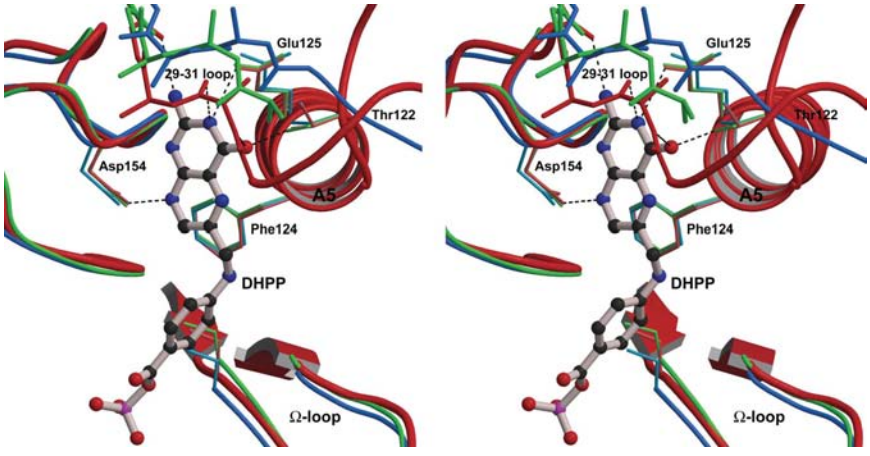

Figure 6
Superposition of the pterin-binding pocket in the *E. coli* bifunctional DHFS/FPGS (red) onto the corresponding putative binding pockets in the *L. casei* (blue) and *T. maritima* (green) structures. The phosphorylated dihydropteroate observed in the *E. coli* structure is shown in ball-and-stick representation. The side chains of the conserved residues associated with pterin binding (see text) are shown as thin bonds in the same color scheme. The loop between helices *A*1 and *A*2 (the 29–31 loop in *E. coli*) is at the top of the figure, shown as thin bonds in the same colors for the three enzymes.

pocket flanked by the loop between helices *A*1 and *A*2 (secondary-structure nomenclature as described by Sun *et al.*, 1998), the N-terminus of helix *A*5 (residues Thr122, Phe124 and Glu125, *E. coli* numbering) and a conserved aspartate (Asp154) from the loop connecting strands *B*4 and *B*5. The *p*-aminobenzoic acid (*p*ABA) moiety passes across the N-terminal end of the Ω-loop, coming close to the position of Ser73, and projects the phosphate group into the active site. Superposition of the three bacterial structures shows that this binding pocket is present in all three enzymes (Fig. 6), although the loop between the first two α-helices is disordered in the *T. maritima* structure. The residues specifically involved in binding the pterin moiety in *E. coli*, Thr122, Phe124, Glu125 and Asp154, are very highly conserved in all FPGS sequences, the first three residues being part of a T/S-*x***FE** consensus sequence. It is highly likely that all FPGS enzymes will have a pterin-binding pocket in this location.

One likely scenario might be that there are two pterin-binding sites in the DHFS/FPGS enzymes. The first is a site such as that observed in *E. coli*, which is used for the addition of the first glutamate residue but not for the addition of the subsequent polyglutamate tail. This cavity seems very tight compared with the putative pocket in *L. casei* and *T. maritima* (Fig. 6). The loop connecting helices *A*1 and *A*2 is approximately 4 Å closer to the active site in *E. coli* and shortens the cavity considerably in such a way that it may only allow for the binding of dihydropteroate. It seems unlikely that the pterin and *p*ABA moieties could move away from the active site following the formation of DHF to allow subsequent additions of L-glutamate. The second site could be similar to the binding site observed for 5,10-mTHF in the *L. casei* enzyme, which appears to be a somewhat more open and possibly more flexible site. This would be optimal for the polyglutamylation reaction in that it could allow for the looping out of the growing polyglutamate tail and it has been suggested that the site associated with the Ω-loop could possibly become more important as the polyglutamate chain length increases (Tan & Carlson, 2005).

In the specific case of the *L. casei* enzyme, in its current position the distance from the end of the *p*ABA ring of the 5,10-mTHF molecule to the γ-phosphate of the ATP is about 11 Å, roughly equivalent to two glutamate residues. If a monoglutamate 5,10-mTHF were bound here, it would not be possible for the terminal carboxylate to reach into the active site and interact with the ATP and modeling of a monoglutamate 5,10-mTHF in this site supports this (Tan & Carlson, 2005). It was initially suggested that a conformational change in the Ω-loop could possibly bring the monoglutamate 5,10-mTHF closer to the active site for phosphorylation (Sun *et al.*, 2001), but a much more plausible explanation is that the *L. casei* enzyme has a second binding site similar to that observed in *E. coli*, where the monoglutamate 5,10-mTHF may be bound for the addition of the second glutamate. When monoglutamate 5,10-mTHF is modeled into this second site, the α-carboxylate group could potentially form hydrogen bonds with Ser73 and Arg82. Furthermore, when the di- and triglutamate derivatives are modeled, these two residues are also implicated in hydrogen bonding to the terminal α-carboxylate group of the polyglutamate tail (Tan & Carlson, 2005). If the *L. casei* enzyme does have such a site in the same location as the DHPP site in *E. coli*, then it is conceivable that the Ser73 residue (along with Arg82) may play an important role in orienting the terminal γ-carboxylate group in the active site and this might explain the 20-fold increase in folate K_m observed for the S73A mutant. Clarification of the presence of two binding sites must await further substrate-binding, mutagenesis and modeling studies.

5. Conclusions

The conformation of the Ω-loop adjacent to the active site is critical for FPGS activity in that it not only provides one of the ligands for the 'classical' mononucleotide magnesium ion, but also forms part of the folate-binding pocket. A major disruption of the interactions which hold the loop in its wild-type conformation directly affect Mg^{2+} and folate binding and has an indirect yet profound effect on the binding of the incoming L-glutamate substrate. There is a fine balance between the two configurations of the Ser73-Pro74 peptide bond and these results give a clear illustration of the significance of hydrogen-bonding interactions in enzyme structure and function.

We thank Steve Shewry for assistance with protein purification. The support of Ted Baker is also gratefully acknowledged. This work was supported by grants from the Auckland Medical Research Foundation and the New Zealand Health Research Council to CAS.

References

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
Bertrand, J. A., Auger, G., Martin, L., Fanchon, E., Blanot, D., Le Beller, D., van Heijenoort, J. & Dideberg, O. (1999). *J. Mol. Biol.* **289**, 579–590.
Bognar, A. L. & Shane, B. (1983). *J. Biol. Chem.* **258**, 12574–12581.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Cichowicz, D. J. & Shane, B. (1987). *Biochemistry*, **26**, 513–521.
Cody, V., Luft, J. R., Pangborn, W., Toy, J. & Bognar, A. L. (1992). *J. Mol. Biol.* **224**, 1179–1180.
Coleman, D. E., Berghuis, A. M., Lee, E., Linder, M. E., Gilman, A. G. & Sprang, S. R. (1994). *Science*, **265**, 1405–1412.
Collaborative Computional Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Esnouf, R. M. (1997). *J. Mol. Graph. Model.* **15**, 132–134.
Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
Mathieu, M., Debousker, G., Vincent, S., Viviani, F., Bamas-Jacques, N. & Mikol, V. (2005). *J. Biol. Chem.* **280**, 18916–18922.
McGuire, J. J. & Coward, J. K. (1984). *Folates and Pterins*, edited by R. L. Blakley & S. J. Benkovic, pp. 135–190. New York: Wiley Interscience.
Merritt, E. A. & Murphy, M. E. P. (1994). *Acta Cryst.* D**50**, 869–873.

Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S. & Dodson, E. J. (1999). *Acta Cryst.* D**55**, 247–255.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pai, E. F., Krengel, U., Petsko, G. A., Goody, R. S., Kabsch, W. & Wittinghofer, A. (1990). *EMBO J.* **9**, 2351–2359.
Schulz, G. E. (1993). *Curr. Opin. Struct. Biol.* **2**, 61–67.
Shane, B. (1989). *Vitam. Horm.* **45**, 263–335.
Sheng, Y., Cross, J. A., Shen, Y., Smith, C. A. & Bognar, A. L. (2002). *Arch. Biochem. Biophys.* **402**, 94–103.
Sheng, Y., Sun, X., Shen, Y., Bognar, A. L., Baker, E. N. & Smith, C. A. (2000). *J. Mol. Biol.* **302**, 425–438.
Smith, C. A. & Rayment, I. (1996*a*). *Biophys. J.* **70**, 1590–1602.
Smith, C. A. & Rayment, I. (1996*b*). *Biochemistry*, **35**, 5404–5417.
Sun, X., Bognar, A. L., Baker, E. N. & Smith, C. A. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 6647–6652.
Sun, X., Cross, J. A., Bognar, A. L., Baker, E. N. & Smith, C. A. (2001). *J. Mol. Biol.* **310**, 1067–1078.
Tan, X. J. & Carlson, H. A. (2005). *J. Med. Chem.* **48**, 7764–7772.
Toy, J. & Bognar, A. L. (1994). *Arch. Biochem. Biophys.* **314**, 344–350.
Tronrud, D. E., Ten Eyck, L. F. & Matthews, B. W. (1987). *Acta Cryst.* A**43**, 489–501.